Next-Gen Predictive Maintenancepage310 320

Article · July 2025

1 author:

Vijay Shelake
University of Mumbai

VOLUME 48 • SPECIAL ISSUE • NO. 1 • JUNE 2025

ISSN 0971-3034

THE INDIAN JOURNAL OF TECHNICAL EDUCATION

Published by

INDIAN SOCIETY FOR TECHNICAL EDUCATION

Near Katwaria Sarai, Shaheed Jeet Singh Marg,

New Delhi - 110 016

INDIAN JOURNAL OF TECHNICAL EDUCATION

Volume 48 • Special Issue • No. 1 • June 2025

INDIAN JOURNAL OF TECHNICAL EDUCATION

Published by
INDIAN SOCIETY FOR TECHNICAL EDUCATION
Near Katwaria Sarai, Shaheed Jeet Singh Marg
New Delhi - 110 016

Editorial

Greening the Globe: The increased use of clean and renewable energy technologies, especially solar, wind, and green hydrogen, is one of the most noticeable trends. India has committed to achieving 500 GW of non-fossil fuel capacity by 2030 under its Nationally Determined Contributions (NDCs). STEM (Science, Technology, Engineering, and Mathematics) innovation is driving this shift through advancements in solar photovoltaic technology, artificial intelligence (AI)-powered smart grids, and advanced battery storage systems. Notably, green hydrogen has emerged as a focal point. By using solar and wind energy to create clean hydrogen for industrial usage, the National Green Hydrogen Mission seeks to establish India as a global centre for hydrogen production. Researchers are working on cost-effective catalysts to increase production efficiency, while both energy companies and startups are investing in hydrogen electrolysers.

Artificial Intelligence (AI), Internet of Things (IoT), and data science are reshaping the sustainability landscape across sectors. AI-powered systems are being used for crop monitoring, weather prediction, energy demand forecasting, and even biodiversity conservation. In urban areas, AI is being employed to create detailed heat vulnerability maps, identifying hotspots where communities are most at risk from extreme temperatures. For example, in cities like Delhi, AI models are integrating satellite imagery and ground-level data to guide policymakers in implementing heat mitigation strategies, such as green roofing and reflective surfaces. Meanwhile, IoT-based solutions are helping to monitor air and water quality in real time, offering critical data for timely environmental interventions.

India’s Smart Cities Mission has sparked a STEM-led revolution in public infrastructure and governance in the field of urban sustainability. The mission encompasses about 8,000 projects that emphasize integrated command-and-control centers, energy-efficient structures, and sustainable mobility. To improve resource efficiency, cities are implementing sensor networks to track water leaks, waste collection, and traffic flow. Furthermore, to battle urban heat and pollution, new state-level policies—like Uttar Pradesh’s Urban Green Policy—are requiring rooftop gardens, Miyawaki woods, and vertical greening. These efforts are backed by environmental engineering, landscape science, and geospatial technologies that enable better city planning.

Despite technological improvements, challenges still exist. Scalability may be hampered by implementation flaws, a lack of capital for early-stage STEM firms, and a shortage of employees with the necessary skills for frontier technologies. However, there is a hope due to India’s innovation momentum and policy orientation. To bridge these gaps, international knowledge connections, digital infrastructure development, and public-private partnerships are essential. STEM is the key to unlocking scalable, inclusive, and long-term solutions for sustainable development in India as it develops and integrates across disciplines.

New Delhi **Editor**

30th June 2025

ATHARVA EDUCATIONAL TRUST'S

ATHARVA COLLEGE OF ENGINEERING

(Approved by AICTE, Recognized by Government of Maharashtra & Affiliated to University of Mumbai - Estd. 1999 - 2000) ISO 21001:2018 ISO 14001:2015 ISO 9001:2015, NAAC Accredited

Contents

Silicon Nitride - Barium Fluoride based VOA Suitable for C-band and L-band Wavelength

Shekhar M. Mane
Assistant Professor
Bharati Vidyapeeth college of Engineering
Navi Mubai, Maharashtra
✉ maneshekhar13@gmail.com

Rajendra B. Mohite
Assistant Professor
Bharati Vidyapeeth college of Engineering
Navi Mubai, Maharashtra
✉ mohiterajendra52@gmail.com

J. Jeyavel
Associate Professor
Department of CSE
Amity School of Engineering and Technology
Amity University
Mumbai, Maharashtra
✉ jeyavelj369@gmail.com

Zunzarrao V. Thorat
Assistant Professor
Bharati Vidyapeeth college of Engineering
Navi Mubai, Maharashtra
✉ zunzar_thorat@rediffmail.com

ABSTRACT

The domain of Photonic integrated circuits based on silicon is a fairly new one, having the aim of designing and manufacturing myriad silicon based optical devices. . A number of factors which includes the availability of a well-established fabrication infrastructure and the fairly high abundance and low cost of the Silicon substrate, this particular strategy seems to be extremely appealing [1]. Variable optical attenuators (VOA) have become an indispensable device in the field of optical networking, which are used in simulation of optical links during verification of link power budgeting and tests such as white-box testing and black-box testing, reducing the output signal intensity from laser diodes in wavelength division multiplexing (WDM). Depends on specific wavelengths, optical amplifiers are used for adjusting the magnitude of incoming WDM signal. We present a variable optical attenuator having low power depending on the configuration of a silicon nitride - Barium Fluoride (Si3N4 -BaF2) S-bend waveguides. For the development of a VOA the characteristics of the thermo-optic coefficient of BaF2 and Si3N4 materials are discussed. Based on the VOA working principle, a mathematical model was designed. The results of simulation are validated with theoretical results. An optical signal attenuation factor of up to 20 dB with a minimum value of 5 dB insertion loss has been achieved for a 10 µm length component which will be compatible with C band and L band range applications.

***KEYWORDS** : Variable optical attenuator (VOA), Electro-optic effect.*

INTRODUCTION

In the control of features such as gain control, channel energy monitoring and controlling, stabilization of cross coupled WDM nodes along with transmission system VOAs are used. The features linked to a fast response, an excellent attenuation range together with a reduced insertion loss are essential for any performance of VOA. Variable optical attenuators (VOA) have wound up being a significant optical communication device in dense wavelength division multiplexing (DWDM) systems especially for managing power requirement [2, 3]. The application location could be rapidly expanded as the primary key element of an optical transmission system. Various additional advantages of waveguide items happen to be when their integration with other optical switches and wavelength multiplexers carried out. In the past few years, a number of ways for VOAs, like micro electromechanical Systems (MEMS) [4,5], micro fluidic [8,9], liquid crystals (LC) [10,11], as well as planar light wave circuits (PLC) [10], are advised. PLC - based VOAs are better when compared with others as an outcome of the compact size of theirs, large scale integration, along with prevention of moisture and vibration. It's, in turn, quite conveniently incorporated with PLC components that are additionally used to create an advanced and sophisticated system [12,14]. Additionally, PLC technological development is especially ideal for the improvement of VOA as Thermo-Optic (TO)

[4,5] and electro-optic (EO) [17,18] properties can be assimilated into it. These optical equipments are produced with organically available polymer- based materials [13, 14, 19, 20, 21], together with inorganic elements as silica [12], silicon oxinitride (SiON), plus silicon-on-insulator (SOI) [ten, eighteen]. Generally, the inorganic materials platform provides excellent balance and also convenience, though the energy usage is on the higher side, and the naturally polymer material environments offer limited energy usage, though the performance may go down over a period of time as an outcome of the precious time connected rest of the supplies [13]. To be able to combine the benefits of the inorganic and organic material platforms with their relatable disadvantages, a combination of organic and inorganic hybrid material combination might be utilized. A VOA made from a pure polymer silicon-based energy flow region waveguide was fabricated and its performance evaluated [17]. Table 1.1 lists the characteristic performance of the VOAs mentioned above. A great deal of Applications, like telecommunications and data, additionally to in biological parameter sensing, Global Positioning system (GPS) Applications, and signal processing is provided using Photonic Integrated Circuits (PICs).

Various techniques have been eventually be recognized in implementing a VOA. However, many of them deliver excellent reading parameters and trustworthiness; VOA integrated on photonic silicon waveguides [two, three] is frequently believed being selected because the fabrication of such things calls for low and small only price changes to present fabrication processes. These silicon nitrate based technologies ensure low optical attenuation and lesser lossy waveguides that can be used with other platforms and are a replacement to existing SOI based PICs.

In this paper, we suggest a VOA according to the Si3N4 BaF2 S-bend waveguide. Operating wavelength range for Si3N4 is from 0.3 µm to 4.6 µm and for BaF2 id from 0.14 µm to 13 µm. Thermal conductivity for Si3N4 is 33 W/m K and for BaF2 is 12 W/m K. In this particular framework, Si3N4 layer is needed as the core and in addition BaF2 layer can be used as cladding over the waveguide. The various signal weakening outcomes is attained by getting the Thermo-Optic (TO) outcome that can be utilized leading to the distant relative RI distinction among the Si3N4 facility and the BaF2 shell. The difference in Refractive Index (RI) of the inner core and outer cladding changes as temperature varies; consequently the attenuation is realized [23]. Since BaF2 having negative TO (dn⁄dT ~ −10^-6) [21], and Si3N4 is having positive TO (dn⁄dT ~ 10^-5) [22], with this VOA setup applied external voltage will lead to even bigger change within the distant relative RI distinction as well as, consequently, additional signal strength reduction. An optical signal weakening also called as attenuation of minimum 5 and a maximum of 20 dB may be accomplished in software simulations. Mathematical model had been made based and simulation based outputs are compared and evaluated with theoretical results.

LITERATURE SURVEY

The polymer material used in optical communication devices have superior thermo-optic effect and based on this finding, an innovative and novel low-loss fluorinated polymer variable optical attenuator (VOA) with 1.0 dB insertion loss, less absorption loss at 1550nm and also operating power of less than 30 mW is demonstrated [21]. For the purpose of practical attenuation, the characteristics of a device with a control circuit offering a closed-loop feedback are evaluated. These characteristics include a low range wavelength and a bare minimal temperature dependence factor.

A polymer/silica based hybrid waveguide thermo–optic attenuator operating with a multimode interference (MMI) coupler that operates at 650 nm is designed and constructed in paper [24]. The optimal heater placement angle is determined by thermal analysis based on material properties. At 650 nm, the manufactured VOA has a 26.5 dB attenuation with a 21 mW electrical input power. The average rise time was found to be 51.99 seconds, and the average fall time is 192 seconds. The results of the time–stability measurement demonstrate that it works reliably.

The Long-range surface plasmon polariton (LRSPP) waveguide-based thermal VOAs with multimode interference architecture were constructed and characterized [25]. Atomic force microscopy was used to investigate the gold (Au) stripe's surface morphology and waveguide configuration. The process of end-fire excitation at 1550nm using the LRSPP mode guiding along with the Au stripe indicates that approximately 12 dB extinction ratio of attenuator at a signal driving power of 69 mW was found. It was confirmed that the fluctuation of poly (methylmethacrylate-glycidyl-methacrylate) polymerised cladding was found to be not more than 3. These good properties guarantee possibilities of this device in the utilization of optically interlocked mechanism.

A VOA based on graphene that operates at an 855 nm wavelength and is integrated onto a photonic Si3N4 waveguide is constructed and characterized in the paper [26]. For a device length of 700 m, a maximum attenuation of 17 dB is achieved at gate voltages of -3 V.

A variable optical attenuator (VOA) based on opto-fluidics that uses a narrowing, uni- mode optical fiber attached with polished side to an electro wetting-on-dielectric (EWOD) platform is described in a paper [27]. An opto-fluidic continuous-fiber VOA typically provides a dependent loss based on wavelength of less than 1.1 dB and a maximum range of 26 dB of broadband attenuation in the 1550 nm transmission window. Opto-fluidics enables the miniaturization and integration of optical devices, as well as the low-cost production of low-power devices by various methods. Using optofluidics, only a small number of fiber optic VOAs have been presented to date. The fluidic VOA has been actuated in a variety of different ways among the devices that have been reported.

On a Silicon-On-Insulator (SOI) wafer, a thermo-optical VOA comprising of three cascaded three-waveguide directional couplers was developed [28]. At the wavelength of 1550 nm, thermo-optical analysis for static and dynamic conditions, and optical propagation characteristic simulation were presented. A RI variation of was used to achieve the range of 0.14 dB to 50 DB (the corresponding temperature variation: 0 to 30 ℃) suitable for dynamic attenuation. The Si/Al2O3/Si structure was blamed for the 5-second response time. As the fundamental component, the device could be easily integrated into silicon optoelectronic circuits to create multi-channel VOA.

This paper [29] proposes a low-power VOA with dual grooves on either sides of the S-bend waveguide core and a hybrid SiON-polymer waveguide configuration, which has S-bend structure. The contradictory properties of SiON and polymer materials, having thermo-optic charactersictics can be fully utilized in this configuration. As a result, the efficiency with which heat is used is improved. An Opto-fluidic chip of the uni-mode fiber VOA is projected [30] to realize an optical attenuation of 50 dB at a 3.6 mW applied electrical power. A fabricated device that is available in the market usually has the dimension of 8 mm length and 46 dB maximum optical attenuation with an 16.2 mW of electrical power and 5.4 dB as insertion loss. A typical chip uses a combination of microfluid and air to perform optical attenuation, with air acting as a stopper, and has a straightforward structure. At 1310 and 1550 nm, the uni-mode characteristics of the VOA chip are discussed. The results of the experiment show that the proposed chip can be used to control optical attenuation. This chip's VOA has a wide operating waveband and a high optical attenuation range (82 dB), ranging from the visible to near-infrared wavelengths.

In this paper [31], two strategies for acquiring a ultra-smaller variable all-optical attenuator in light of multimode obstruction (MMI) utilizing silicon waveguides are proposed. Utilizing two connected silicon waveguides is the first approach. Utilizing a multimode interference coupler is the second approach. A variable attenuator can be achieved by creating a surface outline over the MMI region at specific points of linking waveguides.

Using an X-patterned photonic crystal ring resonator, a simulation model of the C-band VoA is proposed [32]. Variation of the refractive index is how tuning is done. For 0.0022 changes in refractive index, a 30 dB attenuation is reached.

We propose and numerically investigate a novel graphene plasmonic gap waveguide structure that serves as both a variable optical attenuator and modulator [33]. By adjusting the gate supply voltage, the Propagation and attenuation mode of the structure can be achieved. The outcome demonstrates that it is capable of achieving an operational range of frequencies more than 8 THz for both the VOA and the modulator, as well as a high modulation depth per micrometer. A VOA based on Dynamic Scattering Mode (DSM) in ion-doped liquid crystals with negative dielectric anisotropy is demonstrated in Paper [34]. The instrument of weakening comes from optical dissipating, which is created by the electrically instigated precariousness of undulation of LC surfaces. The VOA can also selectively block a specific wavelength of mid-IR light while allowing other wavelengths of light to pass through. The proposed VOA is expected to have spectral discrimination, insensitivity to polarization and high on/off difference are just a few of the superior optical switching properties of such a VOA; consequently, it has the potential to be incorporated into real-world optical systems.

SILICON PLATFORMS

Silicon photonics typically uses silicon on insulator that guarantees a high-index-contrast system. On the other hand, an alternative system with average-index-contrast can be produced in the same Complementary Metal oxide semiconductor (CMOS) environment using the silicon nitride- based silicon photonic device. The relative advantages of the two platforms are discussed in this paper [35].

When the primary characteristics of a SiN-based PIC platform with those of a SOI-based PIC platform is compared based on the range of transparency, it has been found that the waveguides fabricated using SOI have low assimilation issues in the frequency range from 1.1 µm to around 3.7 µm. When it is used in applications that require shorter wavelengths (such as sensors operating in the therapeutic window, data communication at 850 nm, and so on), it is not possible to be implemented using SOI. Therefore, SiN is a viable candidate for the implementation of "silicon photonics" at wavelengths below 1.1 m because it is capable to be operated in the majority of the visible spectrum [35]. This has apted showings of spectroscopic capabilities [36-37], spectroscopy-on-chip capabilities proposed by Dr. C.V.Raman [38-39] and reconciliation with colloidal quantum specks radiating in the noticeable list. In case of contrast the fact that the SiN system has a lower index contrast (2 versus 1.5) than the SOI system (3.5 versus 1.5) because. This fact suggests that it is more challenging to produce for out-of-plane optical input and output highly efficient grating couplers. This can be a very useful feature that enables alignment-tolerant fiber coupling and wafer-level testing [40]. However, there have been reports of grating coupler efficiencies below 3 dB. In case of damage, typically for SOI-based photonic wires, silicon strip-based waveguide losses are of 1 to 2 dB/cm, which are completely surrounded by a cladding made of silica material. This is largely due to scattering losses caused by rough sidewalls. These SiN waveguides have been shown to suffer losses as low as 1 dB/m [41].

Modularity in manufacturing: Because SiN can be deposited using either LPCVD or PECVD the SiN waveguide is more compatible with photonic structures other than SOI.

Nonlinearity of the third order: Despite silicon's enormous Kerr nonlinearity, two-photon absorption (TPA) renders it useless in the 1300/1550 telecom band. Because at high power, it causes additional waveguide losses, TPA itself is a problem in and of itself. Despite the material's large bandgap and weaker Kerr nonlinearity, silicon nitride has a virtually zero TPA. As we increase the SOI-based device dimensions to prolong the operational wavelength to the 2-4 um also called the Mid-infrared (MIR) range has now become emerging field of scientific exploration [14–17] due to the numerous prospective application areas such as lab-on-chip sensors and open space communications. Numerous platforms and geometric structures such as were studied and demonstrated to achieve low propagation loss.

DESIGN AND SIMULATION

Utilizing OptiFDTD Software, we have developed a number of configurations of Ridge and Planar waveguides, using various materials as Silicon, Titanium oxide, Lithium nitrate, Silicon Nitride, Silica, etc. The modes were work out by using vector (full or semi) techniques. Using the technique proposed by Solver, the techniques or modes inside the waveguides are based on the Alternative Direction Implicit (ADI) technique. To resolve matrices proposed by Sylvester, ADI which is an iterative strategy used. To identify the fundamental mode, a vector modal analysis was carried out. We chose Si3N4 as the waveguide material since the planar waveguides dependent on Si3N4 result in insertion loss as well as propagation loss.

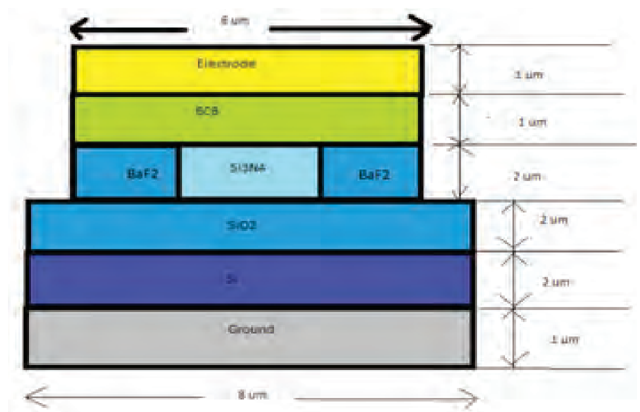


Fig 1 (a) Si_3N_4 core with BaF_2 Cladding waveguide cross-section

The constructed S-bend Si3N4 VOA provides a single method when a polarized light of 1550nm TE is triggered through the device structure through optimum waveguide-based coupling.

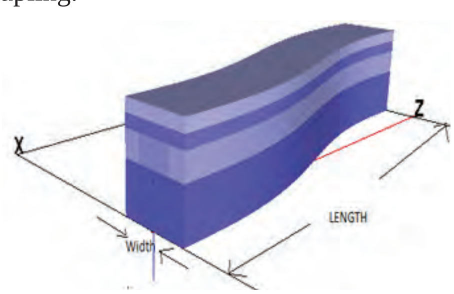


Fig. 1 (b) S-bend VOA

As can be seen in the figure 1a and 1b, S - bend shaped waveguide with Si3N4-BaF2 configuration. A single mode is observed during the simulation. We have observed the effective modality index profile and the RI profile.

ANALYSIS

The attenuation or signal weakening of our VOA is determined and denoted as [42]

$$\propto (dB) = \frac{20\pi\sqrt{2}\, hC_1}{C_2\, l\, \ln 10}\, e^{-\gamma}\left(1 - e^{-\frac{\gamma}{2}}\right) \tag{1}$$

Where $\gamma = \frac{C_2 l^2}{2\pi h}$

In which C_1 and C_2 can be expressed as

$$C_1 = \frac{\varepsilon_1}{2Z_c\varepsilon_2} \tag{2}$$

$$C_2 = \frac{2\pi[2(N_{eff} - n_2)]^{3/2}}{\lambda\sqrt{2}} \tag{3}$$

$$Z_c = \frac{n_2}{2\lambda}\left[a + \frac{2}{k_0\sqrt{(N_{eff}^2 - n2^2)}}\cos\left(\frac{k_x a}{2}\right)\right]^2 \tag{4}$$

$$\varepsilon_1 = \frac{a}{2} + \frac{1}{2k_x}\left[\sin k_x a + \frac{2}{k_0\sqrt{(N_{eff}^2 - n_2^2)}} * \cos^2\left(\frac{k_x a}{2}\right)\right]^2 \tag{5}$$

$$\varepsilon_2 = \frac{1}{2k_0\sqrt{(N_{eff}^2 - n_2^2)}} * \cos^2\left(\frac{k_x a}{2}\right)\exp\left(a k_0\sqrt{(N_{eff}^2 - n_2^2)}\right) \tag{6}$$

$$k_x = k_0\sqrt{(N_1^2 - N_{eff}^2)}, \tag{7}$$

Where K_0 is free space propagation constant, Neff may be the real RI of the S-bend N1 and waveguide may be the effective RI of the comparable waveguide. The table 1. Below gives the RI of various components at 1550nm.

Table 1: Refractive Index (RI) values

S. No	Component	Refractive index
1.	Core Si_3N_4(n1)	1.9904
2.	SiO_2	1.444
3.	Si	3.4757
4.	BaF_2(n2)	1.4757
5.	TiO_2	2.584
6.	Benzocyclobutene (BCB)	1.5254
7.	Air	1

Keeping the width of center region is 4µm and the thickness of center is 2 µm so that single mode can be observed for different Δ12. We have taken the thickness of BCB as 2µm and the denseness of electrode heater is 1µm. Electricity is provided to the electrode. Due to the voltage given from the electric energy source, the RI of n1 and n2 undergo changes due to electro-optic effect and also temperature of material changes. Therefore, the end result of the modification is the signal attenuation.

Table 2. Summary of simulation results

Sr. No.	Structure	Perpendicular offset (um)	Length (um)	Output Amplitude	Attenuation (dB)
1	S-bend Ridge waveguide Si_3N_4 Core	2	100	0.12	18.41
		1.5	100	0.38	8.4
		1	100	0.71	2.97
		2	50	0.07	21.93
		1.5	50	0.67	3.47
		1	50	1.01	-0.08
2	S-bend Si_3N_4 Core with SiO2	2	100	0.84	1.51
		1.5	100	0.63	4.013
		1	100	0.32	9.89
		2	50	0.61	4.29
		1.5	50	0.66	3.6
		1	50	0.99	0.08
3	S-bend Si_3N_4 Core with BaF_2 side grooves	1.5	20	1.18	-1.43
		1.5	100	0.65	3.74
		1.5	940	0.19	14.42
4	Si_3N_4- BaF_2 straight	-	22	0.48	6.37

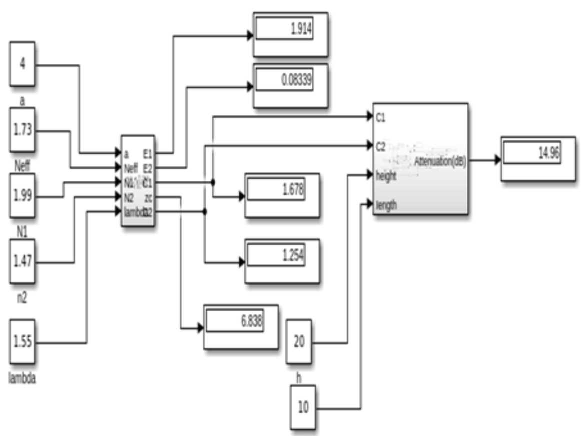


Fig. 2 Mathematical model of VOA

By putting to use the electrical energy in opto-couplers, we can control the strength of the signal towards the end. We were able to develop a VOA in this way using the

Si3N4-BaF2 configuration. Furthermore, using Opti-BPM software, analysis and performance of VOA is done [43]. It was observed that a very less amount of propagation loss is attained without using applied voltage. A radiation loss in the range of 10-20dB has been observed, while the light travels through the waveguide whenever external voltage is provided. Based on equations 1 to 6, as described in fig.2 a mathematical model was created Based on the TO coefficients of the material, a modification in the RI with respect to temperature variations has been observed, as shown in fig. 3.

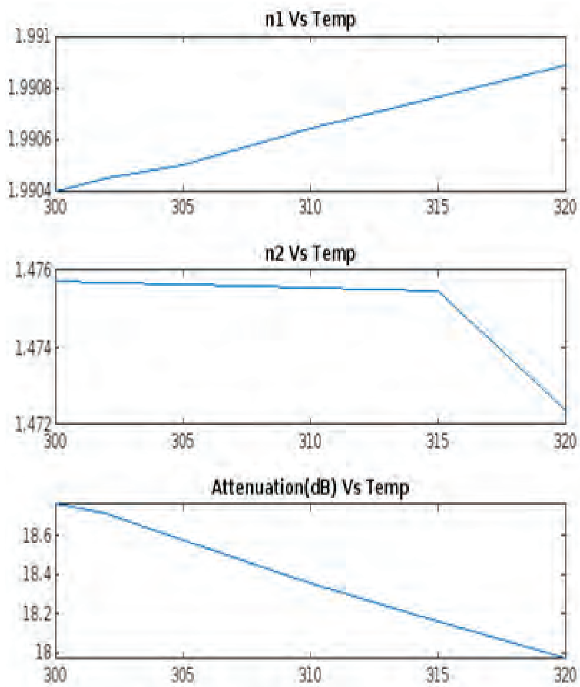


Fig. 3 Change in RI and hence Attenuation w.r.t temperature

RESULTS AND DISCUSSION

From the previous section, it can be seen that various configurations of waveguides were designed with materials ranging from Silica and Silicon Nitrate to Titanium oxide were selected using OptiFDTD Software By means of full or semi-vectorial techniques, the operative modes have been computed. It has been observed that as attenuation increases with value of vertical offset length for device lengths 50µm and 100µm, as shown in fig.4. Utilizing Effective index and FDTD method, the simulations of several VOAs were conducted. The amplitude of signal for the lengths of 50µm and 100 µm with vertical offset from a minimum of 1µm, up to 2µm for S- bend shaped Si3N4 core containing VOA was simulated. Results will be used for waveguide coupling is presented in Table.1. Based on the comparative analysis of the signal weakening levels in various configurations, we can select the preferred VOA for Free space Optics.

From fig. 4, it has been discovered that the amplitude on the papers moves on boosting as vertical offset is increased case of 50µm length but there is an output occurs in a reverse way with respect to 100 µm length shown in fig.4. From this we are competent to assess amount of signal weakening in various configurations additionally we are in a position to choose the preferred VOA for free space Optics.

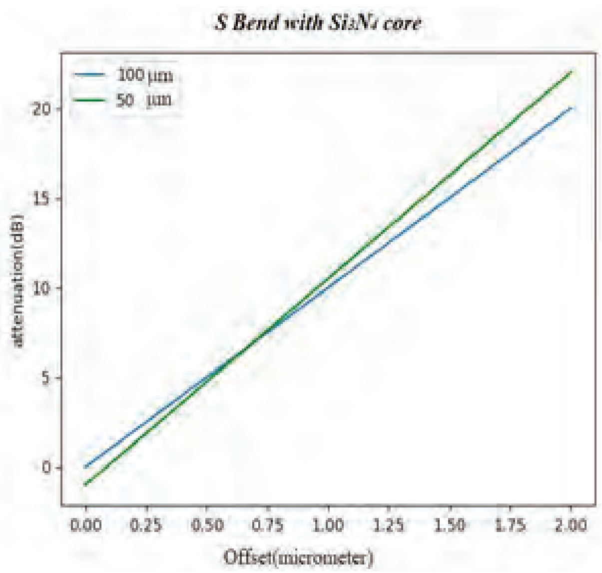


Fig. 4 Results of S-bend waveguide Si3N4 Core, Perpendicular offset bend (µm) vs. loss (dB) for 50µm and 100 µm length of device

We assume that Δ12 is the distant relative RI distinction between the Si3N4 -BaF2 cladding respectively. Whenever the temperature of electrode increases in place by utilizing the electric power source, as the temperature steadily rises, the RI of BaF2 reduces as a result of its negative TO coefficient, and also the RI of Si3N4 increases on account of its useful TO coefficient, as a result there is a reduction in the modal analysis which causes a significant increase in the bending loss. In mathematical model, keeping a, h, l parameters of VOA constant, we can change the value of lambda and we can observe attenuation in VOA for C-band and L band ranges. For temperature range 300K to 320 K, we get the attenuation of 1dB for desired operating wavelength as per the mathematical model.

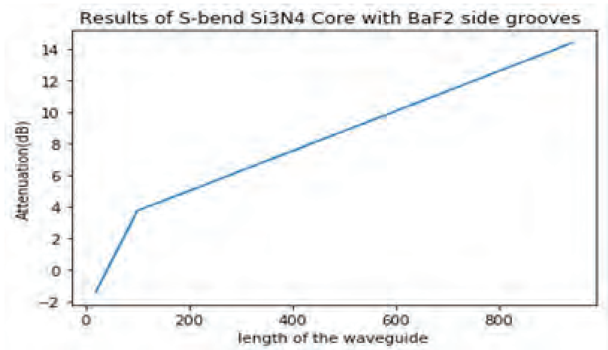


Fig. 5 Device length (µm) vs. Attenuation Loss (dB) for S-bend Si_3N_4 Core with BaF_2 side grooves

For C band range, proposed VOA gives attenuation of 14dB to 15 dB variation with respect to changes in temperature of material from 300K to 320K as shown in table2. And for L band range, proposed VOA gives attenuation of 16 dB to 17 dB for same temperature variation. The device that we propose to implement, with a length of 10 µm, reveals a characteristic attenuation of 14 to 17 dB.

Table 2. Observations for wavelength lambda=1550 nm.

S. No	Change in Temperature	n1	n2	Attenuation dB
1.	300	1.9904	1.4757	15.37
2.	305	1.9905	1.4756	15.23
3.	310	1.9906	1.4755	15.09
4.	320	1.9908	1.4753	14.81

Table 3. Observations for wavelength lambda=1625 nm.

S. No	Change in Temperature	n1	n2	Attenuation dB
1.	300	1.9904	1.4757	17.02
2.	305	1.9905	1.4756	16.87
3.	310	1.9906	1.4755	16.71
4.	320	1.9908	1.4753	16.41

APPLICATIONS AND FUTURE ASPECTS

Silicon photonics being new technology satisfies the requirements for low power consumption, speed and efficiency at affordable prices. In the silicon–based technology, a number of opto-electronic devices are integrated onto a single substrate and linked to one another via very fine waveguides. These opto-electronic circuits may be used to set up transmission at a high speed, increase bandwidth, use less power, and solve problems with latency. It might possibly expand the transfer speed limit by giving small size, super low power gadgets. Optical data transmission reduces electromagnetic interference and speeds up data transfer. Based on these information rates, it can be envisioned that videoconferencing with a high goal that the entertainers or relatives have all the earmarks of being in the room with you. Data can be transferred more quickly and over longer distances using optical links than copper technology does today; 50 Gbps is possible. Intel Corporation recently introduced its 100Gbps optical link connection [27]. For optical trans-receivers, well known company named Fujitsu Laboratories recently pioneered a 4 wavelength integrated silicon laser. The innovative optic-interconnection of external devices to PCs is currently on the anvil. In 2009, Intel introduced Light Peak, a replacement for USB that transmits data at speeds of up to 10Gbps. Researchers' advancements in silicon nano-photonic technology lead to novel concepts for the architectures of future computing systems. Waveguide diodes, Optical Isolator and Circulator based on mirroring techniques and Silicon Photonics-based Optical Trans-receiver are just a few silicon optical devices that have been developed with great success by researchers.

CONCLUSION

We have developed and examined different waveguides based upon silicon nitride which are required for fabrication of VOA. Based on our findings, S-bend Si3N4 - BaF2 is selected among almost all configurations. Latter is grooved and attenuation can be handled. In Si3N4-BaF2 configuration, un-related and different thermos-optic characteristics of resources are entirely implemented and various attenuation readings are performed by electro optic effect. . For temperature range 300K to 320 K, we get the attenuation of 1dB for desired operating wavelength as per the mathematical model. For C band and L band range, proposed VOA gives attenuation 14 dB to 17 dB with respect to change in temperature variation. The device that we propose to implement, with a length of 10 µm, reveals a characteristic attenuation of 14 to 17 dB. Si3N4 and BaF2 both materials are compatible for C band and L band wavelength ranges. While undertaking the fabrication of this specific VOA design, we can use this unit in many VOA applications with wide range of wavelengths of C band and L band.

REFERENCES

1. Ashish Diman,"Silicon Photonics: A Review", e-ISSN: 2278-44861, Volume 3, Issue 5 (Mar. - Apr. 2013), PP 67-79

2. M. Lenzi, S. Tebaldini, D. D. Mola, S. Brunazzi, and L. Cibinetto, “Power control in the photonic domain based on integrated arrays of optical variable attenuators in glass-on- silicon technology,” IEEE J. Sel. Top. Quantum Electron. 5, pp.1289–1297 (1999).

3. H. Uetsuka, T. Hasegawa, and M. Ohkawa, “Variable optical attenuators combined with an arrayed waveguide grating filter for next-generation WDM system,” Hitachi Cable Rev. 20, pp15–18 (2001).

4. T. S. Lim, C. H. Ji, C. H. Oh, H. Kwon, Y. Yee, and J. U. Bu, “Electrostatic MEMS variable optical attenuator with rotating folded micromirror,” IEEE J. Sel. Top. Quantum Electron. 10, pp558–562 (2004).

5. D. Noraica, M. Emmanuelle, and S. Nelson, “Electronically variable optical attenuator enabled by self-sensing in vanadium dioxide,”IEEE Photon. Technol. Lett. 26, pp.1011–1014 (2014).

6. M. I. Lapsley, S. C. S. Lin, X. L. Mao, and T. J. Huang, “An in plane, variable optical attenuator using a fluid-based tunable reflective interface,” Appl. Phys. Lett. 95, 083507(2009).

7. D. Anna, B. Robert, Z. Michele, S. George, and U. Deepak, “Modeling and characterization of an electrowetting-based single-mode fiber variable optical attenuator,” IEEE J. Sel. Top. Quantum Electron. 21, 4500209 (2015).

8. G. Zhu, B. Y. Wei, L. Y. Shi, X. W. Lin, W. Hu, Z. D. Huang, and Y. Q. Lu, “A fast response variable optical attenuator based on blue phase liquid crystal,” Opt. Express 21,pp 5332–5337 (2013).

9. M. Blasl, K. Bornhorst, and F. Costache, “Polarization insensitive variable optical attenuator based on field induced waveguides with a liquid crystal core,” in Proceeedings of the Optical Fiber Communication Conference (OSA, 2015), paper W4A.2.

10. Q. Fang, J. F. Song, G. Zhang, M. B. Yu, Y. L. Liu, G. Q. Lo, and D.-L. Kwong, “Monolithic integration of a multiplexer-demultiplexer with a thermo-optic VOA array on an SOI platform,” IEEE Photon. Technol. Lett. 21, pp.319–321 (2009).

11. S. Tsunashima, F. Nakajima, Y. Nasu, R. Kasahara, Y Nakanishi, T. Saida, T. Yamada, K. Sano, T. Hashimoto, H. Fukuyama, H. Nosaka, and K. Murata, “Silica based, compact and variable-optical attenuator integrated coherent receiver with stable optoelectronic coupling system,” Opt. Express 20, pp. 27174–27179 (2012).

12. R. Patrick, S. Stefan, S. Angela, J. Klemens, S. Jens, T. Dirk, and L. N. Mads, “Monolithic InP receiver chip with a variable optical attenuator for colorless WDM detection,” IEEE Photon. Technol. Lett. 26,pp.349–351, (2014).

13. J. Sun, X. B. Wang, J. W. Sun, X. Q. Sun, C. M. Chen, F. Wang, and D. M. Zhang, “Fabrication of polymer optical attenuator with ultra-low power consumption,” Mod. Phys. Lett. B 28, pp.50113–50120, (2014).

14. M.-N. Alejandro, Z. Y. Zhang, I. Gelani, P. Andrzej, M. Tim, D. F. David, K. Moritz, B. Walter, Z. Crispin, and K. Norbert, “Thermally optimized variable optical attenuators on a polymer platform,” Appl.Opt. 54, pp 569–575 (2015).

15. X. Jiang, X. Li, H. Zhou, J. Yang, M. Wang, Y. Wu, and S. Ishikawa, “Compact variable optical attenuator based on multimode interference coupler,” IEEE Photon. Technol. Lett. 17, pp2361–2363 (2005).

16. Z.-L. Liao and R. W. Chuang, “1× 3 silicon oxynitride tunable optical waveguide attenuators based on the multimode interference effect,”Jpn. J. Appl. Phys. 48, 04C118 (2009).

17. P. P. Sahu, “Variable optical attenuator using thermo-optic twomode interference device with fast response time,” Appl. Opt. 48, 4213–4218 (2009).

18. P. F. Qu, W. Y. Chen, F. M. Li, C. X. Liu, and W. Dong, “Analysis and design of thermo-optical variable optical attenuator using three waveguide directional couplers based on SOI,” Opt. Express 16, 20334–20344, (2008).

19. N. Ghada, W. F. Ho, and H. P. Chan, “Experimental study on the performance of a variable optical attenuator using polymer dispersed liquid crystal,” Appl. Opt. 52, E15–E21 (2013).

20. J. S. Li and J. R. Li, “Variable optical attenuator based on a bent polymer/SION waveguide,” Proc. SPIE 6781, 678158 (2007).

21. Young-Ouk Noh, Chul-Hee Lee et.al”Polymer waveguide Communications 242 (2004) 533–540 [22] Marvin J. Weber Handbook of optica Materials ,2003,CRC \ Press LLC.

23. LINGFANG WANG, QIANQIAN SONG et.al” Low power variable optical attenuator based on a hybrid SiON-polymer S-bend waveguide” Vol. 55, No. 5 / February 10 2016 /Applied Optics

24. Yue-Yang Yu, Xiao-Qiang Sun et al” The 650-nm variable optical attenuator based on Polymer/silica hybrid waveguide,Chin. Phys. B Vol. 25, No. 5 (2016) 054101

25. Xiaoqiang Sun, Ying Xie, Tong Liu,”Variable Optical Attenuator Based on Long- Range Surface Plasmon Polariton Multimode Interference Coupler”, Research article published on Published 22 April 2014.

26. Muhammad Mohsin, et al.,"Graphene based on-chip variable optical attenuator operating at 855 nm wavelength", Vol. 25, No. 25,11 Dec 2017, Optics Express 31660

27. A. Duduś, R. Blue, Member, IEEE, M. Zagnoni, G Stewart and D. Uttamchandani, Senior Member, IEEE,"Modeling and characterization of an electrowetting based single mode fiber variable optical attenuator",1077-260X (c) 2013 IEEE, 10.1109/JSTQE.2014.2382298, IEEE Journal of Selected Topics in Quantum Electronics.

28. Pengfei Qu, Weiyou Chen, Fumin Li, Caixia Liu, Wei Dong,"Analysis and design of thermo-optical variable optical attenuator using three-waveguide directional couplers based on SOI",8 December 2008 / Vol. 16, No. 25 / OPTICS EXPRESS 20334.

29. Lingfang Wang, Qianqian Song, Jieyun Wu and Kaixin Chen," Low- power variable optical attenuator based on a hybrid SiON–polymer S-bend waveguide", Vol. 55, No. 5 / February 10 2016 / Applied Optics

30. Jing Wan, Fenglan Xue, Boyu Chen, Han Cao, "Optofluidic Chip of a Single-Mode Fiber Variable", Vol. 9, No. 2, April 2017,2017 IEEE

31. Trung-Thanh Le," Ultra-Compact Variable All-optical Attenuator Based on Multimode Interference Couplers on an SOI Platform", IACSIT International Journal of Engineering and Technology, Vol. 4, No. 3, June 2012.

32. U. Poorna Lakshmi, D. M. S. Gautham, V. Vishnuvardhan Iyer," Variable Optical Attenuator using X-Shaped Photonic Crystal Ring Resonator", The International Conference on Fiber Optics and Photonics 2016 © OSA 2016.

33. Feng chao ni et al.,"Variable optical attenuator and modulator based on a graphene plasmonic gap waveguide" Optics Communications 426 (2018) pp 251–256,Elsevier publication.

34. Sheng-Ping Chiang et al,"Selective variable optical attenuator for visible and mid-Infrared wavelengths",Vol. 26, No. 13, 25 Jun 2018 ,OPTICS EXPRESS 17009.

35. Roel Baets, Ananth Z. Subramanian, "Silicon Photonics: Silicon Nitried vs Silicon-on-Insulator", OSA2016.

36. R. Heideman et al, Large-scale integrated optics using TriPleXTM waveguide technology: from UV to IR, Proceedings of SPIE, 7221, p.72210R-1-14 (2009).

37. A. Z. Subramanian et al, Silicon and silicon nitride photonic circuits for spectroscopic sensing on-a-chip Photonics Research, 5(3), p.B47 (2015).

38. D. Martens et al, Compact Silicon Nitride Arrayed Waveguide Grating for Very Near-infrared Wavelengths, Photonics Technology Letters, 27(2), p137-140 (2015).

39. A. Dhakal et al, Efficiency of evanescent excitation and collection of spontaneous Raman scattering near high index contrast channel waveguides, Optics Express, 23(21), p.27391-27404 (2015).

40. A. Dhakal et al, Evanescent excitation and collection of spontaneous Raman spectra using silicon nitride nanophotonic waveguides, Optics Letters, 39(13), p.4025-4028 (2014).

41. A. Z. Subramanian et al, Near infrared grating couplers for silicon nitride photonic wires, Photonics Technology Letters, 24(19), p.1700- 1703 (2012).

42. J. M. William, K. K. Steven, and A. C. Rod, "Low-loss Ti: LiNbO3 waveguide bends at λ = 1.3 μm," IEEE J. Quantum Electron.18,1802–1806 (1982).

43. OptiBPM Technical Background and Tutorials PDF, version 13.1.

Survey of Electromyogram based Emotion Recognition

S. A. Mithbavkar
Department of Computer Engineering
Datta Meghe College of Engineering
Navi Mumbai, Maharashtra
✉ shraddha0306@gmail.com

M. S. Shah
Dept. of Electronics and Telecommunication Engg.
Fr. C. Rodrigues Institute of Technology
Navi Mumbai, Maharashtra
✉ milind.shah@fcrit.ac.in

ABSTRACT

Emotions have a strong influence on our daily life. Video and image processing can be used to detect emotions but light and camera alignment restrict its performance. Electromyogram (EMG) based emotion recognition can overcome this problem. EMG signals are acquired either in ceasing or in unceasing expression form. This paper provides detailed research on both these methods' strengths and challenges. This paper gives a critical investigation of EMG based emotion recognition concerning the number of participants, number of EMG channels, and emotions.

***KEYWORDS** : EMG, Emotion recognition, EMG acquisition unit, Neural network, Classifier.*

INTRODUCTION

Emotion is a strong feeling deriving from one circumstance and temper [1]. In emotion recognition systems, researchers detect emotions mainly from visual, audio, body pose, physiological signals, and text. Many authors used audio mode for emotion recognition [2-10]. Li et al. [2] recognized seven emotions with 74.28% classification accuracy. Xu et al. [3] recognized four emotions with 76.18% classification accuracy. Fahad et al. [7] recognized four emotions with 65.93% classification accuracy. Wani et al. [8] recognized four emotions with 87.8% classification accuracy. Mustaqeem et al. [9] recognized four emotions with 77% classification accuracy. Kerkeni et al. [10] recognized six basic emotions with 94% classification accuracy. This audio mode of emotion recognition is only applicable to verbal communication. All these authors used speech processing to detect emotions with 65% to 94% classification accuracy [7, 10]. This emotion classification is affected by the noisy condition.

In nonverbal communication, researchers detect emotions using body pose and physiological signals. For visual and body-pose conditions, researchers used video and image processing techniques to identify the emotion. Authors recognized emotions with 79.8% to 95% accuracy [11, 12]. In this research work, a camera must be held in front of the face. Proper lighting conditions and background are required to detect facial expressions. As shown in Fig. 1, a person may cover their face while showing emotions. In this case, it is difficult to identify emotions using image processing or video processing.

Researchers also recognized emotion or state of mind using physiological signals such as Electrocardiogram (ECG), Blood pressure, Electroencephalogram (EEG), Electrodermal Activity (EDA), EMG, and Respiration signal [16].

Fig. 1 Pictures of persons covering their faces while showing their emotions [13-15]

Many authors used EEG signals to find emotional status, namely arousal and valence or positive and negative. Li et al. [17] recognized positive and negative emotions using EEG signals at 73.10%, and the same author recognized valence at 72.06% and arousal emotions at 74.12% using EEG signals [18]. Liu et al. [19] recognized valence with 69.9% and arousal emotions with 71.2% using EEG signals. However, in these researches, emotion recognition like disgust, sadness, fear, and surprise has not been done.

Song et al. [20] recognized joy, anger, fun, fear, sadness, neutrality, and disgust using EEG and Galvanic Skin Response (GSR) signals with 40.64% classification accuracy. The same author achieved 38.74% classification accuracy using EEG signals. Song et al. [20] recognized positive, negative, and neutral with 75% classification accuracy using EEG and GSR signals and 72% classification accuracy using EEG signals. In this research, the author recognized seven emotions with low classification accuracy and basic emotions like positive, negative, and neural with comparatively high classification accuracy.

Many researchers used only ECG signals for emotion recognition and recognized valence and arousal emotions with 83.55% [21, 22]. Many researchers used more physiological signals like EEG, GSR, BVP, and EMG for emotion recognition [23-27]. In this research, the placement of multiple sensors on the human body makes the subject inconvenient.

Out of the physiological signals discussed above, EMG signals only relate to muscle movement [28]. As shown in Fig 2, EMG signals are collected while expressing emotions. EMG signal is then filtered, rectified, and segmented in the signal processing stage. Different time or frequency domain features are selected in the feature selection stage. Different neural networks or machine learning algorithms are used to classify emotions. Further sections highlighted some important methods and concepts of EMG based emotion recognition.

EMG signals are acquired either in ceasing or in unceasing expression form. This section gives detailed research work on both these methods' strengths and challenges.

Fig. 2 Process of Electromyogram based emotion recognition

Section 2 describes EMG-based emotion recognition when a person expresses emotion in ceasing expression form. Section 3 provides EMG-based emotion recognition when a person expresses emotion in unceasing expression form. Section 4 covers a critical investigation of EMG based emotion recognition.

EMG BASED EMOTION RECOGNITION: WHEN A PERSON EXPRESSES EMOTION IN CEASING EXPRESSION FORM

Fig. 3 shows different facial muscles. Ang et al. [30] acquired EMG signals from the masseter, levator, and corrugators' supercilii from six subjects. All subjects expressed sad, angry, and happy emotions and held those emotions for 2 seconds. The author achieved 94.44% classification accuracy by using the Minimum Distance Classifier (MDC) and power density spectrum. Gibert et al. [31] acquired EMG signals from eight muscles namely, levator labii, venter frontalis, corrugators' supercilii, orbicularis oculi, depressor anguli oris, massete, mentalis, and zygomaticus-major of a single subject. The author recognized six emotions: surprise, disgust, happiness, anger, sadness, and neutral with 91% accuracy. The author used Gaussian models for six emotions.

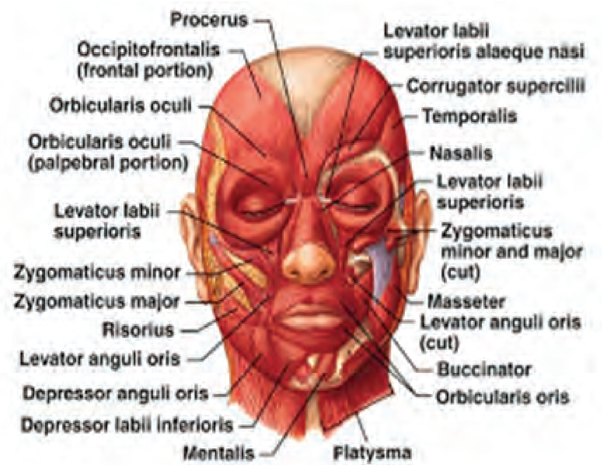


Fig. 3 Facial muscles [29]

Gruebler et al. [32, 33] recognized glowering, smiling, biting, neutral face, concurrent biting, and smiling expressions, in ceased expressions from ten subjects. The author achieved 87.52% for the test set and 91.04% for the training set by using Mean Absolute Value (MAV) and Artificial Neural Network (ANN). Chen et al. [34] used a specially designed headband for recording the EMG signals from the frontalis and corrugators supercilii muscles of six subjects. The author recognized sadness, anger, fear, surprise, and disgust. The author extracted the variance, mean, and maximum values of Root Mean Square (RMS) features from EMG signals. The author achieved 95.24% recognition accuracy using a Back propagation neural network and 96.12% recognition accuracy Elman neural network (ENN).

Xu et al. [35] recognized three emotions, namely, neutral, sadness, and fear from twenty-seven subjects using three-channel EMG signals. The author developed a novel spatiotemporal deep forest (STDF) model that processed 2D images of EMG and classified them. The author achieved 97.41% classification accuracy.

EMG based emotion recognition in ceased expression achieved classification accuracy in the 91% to 97% range [31, 34, 35]. In the ceased way, a muscle is in a constant state. The features extracted during the constant state can be classified more correctly than features extracted during the continuous state. In the case of a dancer or athlete's emotional recognition, the person shows expression continuously. Hereafter, it is needed to consider EMG-based emotion recognition unceasingly.

III. EMG BASED EMOTION RECOGNITION: WHEN A PERSON EXPRESSES EMOTION IN UNCEASING EXPRESSION FORM

Picard et al. [36] used the MIT dataset which included 20 EMG signals of hate, anger, reverence, romantic love, no emotion, platonic love, joy, and grief emotion. The author extracted the Mean Absolute Value (MAV) features and trained the K Nearest Neighbor (KNN) classifier. The author classified three emotions with 83% classification accuracy and eight emotions with 46% classification accuracy.

Cheng et al. [37] used the Augsburg Bio-signal Toolbox (AUBT) dataset. It included 25 EMG signals of a single subject for sadness, joy, pleasure, and anger. EMG signals are collected from zygomaticus-major under a 2-minute musical condition. The author used wavelet-transform with db5 base function on the EMG signal and found maximum and minimum wavelet coefficients. The author achieved 75% classification accuracy using a Back Propagation Neural Network (BPNN). Yang et al. [38] used the same AUBT dataset. The author used the same on BPNN and recognized emotion with 83.3%. The author introduced the emotional recognition method using a Least Squares Support Vector Machine (LSSVM). The author recognized only four emotions using the LSSVM classifier with 91.67% classification accuracy [38].

Takano et al. [39] proposed a unique wearable device for recognizing smiling and neutral facial expressions. The intensity of EMG signals applied to SVM and classified smile emotion with 90% accuracy. Here, the author detected only two emotions. Jerritta et al. [40] detected sad, neutral, disgusted, surprised, afraid, and happy emotions. The author used Principal Component Analysis (PCA) and KNN classifier for classification. The author achieved 69.5% classification accuracy.

Kehri et al. [41] acquired EMG signals from corrugators' supercilii and zygomaticus-major for 10s. The author acquired EMG signals from 12 subjects while expressing disgust, anger, and fear emotions. The author achieved 91.66% classification accuracy by using wavelet packet transform and SVM. Here, the author recognized only three emotions. Kehri et al. [42] acquired EMG signals from different facial muscles for different emotions. The author collected EMG signals from ten subjects. The author extracted the Kurtogram feature, and compared Convolutional Neural Network (CNN), KNN, Artificial Neural Network (ANN), and SVM classifier. The author achieved 93%, 90%, 91.66%, and 81% classification accuracies, respectively.

Mithbavkar et al. [43] collected 2160 EMG signals from three Kathak dance-trained subjects using four-channel acquisition units. The author acquired EMG signals from SCM muscles, zygomaticus-major, frontalis, and corrugators' supercilii muscles. They extracted the variance, mean, and maximum of RMS features from the EMG signal and applied them to the LSSVM classifier. Classifier classified Shringar (affectionate smile), Raudra (ferocious), Karuna (sad), Hasya (humorous), Bhayanaka (fearful), Veer (courageous), Shant (peaceful), Bibhatsa (repulsion), and Adbhut (amazed) Kathak emotions with 80.3% classification accuracy [44]. Mithbavkar et al. [45] collected 7920 EMG signals from eleven Kathak-trained subjects. The author used the variance, mean, and maximum of the 'Difference in IEMG feature' to train LSSVM, Long and Short Term Memory (LSTM), and a Nonlinear Autoregressive Network with Exogenous Inputs (NARX) classifiers. LSSVM, LSTM, and NARX classifiers achieved 64.29%, 93.63%, and 81.27% classification accuracy, respectively. The LSTM classifier has a forgetting layer, which remembers necessary data and deletes unnecessary datasets. This behavior makes LSTM better than the other classifiers.

CRITICAL INVESTIGATION OF EMG BASED EMOTION RECOGNITION

When a person expresses emotions in ceased form, Ang et al. and Xu et al. [30, 35] recognized three emotions for 6 participants with 94.44% and 27 participants with 97.41% classification accuracy, respectively, and Gibert et al. [31] recognized six emotions with 92% accuracy but for single subjects. Here, if the number of emotions is high, accuracy is less. In all the above research work, one thing is common: that person expressed emotion in a ceased form. Due to this, steady EMG magnitude is captured during that cease condition, and the range of accuracy lies between 87% to 97%.

When a person expresses emotions in unceasing expression form, Picard et al. [36] and Chen et al. [34] recognized eight emotions for single subjects and six emotions for eleven subjects, respectively. For eight and six emotions, both researchers achieved 46% and 37.09% classification accuracy, respectively [36, 46]. Jerritta et al. [40] recognized six emotions for fifteen subjects with 69.5% classification accuracy using PCA and KNN classifier. Here, we found comparatively less classification accuracy in unceasing expression form. Kehri et al. [41] used a CNN classifier for the classification of three emotions of 12 subjects and achieved 93% classification accuracy. Mithbavkar et al. recognized [43-44] nine emotions for eleven subjects with 93.63% classification accuracy using an LSTM classifier with differences in IEMG features. In this research work, the author used facial muscles with SCM neck muscles. This research work incorporated facial expression and head movement for emotion recognition, along with LSTM classifier and difference in IEMG feature [44]. Kehri et al. [41] and Mithbavkar et al. [45] improved classification accuracy using CNN and LSTM classifiers.

CONCLUSION

From the investigation reported in the above sections, it is concluded that many researchers effectively recognized emotions in either ceasing or unceasing expression form. EMG-based emotion recognition in ceased expression achieved classification accuracy in the 91% to 97% range [23, 31, 34, 35], and in unceasing expression, achieved classification accuracy in the 46% to 93% range [36, 43-45]. In the ceased way, a muscle is in a constant state. Muscle movement is static. The features extracted during the ceased state achieved higher classification accuracy than features extracted during the unceasing state. From this survey, we found that deep neural networks like CNN and LSTM are used to improve classification accuracy for multiple subjects in an unceasing state.

REFERENCES

1. P. K. Srimani and T. Hegde, “Analysis of facial expressions with respect to Navras as in Bharathanatym styles using image processing,” Int. J. Knowl. Eng., vol. 3, no. 2, pp. 193–196, Nov. 2012.
2. L. Li, Y. Zhao, D. Jiang, Y. Zhang, F. Wang, I. Gonzalez, E. Valentin, and H. Sahli, “Hybrid deep neural network, hidden Markov model (DNNHMM) based speech emotion recognition,” in Proc. Humaine Assoc. Conf. Affect. Comput. Intell. Interact., Sep. 2013, pp. 312-317.
3. M. Xu, F. Zhang, and S. U. Khan, “Improve accuracy of speech emotion recognition with attention head fusion”, in Proc. 10th Annu. Comput. Commun. Workshop Conf., Jan. 2020, pp. 1058-1064.
4. S. Lalitha, A. Madhavan, B. Bhushan, and S. Saketh, “Speech emotion recognition,” International Conference on Advances in Electronics Computers and Communications, Bangalore, 2014, pp. 1-4.
5. S. T. Shikler, “Analysis of affective expression in speech,” Technical Report, University of Cambridge, UCAM-CL-TR-740 ISSN 1476-2986, Computer Laboratory, 2009.
6. C. Busso, and S. Lee, “Analysis of emotionally salient aspects of fundamental frequency for emotion detection,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 17, no. 4, pp. 582-596, 2009.
7. M. S. Fahad, A. Deepak, G. Pradhan, and J. Yadav, “DNN-HMM-based speakeradaptive emotion recognition using MFCC and epoch-based features,” Circuits, Syst., Signal Process, vol. 40, no. 1, pp. 466-489, Jan. 2021, doi: 10.1007/s00034-020- 01486-8.
8. T. M. Wani, T. S. Gunawan, S. A. A. Qadri, H. Mansor, M. Kartiwi, and N. Ismail, References 99 “Speech emotion recognition using convolution neural networks and deep stride convolutional neural networks,” in Proc. 6th Int. Conf. Wireless Telematics, Sep. 2020, pp. 1-6, doi: 10.1109/icwt50448.2020.9243622.
9. Mustaqeem, M. Sajjad, and S. Kwon, “Clustering-based speech emotion recognition by incorporating learned features and deep BiLSTM,” IEEE Access, 2020, doi: 10.1109/ACCESS.2020.2990405.
10. L Kerkeni, Y. Serrestou, M. Mbarki, K. Raoof, M. Mahjoub, C. Cleder, “Automatic speech emotion recognition using machine learning,” Social Media and Machine Learning, book chapter. intechopen publication, 25th March 2019.
11. A. Bhadane, A. Dixit, V. Ingle, D. Shastri, “Facial expression recognition using image processing,” IJARIIE, vol. 5, no. 3, pp. 266-271, 2019.
12. B. Niu, Z. Gao, and B. Guo, “Facial expression recognition with LBP and ORB features,” Hindawi Computational Intelligence and Neuroscience, Volume 2021, Article 8828245, pp.1-10, 2021.
13. https://www.freepik.com/free-photo/portrait-tired-depressed-caucasian-guy-hiding-his-face-facepalm-sobbing-feel-distressed-emotional-burnout-working-remote-from-home-during-pandemic-quarantine-white-

wall_16224891.htm. Accessed: June 2023.

14. https://www.istockphoto.com/photos/covering-smile. Accessed: June 2023.

15. https://stock.adobe.com/in/images/young-african-american-man-over-isolated-background-with-sad-expression-covering-face-with-hands-while-crying-depression-concept/227140620. Accessed: June 2023.

16. K. Kim, S. Bang, and S. Ki, "Emotion recognition system using short-term monitoring of physiological signals," Medical and Biological Engineering and Computing, vol.42, no.3, pp. 419–427, 2004

17. Y. J. Li, J. J. Huang, H.Y. Wang, N. Zhong, "Study of emotion recognition based on fusion multi-modal bio-signal with SAE and LSTM recurrent neural network," Tongxin Xuebao J. Commun. 2017. vol. 38, pp.109–120.

18. X. Li, D. Song, P. Zhang, G. Yummy, Hou, B. Hu, "Emotion recognition from multi-channel EEG data through convolution recurrent neural network," Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, Shenzhe, China. 15–18 December 2016. pp. 352–359.

19. J. Liu, H. Meng, A. Nandi, M. Li, "Emotion detection from EEG recordings," Proceedings of the 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD); Changsha, China. 13–15 August 2016; pp. 1722–1727.

20. T. Song, W. Zheng, C. Lu, Y. Zong, X. Zhang, Z. Cui, "MPED: A multi-modal physiological emotion database for discrete emotion recognition," IEEE Access, vol. 7, pp. 12177-12191, 2019, doi:10.1109/ACCESS.2019.2891579

21. F. Agrafioti, D. Hatzinakos, A. K. Anderson, "ECG Pattern analysis for emotion detection," IEEE Trans. Affect. Computing, vol. 3, pp. 102–115, 2012, doi: 10.1109/TAFFC.2011.28.

22. G. Valenza, L. Citi, A. Lanatá, E. P. Scilingo, R. Barbieri, "Revealing real-time emotional responses: A personalized assessment based on heartbeat dynamics," Sci. Rep. 2014; 4:4998, doi: 10.1038/srep04998.

23. M. Monajati, S. H. Abbasi, F. Shabaninia, S. Shamekhi, "Emotions state recognition based on Physiological Parameters by Employing of Fuzzy-Adaptive Resonance Theory," Int. J. Intell. Sci., vol. 2, pp. 166–176, 2012.

24. H. F. Garcia, M. A. Alvarez, A. A. Orozco, "Gaussian process dynamical models for multi-modal affect recognition," Conf. Proc. IEEE Eng Med Biol. Soc. 2016, pp. 850- 853.

25. S. Basu, N. Jana, A. Bag, M. Mahadevappa, J. Mukherjee, S. Kumar, R. Guha, "Emotion recognition based on physiological signals using valence-arousal model," Proceedings of the Third International Conference on Image Information Processing, Waknaghat, India, 21–24 December 2015, pp. 50–55.

26. G. Zied, Z. Lachiri, and C. Maaoui, "Emotion recognition from physiological signals using a fusion of wavelet-based features," Proceedings of the International Conference on Modelling, Identification and Control; Sousse, Tunisia. 18–20 December 2015.

27. M. Awais, M. Raza, N. Singh, K. Bashir, U. Manzoor, S. Islam, "LSTM based emotion detection using physiological signals: IoT framework for healthcare and distance learning in COVID-19," in IEEE Internet of Things Journal, 2021, doi: 10.1109/JIOT.2020.3044031.

28. M. Wand, M. Janke, T. Schultz, "Tackling speaking mode varieties in EMG-based speech recognition," IEEE Transaction on Biomedical Engineering, vol. 61, no. 10, pp. 2515-2526, 2014.

29. https://www.pinterest.com/pin/336362665893070260/. Accessed on 29 May 2020.

30. L. B. P Ang, E.F. Bele., R.A. Bernardo., E.R. Boongaling, G.H. Briones, and J.B. Coronel, "Facial expression recognition through pattern analysis of facial muscle movements utilizing electromyogram sensors," TENCON 2004. IEEE Region 10 Conference, vol. 3, pp. 600–603.

31. G. Gibert, M. Pruzinec, T., Schultz, and C. Stevens, "Enhancement of human-computer interaction with facial Electromyographic sensors," in Proceedings of the 21st Annual Conf. of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7. ACM, 2009, pp. 421–424.

32. A. Gruebler, V. Beren, and K. Suzuki, "Emotionally assisted human-robot interaction using a wearable device for reading facial expressions," Advanced Robotics, 26:10, pp. 1143-1159, 2012.

33. A. Gruebler, and K. Suzuki, "Design of a wearable device for reading positive expressions from facial EMG signals," IEEE Transactions on Affective Computing, vol. 5, no. 3, pp. 227-237, 2014.

34. Y. Chen, Z. Yang, and J. Wang, "Eyebrow emotional expression recognition using surface EMG signals," Neurocomputing, vol. 168, Nov. 2015, doi: 10.1016/j.neucom.2015.05.037

35. M. Xu, J. Cheng, J. Li, Y. Liu, X. Chen, "Spatio-temporal deep forest for emotion recognition based on facial electromyography signals," Computers in Biology and Medicine, 2023, vol. 156: 106689. https://doi.org/10.1016/j.compbiomed.2023.106689.

36. R. W. Picard, E. Vyzas, and J. Healey, “Toward machine emotional intelligence: analysis of the affective physiological state,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 10, pp. 1175–1191, Oct. 2001.

37. B. Cheng and G.-Y. Liu, “Emotion recognition from surface EMG signal using wavelet transform and neural network,” J. Comput. Appl., vol. 28, no. 2, Feb. 2008, doi: 10.3724/SP.J.1087.2008.00333.

38. Y. G. Yang and S. Yang, “Study of emotion recognition based on surface electromyography and improved least squares support vector machine,” J. Comput., vol. 6, no. 8, Aug. 2011, doi: 10.4304/jcp.6.8.1707-1714.

39. Y. Takano and K. Suzuki, “Affective communication aid using wearable devices based on biosignals,” in Proceedings of the 2014 Conf. on Interaction Design and children. ACM, 2014, pp. 213–216.

40. S. Jerritta, M. Murugappan, K. Wan, and S. Yaacob, “Emotion recognition from facial EMG signals using higher-order statistics and principal component analysis,” J. Chinese Inst. Eng., vol. 37, no. 3, Apr. 2014, doi: 10.1080/02533839.2013.799946.

41. V. Kehri, R. Ingle, S. Patil, and R. N. Awale, “Analysis of Facial EMG Signal for Emotion Recognition Using Wavelet Packet Transform and SVM,” M. Tanveer R. B. Pachori (eds.), Mach. Intell. Signal Anal. Adv. Intell. Syst. Comput., vol. 748, pp. 247– 257, 2019.

42. V. Kehri and R.N. Awale, “A facial EMG data analysis for emotion classification based on spectral kurtogram and CNN,” International Journal of Digital Signals and Smart Systems, vol. 4, Issue 1-3, pp. 50-61, 2020, https://doi.org/10.1504/IJDSSS.2020.106072.

43. S. A. Mithbavkar and M. S. Shah, “EMG based emotion recognition in Indian classical dance,” Bioscience Biotechnology Research Communications, vol. 13, no. 14, pp. 330– 334, Dec. 2020, doi: http://dx.doi.org/10.21786/bbrc/13.14/76.

44. S. A. Mithbavkar and M. S. Shah, “Recognition of Emotion in Indian Classical Dance Using EMG Signal,” International Journal on Advanced Science, Engineering and Information Technology, vol. 11 (2021), no. 4, pp. 1336-1345, 31 August 2021, doi: 10.18517/ijaseit.11.4.14034.

45. S. A. Mithbavkar and M. S. Shah, “Analysis of EMG based emotion recognition for multiple people and emotions,” 2021 IEEE 3rd Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS), 2021, pp. 1-4, doi: 10.1109/ecbios51820.2021.9510858.

46. J. Chen, T. Ro, Z. Zhu, “Emotion Recognition With Audio, Video, EEG, and EMG: A Dataset and Baseline Approaches,” Digital Object Identifier, 2022. vol. 10. pp. 13229-13242; Doi: 10.1109/ACCESS.2022.3146729.

Fruit Quality Monitoring Using Image Based Freshness Detection

Nitixa Maheshbhai Bhandari
Department of Computer Engineering
Shah & Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ nitixa.17522@sakec.ac.in

Manimala Mahato
Department of Computer Engineering
Shah & Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ manimala.mahato@sakec.ac.in

ABSTRACT

The demand for high-quality and fresh fruits has increased significantly in recent years, driven by consumer preferences for nutritious and visually appealing produce. However, ensuring the freshness and quality of fruits remains a challenge for producers, distributors, and retailers due to the subjective nature of visual inspection and the rapid deterioration of perishable goods. In response to this challenge, the project suggests a novel approach for detection of fruit freshness using image processing techniques. The essential goal of this intended system is to create a computer vision system capable of accurately assessing the freshness of fruits based on visual cues captured in images. The proposed system leverages innovative image processing technique including feature extraction, segmentation, and classification, to analyse key attributes such as colour, texture, and surface defects indicative of freshness or spoilage, thereby enabling accurate and efficient assessment of product quality.

KEYWORDS *: Fruit freshness, Image acquisition, Image analysis, Image classification.*

INTRODUCTION

During the last couple of years, utilization of fruit product items has decisively expanded in numerous nations. It is assessed that around 20% of all fruit products created are lost every year because of value decay. [1] In today's fast-paced world, ensuring the quality and freshness of agricultural produce, particularly fruits, is paramount for agricultural and food processing industries. The ability to accurately assess the freshness of fruits not only guarantees consumer satisfaction but also plays a crucial role in minimizing waste and optimizing resource utilization throughout the supply chain.

This project attempts to build a seamlessly integrated system that is reliable and easy to use that easily integrates with existing workflows for processing fruit quality monitoring systems. By automating freshness-assessed tasks that were previously completed by human inspectors, the system is meant to streamline efficiency, reduce labor expenses, and work at overall proficiency in fruit operations, distribution, and production.

Additionally, the fruit freshness detection system possesses the potential to contribute to good environmental and economic outcomes by decreasing food waste and optimizing resources to promote better crop utilization and guarantee that consumers receive high-quality products. With its ability to continuously monitor capabilities intended for high-performance, low-latency, high throughput, accuracy, and scalability, this sticks to be a very useful tool for stakeholders across the whole supply chain, including farmers, food processors, distributors, retailers, and consumers.

Inspired by the effectiveness of image processing techniques in various fields such as medical diagnostics [2], artificial intelligence [3], and remote sensing [4], the fruit freshness detection project could automate the assessment of fruit freshness and also categorize fruit into fresh and rotten fruit using advanced image processing techniques as discorded in Fig. 1. The system leverages digital imaging technology that can quickly and accurately assess fruit quality, providing stakeholders with crucial information to make judicious choices on harvest, sorting, grading, packing, and distribution. It also assists clients to verify the quality of the fruit.

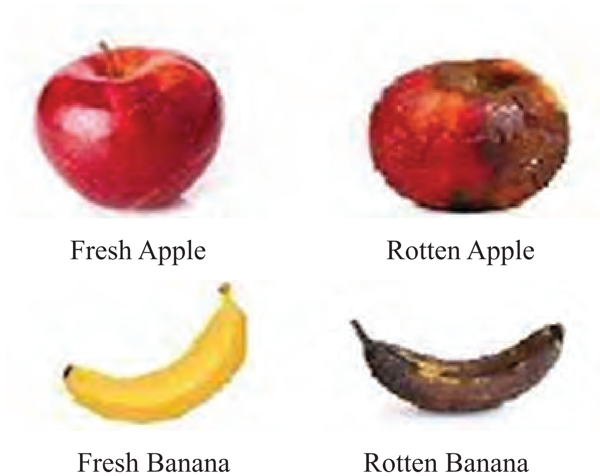


Fig. 1. Fresh Fruit and Rotten Fruit

This system for detecting fruit freshness employs specific techniques of image processing to ascertain the freshness of the fruit in question. It uses the following key steps, which are image preprocessing, image analysis, image classification, and report generation. Its main advantages include being inexpensive, accurate, and simple to implement.

LITERATURE REVIEW

The paper [5] published by Zihe Wei, Miaosen Chang, and Yipeng Zhon proposed a model for finding fruit freshness that is based on the YOLOv8 and SE attention mechanism. To work on this project, the author used the YOLOv8 model, which is an advanced model built upon the YOLO model that has a new feature and improvement, which is channel-wise attention that improves interdependence between channels. The proposed technique accomplishes an average accuracy of 87.8% and maximum accuracy of 95.0% for identifying the freshness of one category of fruit. The current model has limitations, as it is trained on a restricted dataset for apples and bananas, and lacks the capacity to detect a large numbers of fruit varieties.

The paper [6] works on a different technique of machine learning that is Deep Convolution Neural Networks to classify images. This suggested method works on several steps that are related to each other, starting from data gathering to picture resizing, expansion, picture naming, and so forth. After that, the AlexNet model is used with eight layers, from that it contains five work as convolutional layers and other three completely associated layers. This time, moving learning and adjusting the CNN are carried out. At the last stage, a softmax classifier is utilized to classify the image and achieve an accuracy of 98.2%. This proposed model produced a result in 8 ms.

In paper [7], identify the quality of the fruit using an image processing technique. They try to find spoiled fruit from the group of fruit. This technique also requires several steps to find fruit and vegetable quality. This proposed system used different properties of fruit, such as variety, surface, size and shape to differentiate them into different categories. With the help of the K-nearest neighbour algorithm, the author categorizes fruit into good and bad quality. Here the author tries to implement different classifiers to classify fruit into good or bad and compare the results of different classifiers. Results conclude that 89.16% precision of classification has been accomplished by utilizing an SVM classifier. In any case, the KNN classifier shows a better outcome over SVM with a precision of 93.33%.

In this study [8], the author works on a grading method to identify fruit freshness, which is based on Deep Learning. Here the author mainly focuses on saving fruit from being thrown and avoiding fruit wasting. In this article, the author tries to implement deep learning and computer vision techniques for grading the fruit freshness. Here the author implements various deep learning methods, including VGG, ResNet and GoogLeNet. AlexNet, along with the YOLO model, is used for fruit and vegetable reorganization. This YOLO extracts ROI that stands for the region of interest from provided digital images. The result concludes that all deep learning models produced almost same results, but the VGG model was identified as the best during validation.

The paper [9] focuses on producing a model for the automatic classification of fruit. The author proposed two classification branches: one, which is a binary classifier that classifies fresh and rotten fruit, and the second, which identifies types of fruit. Here the author implements a multitasking framework. For the implementation of the model, the author used a transfer learning technique in the model training phase. In this model, the author classified fruit into two categories only, such as fresh or rotten. For the evaluation of his proposed model, the creator utilized straightforward pictures from the current dataset and genuine pictures taken from the web, both addressing new and spoiled fruit for various fruit classifications as their dataset. This model suggested by author attained an accuracy of 98.50% for freshness classification and 97.43% for the different types of fruit categorization.

In paper [10], the author observes that there is a lack of a multi-fruit data set that can be used for real-time food quality examination. To solve this problem, the author provided a new dataset that contains images of fruit. The main objective of the author is to generate a dataset that facilitates the development and evolution of fruit freshness assessment models. The dataset contains 11 fruit images, which are classify into three different classes of freshness and five models, such as ShufeNet, SqueezeNet, EfcientNet, MobileNet-V2 and ResNet18 as baseline models to identify quality of the fruit utilizing this dataset. The proposed dataset is efficiently coordinated and commented, which make it compatible for testing the execution of a new learning classifier. This dataset can be helpful to the exploration of the local area in the fields of PC vision and AI, for example, by applying it to different examination errands like fruit classification and fruit quality acknowledgment. Here in this study, the creator likewise recognizes the impediment of the dataset that contain small amount of data.

In this paper [11], the author proposed a model for Indonesia because of the tropical climate of Indonesia; fruit and vegetables easily grow there. Especially fruit plants. Such a long way, there are as yet many fruit suppliers that send unsatisfactory products because of the absence of exactness in doing the fruit product arranging process done through its workers when the fruit product is taken from the manor and the passage of another fruit products from an ill-advised packaging. The suggested model identifies spoiled fruit in both the stage from production to consumption. The proposed model works on profound learning with the help of CNN (Convolutional Neural Network) model to recognize freshness of the fruit. The proposed model has some impediments. This model has restricted training on an informational index, for example, fresh apples, rotten apples, fresh oranges, rotten oranges, fresh bananas, and rotten bananas; hence, this proposed model can anticipate these 6 classes.

The paper [12] suggests a model that uses a machine learning technique to grade fruit freshness. This proposed model implements classification of images using a Convolutional Neural Network (CNN) to calculate freshness of the fruit. Convolutional Neural Network (CNN) is implemented to recognize the size, shape and colour of different fruit samples. Datasets are very important for the success of any model, so here the author tries to create their own dataset with different images of fruit. This model categorized fruit into three categories, such as ripe, unripe and overripe depending on the freshness of the fruit.

The paper [13] author proposed a model that automatically recognizes the fruit and evaluates fruit freshness. In this proposed system, the author focuses on the fruit supply chain that includes fruit classification and fruit freshness detection. Here the author tries to generate output in audio format. In this model, the author mainly focuses on Apple and categorizes Apple in different categories such as Apple A, Apple B, Apple C, Apple D, Apple E and Apple F using VGGnet architecture. After identifying types of fruit, the author used YOLOv3 that is an advanced deep neural network model, to evaluate given apple is fresh or not.

In study [14], the author presents a deep learning model aimed at delivering vegetables and fruits that are both flavorful and nutritious to customers. The research explores various deep learning architectures, including GoogleNet, DenseNet-201, and ResNext-101, to assess the freshness of the produce. Additionally, Principal Component Analysis (PCA) is utilized for dimensionality reduction. This model successfully attained an accuracy rate of 96.98% in identification of fruits and vegetable freshness.

The paper [15] suggests a model for identifying the freshness of the fruit using an image sensor and edge processors with a lightweight CNN architecture. This model expects ON Semiconductor's 13 MP AR1335 image sensor along with their global screen-based, 12 MP, 90-frame-per-second image sensor (XGS-12). The continuous setup with ON Semiconductor's AR1335 13 MP camera associated with Xavier is displayed in Fig. 2. This model classifies fruits into two categories: fresh and rotten, specifically focusing on three types: bananas, oranges, and apples. In high-stress scenarios, it utilizes ON Semiconductor's 13 MP camera equipped with an AR1335 sensor to achieve an accuracy rate of 97%.

In paper [16], a proposed model that detects the freshness of fruit using Raspberry Pi. The proposed model has a detector system that identifies the freshness of fruit using Raspberry Pi. This proposed framework contains different parts, for example, Raspberry Pi 0W, computerized picture processing, Nearness sensor, Load cell and Gas sensor. At the point when the fruit is being picked, it ought to be set in a transport line, where it is gone through the sensor unit which can recognize and show the total freshness status of the fruit. This proposed system gives 80% accuracy for detecting the freshness of fruit.

Fig. 2. Ongoing connection setup with AR 1335 13 MP associated with NVIDIA Jetson Xavie

During this study, we are going through various deep learning models. All these models play important roles in fruit freshness detection by extracting image-based features and identifying fruit freshness level. Here is an overview of deep learning models.

Deep Learning Model

Deep learning is one of the latest examples in AI and examination in fake intelligence. [17] Numerous enormous forward jumps have been made in this field all through the world. Deep learning [18] is a subfield of simulated intelligence that deals with computations that are moved by the plan and capacity of the brain. With everything taken into account, it is comparable to the workings of our brain. The Deep learning assessment resembles the ways tactile technique is facilitated, where each of the neuron is link together and passes on data. Because of that that Deep learning makes it conceivable to zero in on information portrayals with different degrees of reflection.

In Deep Learning, three types of Neural Networks are utilized, which are Artificial Neural Network (ANN), Recurrent Neural Network (RNN), and Convoluted Neural Network (CNN). This approach has undoubtedly personalized the speech detection system, visual object article identification, object acknowledgment, and considerably more. The justification behind this study is to design a reasonable experiment to expect to cultivate a CNN model to perceive the fruit freshness.

Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are a type of deep neural network primarily used for image processing tasks. They are specifically designed to automatically and adaptively learn spatial hierarchies of features from input images. CNNs have transformed the field of computer vision and are widely used in applications such as image classification, object detection, semantic segmentation, and more. Their ability to automatically learn different levels of representations from raw pixel data makes them incredibly powerful tools for various image processing tasks.

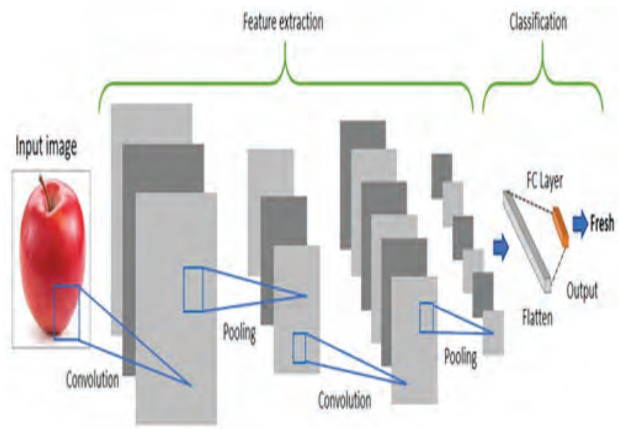


Fig. 3. Design outline of Convolutional Neural Network

Using Convolutional Neural Networks (CNNs) to detect fruit freshness through image processing includes several important steps. First, a detailed dataset is needed, consisting of labeled images of different fruits at various stages of freshness, including both fresh and spoiled instances, is assembled. Subsequently, the collected images undergo preprocessing to standardize their size, normalize pixel values, and enhance dataset diversity through augmentation techniques like rotation and flipping. The design outline of CNN is shown in Fig. 3. The CNN architecture designed for this task typically comprises various convolutional layers executed by pooling layers to actually extricate progressive characteristic from input pictures. It may incorporate well-established CNN architectures such as VGG, ResNet, or Inception, tailored to meet the specific requirements of freshness detection.

Pre trained convolutional neural networks

AlexNet

AlexNet, an innovative convolutional neural network (CNN) architecture, transformed the field of computer vision when it won the 2012 ImageNet Large Scale Visual Recognition Challenge. While not specifically designed for freshness detection, its deep learning principles and architectural innovations paved the way for subsequent CNNs tailored to image processing tasks, including fruit freshness detection. AlexNet's use of ReLU activation, convolutional layers, and fully connected layers laid the foundation for effective extraction of feature and categorization, making it influential in the development of CNN-based models for freshness detection applications, and ultimately inspiring the creation of more specialized and accurate fruit quality assessment systems. In Papers [6] and [8], the authors implement this model for the classification of the fruit.

GoogleNet

GoogleNet, otherwise called Inception v1, is a milestone convolutional brain organization (CNN) engineering created by Google scientists. While not straightforwardly customised for freshness identification, its beginning module with the numerous equal convolutional pathways permit proficient extraction of assorted highlights at various scales. This compositional development empowers GoogleNet to catch complex examples in pictures successfully, making it an important resource in creating CNN-based models for newness recognition errands. GoogleNet's commencement module has enlivened ensuing CNN models, adding to progressions in picture handling and characterization assignments. In Paper [14], the author implemented the GoogleNets model and achieved 96.98% freshness detection.

Table 1. Comparative Analysis of Different CNN Models

Name of pretrain network	Depth	Size	Para-meter	Input image size	Recognition Accuracy Rate	Training Time
AlexNet	8	227	61.0	227x227	95.81	235
GoogleNet	22	27	7.0	224x224	94.76	611
MobileNet	53	13	3.5	224x224	91.06	240
ResNet	101	167	44.6	224x224	94.24	1394
VGG	16	258	138	224x224	99.48	1334

MobileNet

MobileNet is a lightweight convolutional neural network (CNN) architecture intended for proficient sending on mobile and installed gadgets. While not explicitly created for newness identification, its depth-wise distinct convolutions and effective model design empower elite execution with altogether less boundaries and calculations contrasted with conventional CNNs. due to this, MobileNet is suitable for asset-compelled scenarios and has the potential to be used for continuous newness identification applications on mobile devices or edge figuring stages with limited computational resources. After implementing a number of models in study [10], the author used MobileNet to achieve 96.3% accuracy.

ResNet

ResNet, which stands for residual neural network, is a deep convolutional neural network (CNN) generally utilized in image processing because of its astounding depth and excellent accuracy in performing these tasks. This enables effective feature learning and allows ResNet to detect deep features that are important for newness detection in fruit images.The depth and performance of ResNet enables CNN based models for newness detection to leverage its feature extraction capabilities and learn complex hierarchical representations from the input images. In references [8] and [11], the systems were designed using these particular models to which the mentioned results were obtained with accuracy rates of 96.5% and 99.8%.

VGG

VGG, or Visual Calculation Gathering Organization, is an example of an exemplary convolutional neural network (CNN) architecture well known for its simplistic and efficient use in image classification tasks. While not explicitly custom-made for freshness identification, its profound design with various convolutional layers empowers the extraction of progressive features from input images, ultimately enabling accurate classification and assessment of fruit quality. With its direct plan and uniform engineering, VGG can act as major areas of strength for creating CNN-based models for freshness discovery, giving an establishment to catching important examples and surfaces demonstrative of fruit freshness in pictures. In paper [13], the proposed system used this model to identify types of fruit.

LIMITATION

Despite the advancements in fruit freshness detection using image processing, existing systems have several limitations that hinder their effectiveness and practical utility. During this study, we observed various systems that implement various machine learning and artificial techniques to identify the freshness of the fruit. The existing system has several limitations, such as limited accuracy, sensitivity to the environmental factor, hardware requirements like high-resolution cameras and specialized imaging equipment, sensors, some systems that do not work for varieties of fruits, lack of real-time feedback, user interface complexity, etc. By overcoming these challenges, future freshness detection systems can offer improved accuracy, reliability, efficiency, and usability, ultimately enhancing the quality and safety of fresh produce in various applications.

PROBLEM STATEMENT

The detection of fruit freshness is a critical task in the food and agricultural sector to guarantee the quality and safety of fresh produce. Traditional methods of fruit freshness assessment, such as manual inspection and chemical analysis, are often labour-intensive, time-consuming, and subjective. Nowadays, there has been developing interest in utilizing image processing processes to automate fruit freshness detection processes, offering the potential for improved accuracy, efficiency, and scalability. Therefore, we proposed one system that classified fruit into fresh and bad as shown in Figure 1 and also identified the freshness of fruit effectively and accurately.

DATASET

In our proposed model acquire various images from the Kaggle [19]. Kaggle permits clients to find datasets they need to use in building simulated intelligence models, distribute datasets. This dataset contains images of the different fruit such as, Apple, Orange, Banana. This dataset also includes some real time images to improve quality of our dataset. We will try to implement this model for the varieties of fruit. So, this proposed include images of different fruit in dataset so it capable to find freshness of fruit freshness. Fig. 4. shows some sample fruits images.

Fig. 4. Few examples of fresh and rotten fruits images

PROPOSED METHODOLOGY

During this study, we identify various problems such as food wastage, extracting quality of fruit, customer satisfaction, etc. The project on fruit freshness detection using image processing meets the pressing needs of quality control, automation, productivity enhancement, consumer confidence, and technological advancement in the agriculture and food industry. Its successful implementation promises significant benefits for stakeholders across the supply chain, from producers and processors to retailers and consumers. The primary target of this framework is to make a profoundly exact and effective system that creates an automatic fruit freshness detection system with a user-friendly interface. The proposed model is an automatic fruit freshness detection system. Here, this proposed system uses an image processing technique to identify fruit freshness effectively and accurately.

The system aims to address many objectives like,

- It can identify fruit freshness automatically,
- It improves the quality control process in agricultural and food processing industries by providing real-time monitoring and assessment of fruit freshness
- It improves efficiency and productivity
- The system extracts the relevant information from the fruit images and provides to the stakeholders that involved in the harvesting, distribution, and promotion of fruits
- The proposed system offers real time decision making
- The system provides food safety and customer satisfaction.

To implement proposed system, we will utilize various image processing algorithms, including feature extraction, segmentation and classification that analyse key feature of the fruit such as colour, texture, and surface defects that indicate the freshness or spoilage of the fruit. The proposed system may follow various concepts of image processing as shown in Fig. 5. Below, depicts the various modules that are used in our proposed model.

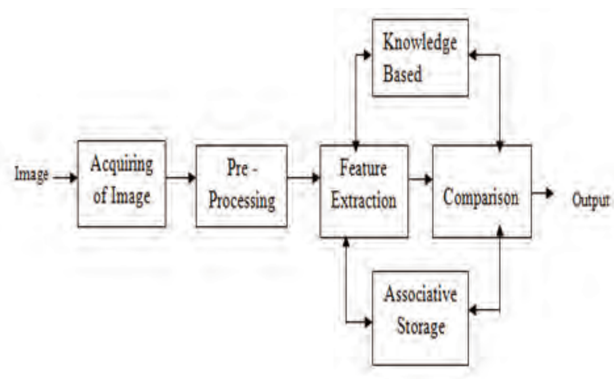


Fig. 5. Flow of development model

Image acquisition: This is the first phase of the fruit freshness detection system. This phase captures high-quality images of fruits; for that, you can use digital cameras or mobile devices. You can also upload an existing image.

Preprocessing: Before using the image for the feature extraction, preprocessing of the image is required to improve the quality of the image. This step involved removing noise, resizing the image, and normalizing features for analysis.

Feature extraction: This process is used to extract relevant features from images that will used to detect freshness of the fruit using various feature extraction techniques. In our proposed model, we try to implement fusion of deep features with SIFT (Scale-Invariant Feature Transform) features, which refers to combining information extracted from different types of feature extraction methods to enhance the accuracy and robustness of detecting freshness in fruits. This extracted feature can be used to identify the freshness of the fruit.

Model Training: The detected feature will be utilized for the model training. To train our proposed image-based freshness detection system, we will use deep neural network (DNN) architecture. In this proposed model, a completely connected neural network (DNN) was executed utilizing the Keras Sequential API. During this process, the softmax activation function is proposed to be used for multi-class classification of various fruit varieties, and the Adam optimizer is used to compile the model, which is known for its adaptive learning rate adjustments to minimize the categorical loss function. The proposed deep learning-based freshness classification model effectively predicts fruit freshness levels with high accuracy.

Reporting: In this stage, a thorough report will be produced. This produced reports that outlined the fruit's freshness examination results. The fruits are also divided into different freshness categories in the report, providing a clear and valuable overview of the assessed produce.

Model Evolution: A number of metrics, including F1-score, confusion matrix, recall, accuracy, and precision, among others, will be utilized to evaluate the proposed fruit freshness detection model.

CONCLUSION

Rest, the fruit freshness detection endeavour project is aimed at changing the standards of quality control management within agriculture and the food processing sector. This project proposes the development of a custom automated system to address the need for the time-appropriate evaluation of fruit freshness availability with the renders of advanced image processing and machine learning. It has armed stakeholders with an effective tool for automating freshness assessment tasks. The system is expected to increase efficiency, reduce waste, and enhance effectiveness in the operations of fruit production and distribution. Such systems have been designed in the manner that makes them highly accurate, inexpensive, easy to implement, and user-friendly.

REFERENCES

1. Guohua, Hui, et al. "Study of peach freshness predictive method based on electronic nose." Food Control 28.1 (2012): 25-32.
2. S. Misar, S. Madlani, S. Tamhane and M. Mahato, "ECG Report Analysis System," 2021 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), Pune, India, 2021, pp. 1-4, doi: 10.1109/SMARTGENCON51891.2021.9645917.
3. N. P. Parab, A. A. Rampurawala, S. S. Khade and M. Mahato, "Personalised AI Based Yoga-Pose Recommendation System Using Game Theory," 2023 International Conference on Advanced Computing Technologies and Applications (ICACTA), Mumbai, India, 2023, pp. 1-6, doi: 10.1109/ICACTA58201.2023.10393658
4. Mahato, Manimala & Gedam, Shirishkumar & Joglekar, Jyoti & Buddhiraju, Krishna. (2019). Dense Stereo Matching Based on Multiobjective Fitness Function--A Genetic Algorithm Optimization Approach for Stereo Correspondence. IEEE Transactions on Geoscience and Remote Sensing. PP. 1-13. 10.1109/TGRS.2018.2883483.
5. Wei, Zihe, Miaosen Chang, and Yipeng Zhong. "Fruit Freshness Detection Based on YOLOv8 and SE attention Mechanism." Academic Journal of Science and Technology 6.1 (2023): 195-197.
6. Amin, Umer, et al. "Automatic fruits freshness classification using CNN and transfer learning." Applied Sciences 13.14 (2023): 8087.
7. K B Mirra, P Pooja, S Ranchani, R Rajakumari. "Fruit Quality Analysis using Image Processing." International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958 (Online), Volume-9 Issue-5, June 2020.
8. Fu, Yuhang, Minh Nguyen, and Wei Qi Yan. "Grading methods for fruit freshness based on deep learning." SN Computer Science 3.4 (2022): 264.
9. Kang, Jaeyong, and Jeonghwan Gwak. "Ensemble of

multi-task deep convolutional neural networks using transfer learning for fruit freshness classification." Multimedia Tools and Applications 81.16 (2022): 22355-22377.

10. Abayomi-Alli, Olusola O., et al. "FruitQ: a new dataset of multiple fruit images for freshness evaluation." Multimedia Tools and Applications 83.4 (2024): 11433-11460.

11. Valentino, Febrian, Tjeng Wawan Cenggoro, and Bens Pardamean. "A design of deep learning experimentation for fruit freshness detection." IOP conference series: earth and environmental science. Vol. 794. No. 1. IOP Publishing, 2021.

12. Harsh, Aniket, et al. "Fruit freshness detection using CNN approach." International Research Journal of Modernization in Engineering Technology and Science 2.6 (2020): 14-18.

13. Kuriakose, Anna Biju, et al. "Automatic Fruit Classification and Freshness Detection." India: Ilahia College of Engineering and Technology Muvattupuzha (2021).

14. Yuan, Yue, and Xianlong Chen. "Vegetable and fruit freshness detection based on deep features and principal component analysis." Current Research in Food Science 8 (2024): 100656.

15. Ananthanarayana, Tejaswini, Raymond Ptucha, and Sean C. Kelly. "Deep learning based fruit freshness classification and detection with CMOS image sensors and edge processors." Electronic Imaging 32 (2020): 1-7.

16. Jayasankar, Krithika, et al. "Fruit freshness detection using raspberry pi." International Journal of Pure and Applied Mathematics 119.15 (2018): 1685-1691.

17. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." nature 521.7553 (2015): 436-444.

18. Minar, Matiur Rahman, and Jibon Naher. "Recent advances in deep learning: An overview." arXiv preprint arXiv:1807.08169 (2018).

19. https://www.kaggle.com/code/faizalkarim/multi-label-classification-name-freshness/input

Future Finder: Market Analysis Tool for Economic Opportunities

Varsha Kulkarni
Student
Lokmanya Tilak College of Engineering

Sharmeen Gardi
Student
Lokmanya Tilak College of Engineering

Mahek Singh
Student
Lokmanya Tilak College of Engineering

Vedika Chavan
Student
Lokmanya Tilak College of Engineering

ABSTRACT

Navigating dynamic job markets and educational opportunities is challenging due to fragmented and generic guidance. This project proposes a machine learning-based platform to deliver personalized recommendations for education and career pathways. By integrating diverse datasets—employment trends, regional economic conditions, skill demand forecasts, and educational resources—the platform provides tailored insights aligned with user-specific goals and preferences. Key features include predictive analytics for future trends, real-time updates on industry changes, and a free, accessible interface for wide adoption. This tool aims to empower individuals with data-driven, actionable insights to make informed, future-ready decisions in their educational and career journeys.

KEYWORDS *: Educational opportunities, Machine learning, Personalized recommendations, Predictive analytics, Data-driven insights.*

INTRODUCTION

Navigating career decisions in today's fast-evolving job market can be challenging for individuals seeking clarity and direction. Whether transitioning to a new career, advancing within an existing role, or starting fresh as students, identifying the most suitable paths often requires costly, inaccessible, and generic career counseling. Existing platforms like Naukri and LinkedIn provide job listings but lack the ability to deliver personalized career guidance tailored to an individual's aspirations and unique profile.

To address these challenges, FutureFinder is an innovative platform designed to offer personalized career recommendations and actionable insights to students, working professionals, and career changers. By integrating advanced algorithms, data collection, and user-centric design, FutureFinder provides a holistic and tailored approach to career planning and development.

The platform focuses on three primary user groups:

Students: Assisting in identifying suitable degree programs, recommending universities, and outlining potential career pathways aligned with their aspirations.

Working Professionals: Offering skill set analysis, career development paths, advanced education options, and salary benchmarking for growth within their current roles or industries.

Career Changers: Evaluating transferable skills, suggesting new career paths, and aligning recommendations with user interests and strengths.

FutureFinder employs a robust framework comprising four key components:

Comprehensive Career Data Collection: Detailed information about users' qualifications, skills, and goals is collected to build a holistic profile.

Career Transition, Growth, or Exploration: Adaptable modules cater to diverse user needs, whether exploring new roles, advancing within existing careers, or initiating a career journey.

Tailored Assessments: Custom evaluations consider salary expectations, geographic preferences, work-life balance, and educational goals to deliver accurate recommendations.

Personalized Results: Users receive detailed reports with career suggestions, potential job roles, industry trends, and geographic insights for informed decision-making.

Unlike traditional career counseling services, the platform democratizes access to tailored advice, empowering users to make informed decisions about their educational and professional futures. This paper outlines the design, implementation, and key features

of FutureFinder while exploring its potential to redefine career discovery and planning.

LITERATURE REVIEW

The reviewed literature highlights the significant potential of machine learning and data mining in career guidance and educational applications.

It further illustrates the effective use of data mining techniques to analyze educational data, which helps improve students' academic performance. The study demonstrates how mining such data can reveal hidden patterns and insights, aiding educators in making informed decisions to enhance learning outcomes. However, its focus remains primarily on educational performance rather than extending to career guidance or broader applications [1]. Machine learning and AI can play a critical role in career decision-making by analyzing student profiles, including interests, skills, and academic performance, to recommend suitable career paths [2,3]. However, current career guidance systems lack comprehensive coverage of subsequent steps, such as college and country selection, living condition analysis, and long-term career planning [2,3]. Furthermore, systems leveraging network visualization and skill-matching techniques excel at identifying and addressing skill gaps in specific industries, such as IT, but often fail to integrate real-time job market trends and industry-specific requirements [4,3]. Additionally, data mining techniques effectively uncover hidden patterns in educational data, which can improve academic performance and decision-making, though they often stop short of applying these insights to career planning [5,6]. Moreover, feature selection techniques are essential for enhancing the accuracy of predictive models in educational data mining, making them critical for building robust career guidance systems [7]. Predictive modeling has demonstrated success in forecasting academic outcomes but has limited application in personalized career recommendations and industry-specific guidance [8,9]. Lastly, while frameworks for student performance monitoring and academic administration using data mining are well-developed, they do not extend to holistic career planning that includes job market data and tailored advice [9,6]. Collectively, the reviewed works underscore the need for integrating real-world job market insights, long-term career pathing, and personalized guidance into career recommendation systems to enhance their relevance and utility [2,3,6].

PROPOSED SYSTEM

The proposed system aims to provide a comprehensive career guidance platform tailored to the needs of working professionals and students by leveraging data collection, assessment modules, and machine learning classifiers.

User Onboarding Component:

User Login/Register: The user logs into the system using their credentials, or registers as a new account and then proceeds to log in the system.

The system begins the Assessment Process, comprising of:

Personality and Aptitude Assessment: Collects data on hobbies, interests, strengths, and weaknesses.

Preference and Experience Assessment: Gathers details on career goals, preferred industries, current job status (if applicable), skills, and experience.

The system generates a detailed Career Report based on user responses.

The report enables the user to explore suitable locations for their career aspirations and browse potential careers or job roles aligned with their qualifications and interests.

Career Prediction Algorithm: This algorithm integrates machine learning to predict and rank suitable career options for users.

The user provides details, including skills, education, experience, and career preferences, during the assessment process.

A detailed assessment collects insights into career motivators, preferred industries, and additional user-specific factors. The data is securely stored in a database for analysis and future reference.

The system queries a career database to retrieve relevant career profiles. These profiles include key attributes such as job descriptions, required skills, salary ranges, and growth prospects.

A machine learning classifier is applied to classify and predict the most suitable career paths for the user. Multiple classifiers are tested (e.g., Decision Trees, Random Forest, Neural Networks) to identify the optimal model with the highest accuracy. The classifier leverages the user's input data and learned patterns to generate predictions.

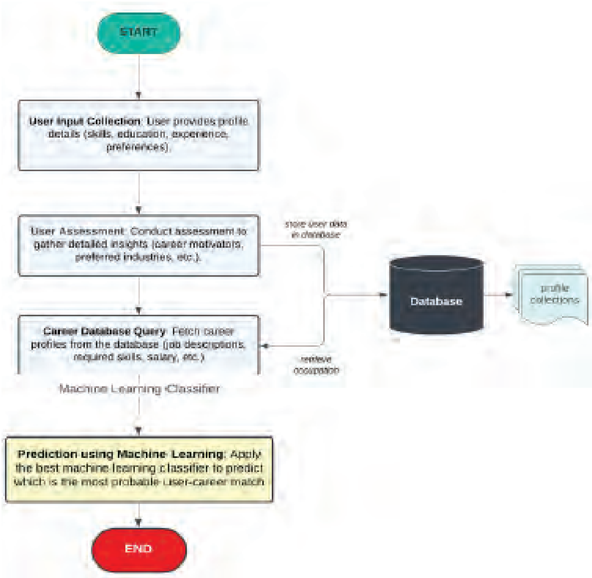


Fig. 1: Proposed System Framework

IMPLEMENTATION DETAILS

The backend implementation consists of two primary aspects: defining the data structure and determining the best machine learning algorithm for career predictions. The input data is carefully structured to include feature variables and target variables. The feature variables capture user profiles with fields such as Education: Bachelor's, Master's, or PhD degrees and certifications, Skills: Hard and soft skills, Work History: Current job title, years in the role, and industry experience, Career Preferences: Career goals, motivators, job type, preferred location, work schedule, travel requirements, and work-life balance importance. Some variables, such as certifications and job history, are optional for students but essential for job seekers.

The target variable is the "suitable career," which represents the best-fit career path based on the input data.

The career profile dataset contains detailed information for each career, such as career title, job description, industry, and company name (optional); required and preferred skills, certifications, and education requirements, experience thresholds, salary range, job requirements, and ideal personality traits and any additional parameters like work location, travel requirements, and preferences.

To enhance exploration, upcoming features will include recommendations for the best countries for specific careers, colleges offering relevant programs, and other supportive data.

Comparative Study on Algorithms

We performed a comparative study on four machine learning algorithms: Decision Tree, K-Nearest Neighbors (KNN), Random Forest, and Support Vector Machine (SVM). The different datasets used for the classification and testing of algorithms are split into sets of test and training models, with 70% as test and 30% as training datasets. The machine learning model is fitted using the train dataset and the test Dataset is used to assess how well a machine learning model fits the data. This study was implemented using three datasets: Studies Career Recommendation Dataset (Kaggle): Provides information on students' academic performance, extracurricular activities, and career aspirations, used for predicting career paths based on student profiles.

IT Career Proficiency Dataset: Surveys proficiency in various IT skills, ideal for classification and clustering tasks to recommend suitable IT roles.

CareerMap Dataset: Combines psychological traits and technical skills for career predictions in computer science, addressing the need for labeled data through an ensemble approach.

Algorithm Testing and Career Prediction Module

The career prediction module processes user input data, including skills, education, and preferences. The module applies machine learning classifiers to analyze patterns and predict career matches. Comparative testing of algorithms was conducted on the datasets to identify the best-performing model for the platform. The performance of Decision Tree, KNN, Random Forest, and SVM was evaluated for accuracy and reliability.

The system then generates a personalized list of career suggestions, which users can explore along with detailed career profiles.

This architecture ensures a robust system capable of offering tailored recommendations, with plans for continuous enhancements like integrating global opportunities and educational pathways.

The following tabular representations depicted the classification accuracy, precision, recall, and F1 score respectively for each algorithm used on the datasets. The values for each metric evaluation vary as the datasets used are random and have different classes and categories, generating variation in the output of the algorithms.

Figures 2 and 3 depict the classification accuracy, and F1 score respectively for each algorithm used on the datasets.

Precision is the proportion of real positive examples among those classified as positive by the model. Recall, commonly referred to as sensitivity, is the percentage of positive examples among all the actual positives. The F1 score integrates the model's recall and precision as the harmonic mean of the two.

All classifiers performed admirably, with high accuracy scores and relatively few variations in the results. However, upon comparing the metrics, we observed that the Random Forest model consistently outperformed the others, delivering the highest accuracy and superior performance across the datasets.

Random Forest demonstrated superior results across all metrics, making it the most reliable choice. Based on the observations, we conclude that Random Forest outperforms Naive Bayes, SVM, Decision Tree, and k-NN, making it the algorithm of choice for our system. We primarily focus on accuracy, so the accuracy results over all datasets for various algorithms were as follows:

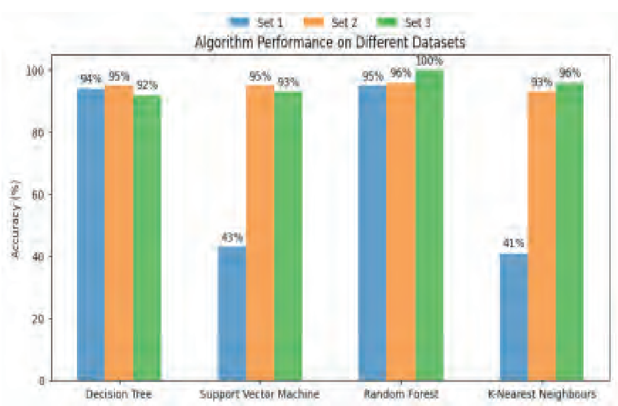


Fig. 2: Accuracy in (%) of the algorithms applied in each dataset

The evaluation metrics employed to assess the performance of each algorithm are as elaborated in the following tables.

Table 1: Evaluation metrics for the classification algorithms decision tree

	Accuracy	Precision	Recall	F1-Score
Dataset 1	0.94	0.94	0.94	0.94
Dataset 2	0.95	0.93	0.94	0.98
Dataset 3	0.93	0.98	0.93	0.96

Table 2: Evaluation metrics for the classification algorithm random forests

	Accuracy	Precision	Recall	F1-Score
Dataset1	0.95	0.95	0.95	0.95
Dataset2	0.96	1.00	0.97	1.00
Dataset3	1.00	1.00	1.00	1.00

While both Random Forest and Decision Tree demonstrate comparable performance with high accuracy scores, Random Forest consistently outperforms Decision Tree across most datasets, albeit with marginal differences. This establishes Random Forest as the more effective model among the two, making it the preferred choice for the given application.

Table 3: Evaluation metrics for the classification algorithm support vector machine

	Accuracy	Precision	Recall	F1-Score
Dataset1	0.43	0.35	0.44	0.35
Dataset2	0.95	0.96	1.00	0.93
Dataset3	0.93	0.99	0.99	0.99

Table 4: Evaluation metrics for the classification algorithm k-nearest neighbour

	Accuracy	Precision	Recall	F1-Score
Dataset1	0.41	0.42	0.41	0.41
Dataset2	0.84	0.86	0.82	0.82
Dataset3	0.96	1.00	0.93	0.92

Finally, the F1 Score is a crucial metric in machine learning that provides a balanced measure of a model's precision and recall. The F1 Score formula is derived from the harmonic mean of precision and recall, making it an essential component in the precision recall F1 score framework.

The F1 Score for different datasets and learning algorithms is as follows:

To conclude, Random Forest consistently outperforms other models in generalizing to new data, as evidenced by its strong cross-validation performance. Hence, this will be the algorithm in use for our system. The future scope of this project involves expanding the platform's capabilities to provide even broader and more personalized career guidance, ensuring accessibility and relevance in a rapidly changing education and job market.

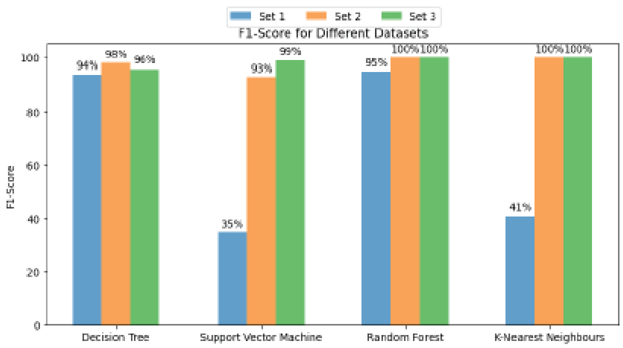


Fig. 3: F1 Score of the algorithms applied in each dataset

CONCLUSION

In conclusion, the FutureFinder project represents a comprehensive and innovative approach to career guidance and education planning. By leveraging advanced technologies such as machine learning and data integration, we aim to create a user-centric platform that empowers students and professionals to navigate their career paths with confidence. Through a structured implementation plan divided into distinct phases—encompassing research, design, machine learning model development, testing, and deployment—we are committed to delivering a robust solution that meets the evolving needs of our users. By focusing on personalized recommendations, real-time job market insights, and user-friendly interfaces, FutureFinder will not only simplify the process of career exploration but also address the current gaps in the market, such as the lack of tailored guidance and accessibility. Our approach incorporates continuous improvement through user feedback, ensuring that the platform remains relevant and effective in an ever-changing job landscape. As we move forward, we remain dedicated to refining our algorithms, expanding our database of occupations, and consulting industry experts to provide the most accurate and valuable recommendations. Ultimately, FutureFinder aims to bridge the gap between ambition and opportunity, helping individuals shape their futures with clarity, purpose, and confidence.

REFERENCES

1. A.M. El-Halee's, and M. M. Abu Tair, "Mining educational data to improve students' performance: A case study," International Journal of Information and Communication Technology Research, 2011, pp. 140-146.M. tech Er. Rimmy Chuchra. " Use of data mining techniques for the evaluation of student performance: a case study ". International Journal of Computer Science and Management Research, 1(3):425-433, October 2012.

2. Dahanke Ajay, Shinde Nilesh, Dhagate Anirudh, Shaikh Huzaif, "An Intelligent Career Guidance System using Machine Learning", International Research Journal of Engineering and Technology (IRJET), March 2022.

3. Tanmay Mathur, Pawar Soniya Dattatray , Mansi Singh, Pawan Kumar, "Intelligent Career Guidance System using Machine Learning", IJARIIE-ISSN(O)-2395-4396, June 2022.

4. S. Karthik M. Sukanya, S. Biruntha and T. Kalaikumaran, "Mining Data mining: Performance improvement in education sector using classification and clustering algorithm," ICCCE In Proceedings of the International Conference on Computing and Control Engineering, 2012.

5. Romero, C., Ventura, S., "Educational data mining: A review of the state of the art," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40(6):601-618, 2010.

6. Han, J., Kamber, M., Pei, J., "Data Mining: Concepts and Techniques," Elsevier, 2012.

7. Ramaswami, M., Bhaskaran, R., "A Study on Feature Selection Techniques in Educational Data Mining," Journal of Computing, 1(1):7-11, 2009.

8. Kotsiantis, S., "Use of Machine Learning Techniques for Educational Proposes: A Decision Support System for Forecasting Students' Grades," Artificial Intelligence Review, 37(4):331-344, 2012.

9. Natek, S., Zwilling, M., "Student data mining solution – knowledge management system related to higher education institutions," Expert Systems with Applications, 41(14):6400-6407, 2014.

Automated Detection of Deficit/Hyperactivity Disorder (ADHD) in Children Using Electroencephalography (EEG) and Hybrid Deep Learning Model (CNN & LSTM)

Fabiya Suhail Chhapra
Department of Computer Science Engg. (Data Science)
Vidyavardhini's College of Engineering and Technology
✉ fabiya.S221396201@vcet.edu.in

Nupur Lalit Vartak
Department of Computer Science Engg. (Data Science)
Vidyavardhini's College of Engineering and Technology
✉ nupur.213118205@vcet.edu.in

Ichhanshu Jaiswal
Department of Computer Science Engg. (Data Science)
Vidyavardhini's College of Engineering and Technology
✉ ichhanshu.jaiswal@vcet.edu.in

ABSTRACT

Attention Deficit Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder in children, often persisting into adulthood, characterized by inattention, hyperactivity, and impulsivity. Diagnosis typically relies on subjective assessments, risking misdiagnosis or delays. EEG, reflecting brain activity, offers objective support for ADHD detection.

This study introduces a novel feature extraction method that evaluates dynamic connectivity tensors among EEG channels, preserving temporal and spatial structures while reducing dimensions. These features are processed by a hybrid neural network combining CNNs and LSTMs to capture spatial electrode dependencies and temporal EEG patterns.

Using an EEG dataset of 800 recordings (30 seconds each) from 61 children with ADHD and 60 controls (ages 7–12), recorded via a 19-channel 10–20 system at 128 Hz during a visual attention task, our method achieved 73% accuracy. This demonstrates its promise as a reliable, non-invasive tool for ADHD diagnosis, advancing EEG-based deep learning in clinical applications.

KEYWORDS *: ADHD, EEG, CNN, LSTM, Neurodevelopmental disorders.*

INTRODUCTION

Attention Deficit Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder that affects approximately 5–7% of children globally [1, 2]. ADHD is characterized by inattention, hyperactivity, impulsivity and careless mistakes. (Fig.1) ADHD significantly impacts a child's academic performance, social interactions, and emotional well-being. If untreated, these issues can persist into adulthood, leading to long-term challenges such as poor career outcomes and relationship difficulties. Early and accurate diagnosis is essential to mitigate these effects and implement effective interventions such as behavioral therapy or medication [2, 3].

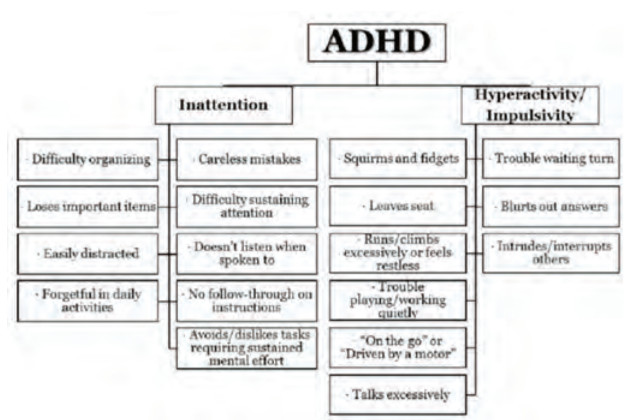


Fig. 1: Behavioral Symptoms of ADHD

Despite the critical need for accurate diagnosis, current clinical practices largely depend on subjective assessments, including interviews, behavioral checklists, and reports from parents and teachers. These methods are susceptible to observer bias and inconsistencies, often resulting in overdiagnosis or underdiagnosis. The reliance on subjective measures not only delays treatment but also fails to capture subtle diagnostic nuances. Addressing this gap requires developing objective tools capable of providing reliable and reproducible diagnostic outcomes [3, 4].

Electroencephalography (EEG) represents the electrical activity in the brain, categorized into different frequency bands. (Fig.2) Delta waves (0.5–4 Hz) are the slowest and occur during deep sleep, while theta waves (4–8 Hz) are linked to light sleep and creativity. Alpha waves (8–12 Hz) indicate relaxation, beta waves (12–30 Hz) are associated with active thinking and stress, and gamma waves (30–100 Hz) relate to learning and high-level cognition. These waves help study brain states, diagnose disorders, and understand neurological conditions. EEG, a non-invasive method for measuring brain activity, has emerged as a promising tool for ADHD diagnosis. Studies reveal that children with ADHD exhibit distinct patterns in EEG brainwave activity, such as an increased theta-to-beta power ratio, particularly in the frontal cortex—a region associated with attention regulation and impulse control [5, 6, 7]. These EEG biomarkers offer valuable insights into the neural underpinnings of ADHD and present an opportunity for leveraging automated techniques to enhance diagnostic precision.

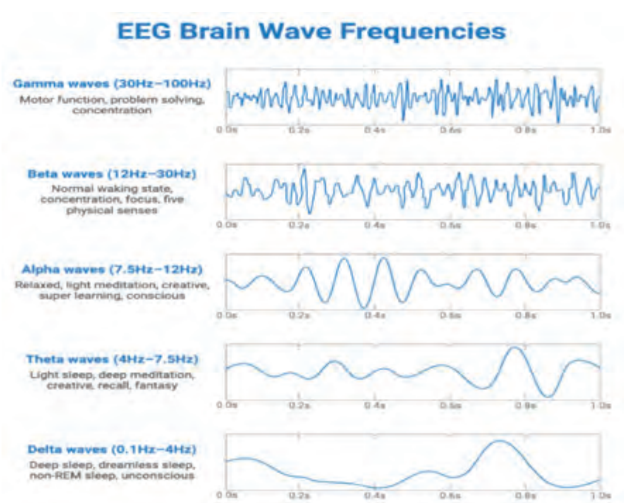


Fig. 2: EEG Brain Wave Frequencies

In recent years, advancements in artificial intelligence (AI) and deep learning (DL) have demonstrated immense potential in medical diagnostics. Convolutional Neural Networks (CNNs) specialize in extracting spatial patterns, while Long Short-Term Memory (LSTM) networks excel at capturing temporal dynamics in time-series data like EEG signals. Hybrid CNN-LSTM architectures combine these capabilities, enabling comprehensive spatial-temporal analysis of EEG data. These models can identify intricate patterns in neural activity, offering enhanced diagnostic accuracy compared to traditional statistical or machine learning methods [8, 9, 10].

This study focuses on designing a hybrid CNN-LSTM deep learning model to classify ADHD and non-ADHD cases using EEG data. (Fig.3) By automated feature extraction and classification, the system aims to reduce the subjectivity inherent in traditional diagnostic approaches. The integration of real-time processing capabilities further ensures timely insights, enhancing the efficiency of ADHD assessments [8, 11, 12].

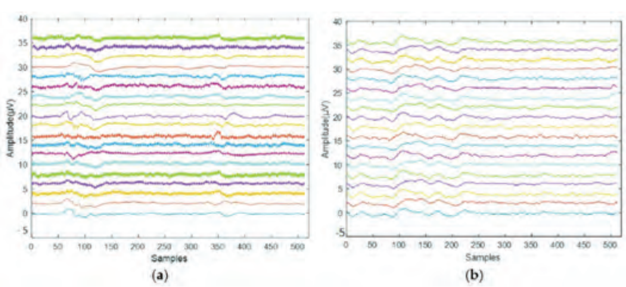


Fig. 3 EEG signals: (a) ADHD class and (b) Normal class

In addition to advancing diagnostic accuracy, this project seeks to contribute a reusable and adaptable framework for analyzing EEG data, potentially applicable to other neurological disorders. By leveraging state-of-the-art deep learning techniques, the proposed system holds the potential to revolutionize ADHD diagnosis, paving the way for more objective, consistent, and scalable solutions in clinical settings [8, 13, 14].

BACKGROUND

Recent advancements in neuroscience and computational methodologies have positioned EEG-based machine learning models as a cutting-edge approach to ADHD diagnosis. This literature review consolidates theoretical and empirical contributions in this domain.

To establish a comprehensive theoretical framework, review papers have extensively explored ADHD's neurobiological basis, the utility of EEG in brain signal analysis, and the incorporation of ML/DL techniques in healthcare. Noroozi and Ahmadi (2022) provide an in-depth analysis of automated ADHD diagnostics using EEG data, outlining common methods, challenges, and future directions [13]. Taspinar & Ozkurt (2024) reviews the use of resting state fMRI (rsfMRI) data and ML methods for ADHD detection, providing insights into the ADHD-200 database, and emphasizing the importance of pre-classification steps in the detection process[11]. Similarly, Zhang et al. (2022) focus on specific ML models' effectiveness and emerging trends in neuroimaging [35]. Collectively, these reviews underscore EEG's diagnostic potential due to its ability to identify atypical neural patterns, such as heightened theta-to-beta ratios, in ADHD patients. They also emphasize the role of ML in enhancing diagnostic accuracy and consistency.

Building on this theoretical base, implementation-oriented studies demonstrate the practical application of ML/DL models in classifying ADHD using EEG data. Zhang et al. (2022) implemented attention-based CNN-LSTM models to capture EEG data's spatial and temporal features, achieving high classification accuracy [30]. Bakhtyari and Mirzaei (2022) extended this approach using ConvLSTM networks combined with attention mechanisms, revealing dynamic connectivity patterns unique to ADHD [5]. Similarly, Alqudah and Hiasat (2021) proposed a hybrid CNN-LSTM framework, highlighting the importance of architectural optimization for enhanced diagnostic performance [12]. These studies exemplify the significant strides made in real-world implementation, offering insights into model training, evaluation, and generalizability.

Despite these advancements, challenges persist. Early research, such as that by Monastra et al. (1999) and Barry et al. (2003), faced limitations like small sample sizes and restricted tasks, which reduced the generalizability of findings [6, 38]. While more recent studies have achieved impressive accuracy rates—up to 99.75% with hybrid CNN-LSTM models (Bakhtyari & Mirzaei, 2022)—issues such as overfitting, inconsistent preprocessing, and computational demands remain barriers [5]. Furthermore, real-world applicability is hindered by limited dataset diversity and the complexity of advanced architectures, as noted in studies like ADHD-AID (Attallah, 2024) and Ulam's Spiral-based feature extraction models (Atila et al., 2023) [33, 39].

Future research must address these limitations by emphasizing the integration of scalable and computationally efficient deep learning frameworks. Efforts should also focus on diversifying datasets to include broader population groups, enabling models to capture the heterogeneity of ADHD. This dual approach holds the potential to transform EEG-based ADHD diagnostics, offering robust, real-time solutions for clinical implementation.

MATERIALS & METHODS

The dataset used in this study includes EEG recordings from 61 children with ADHD and 60 healthy controls (boys and girls, ages 7–12). Children in the control group had no history of psychiatric disorders, epilepsy, or high-risk behaviors. ADHD diagnosis followed DSM-IV criteria and was confirmed by an experienced psychiatrist. ADHD participants were on Ritalin for up to six months. [15]

EEG data was recorded using the standard 10–20 electrode placement system with 19 channels (Fz, Cz, Pz, C3, T3, C4, T4, Fp1, Fp2, F3, F4, F7, F8, P3, P4, T5, T6, O1, O2) at a sampling frequency of 128 Hz. Reference electrodes (A1 and A2) were located on the earlobes. The recording protocol included a visual attention task to assess cognitive deficits in ADHD children. Participants were shown pictures of cartoon characters and asked to count the number of characters in each image. The number of characters varied randomly between 5 and 16, and each image was displayed continuously until the child responded. The total EEG recording duration depended on the response speed of each participant.

Table 1 Data was segmented into 30-second intervals for analysis. This resulted in a total of 800 instances: 400 intervals for ADHD participants and 400 for healthy controls. These intervals were used to train and validate the model.

Group	Number of Subjects	Number of Intervals	Duration (Average)	Sampling Rate
ADHD Children	61	400	30 Seconds	128Hz
Healthy Controls	60	400	30 Seconds	128Hz

Problem Definition

EEG signals are multivariate time series captured across multiple channels. In this study, the dataset comprises $C=19$ channels (corresponding to EEG recording electrodes), and each recording spans $S=30$ seconds, sampled at a frequency of $F=128$ Hz.

The total number of samples per EEG channel is P=S×F=3840.

The recording signal from all channels at a specific time tt is denoted as xt RC, and the complete 30-second EEG data is represented as a matrix X, where one dimension corresponds to the 19 channels, and the other captures the time series across 30 seconds: X=[x1, x2 , … , xP]T.

The dataset contains a total of 800 labeled instances of these multivariate time series: 400 instances correspond to ADHD participants, and 400 to healthy controls. [15] These labeled instances serve as the foundation for developing a machine learning model capable of distinguishing between ADHD and healthy controls. The goal is to leverage this data to train an accurate and robust classification model, facilitating more objective ADHD diagnosis based on EEG signals.

Data preprocessing

EEG signals are highly prone to artifacts, which can occur during recording and impact data quality. [16] These artifacts are classified as external or internal. External artifacts, caused by environmental noise or experimental errors, were addressed using a Finite Impulse Response (FIR) high-pass filter with automatic order. This filter removed low-frequency noise below 4 Hz, often caused by sweating or neck and jaw movements. Power spectral density analysis revealed that gamma-band frequencies above 60–70 Hz and line noise between 45–55 Hz also affected the data. Fig. 4 exhibits the power spectral density of an instance of the dataset. To eliminate these artifacts, we filter the dataset with a low-pass filter at a frequency of 40 Hz with the same configuration in the high pass filter. [17]

Internal artifacts, such as eye movements and blinking, are common during EEG recordings and significantly reduce signal quality. Manual preprocessing is widely used to address these but requires expert knowledge, which limits scalability. Automatic methods like PREP [18] were tested but did not yield better results. Given the dataset's visual stimulus recording protocol, Independent Component Analysis (ICA) was used to isolate brain-related sources from artifacts. [19] IC Label [20] was then applied to detect and remove components related to eye movements with a probability above 0.9. These preprocessing steps were essential for improving the quality of the EEG signals and enhancing the performance of the subsequent analysis.

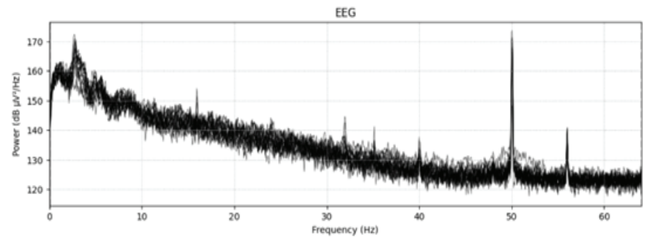


Fig. 4 Power spectral density of an instance of EEG recordings before filtering

Feature Extraction

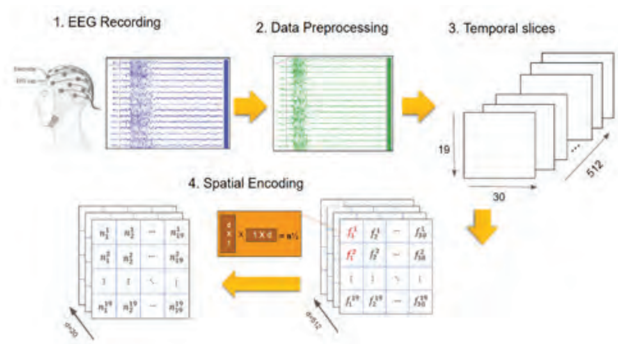


Fig. 5 The proposed feature extraction scheme

Fig. 5 represents a pipeline for processing EEG (electroencephalography) data for use in a machine learning or deep learning application. The process begins with EEG recording, where electrodes placed on the scalp measure electrical brain activity over time, capturing signals from multiple channels corresponding to different locations on the scalp. After recording, the signals undergo preprocessing to clean and prepare them for analysis. [21] This step includes filtering to remove noise or artifacts, re-referencing, and normalization to retain only meaningful signal patterns. Once the data is preprocessed, it is divided into smaller temporal slices, creating a sequence of short time windows, with each slice representing a fixed time segment of the EEG signals (e.g., 512 data points). These slices are essential for capturing time-dependent features, allowing the model to analyze temporal dynamics. Each temporal slice is then subjected to spatial encoding to organize the data based on the arrangement of EEG channels. This step involves mapping the EEG channels to a grid-like structure (19x19 in this case) to mimic the spatial layout of electrodes on the scalp. The grid is further processed using feature extraction or encoding, resulting in spatially encoded data represented as a matrix or tensor. This structured data retains information about the temporal and spatial relationships of the EEG signals.

Ultimately, this pipeline transforms raw EEG recordings into a format suitable for machine learning, emphasizing both temporal and spatial information, which can then be utilized by advanced algorithms to detect patterns, classify mental states, or predict outcomes. [22].

Model

This section introduces a deep neural network architecture combining convolution and Long Short-Term Memory (LSTM) layers. In the proposed model, the convolution layer captures spatial features among EEG channels, while the LSTM layer encodes temporal dependencies across time segments. Finally, the data passes through a dense layer for classification. The following subsections provide a detailed description of each of these layers.

Hybrid Deep Learning Model (CNN & LSTM)

In recent years, deep learning has revolutionized the field of medical diagnostics, enabling the analysis of large, complex datasets with high accuracy. Deep learning models, such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, are particularly effective in processing time-series data like EEG signals. The idea of LSTMs was proposed in 1997 to mitigate recurrent neural network problems [23] CNNs excel at extracting spatial features from EEG data, identifying local patterns within the signal,[24] while LSTMs are well-suited for capturing temporal dependencies, making them ideal for analyzing EEG data over time.

Hybrid models that combine CNNs and LSTMs have shown superior performance in medical applications, as they leverage both spatial and temporal characteristics of data.[25] In this project, we employ a hybrid CNN-LSTM model to process EEG signals, aiming to improve the accuracy and reliability of ADHD diagnosis. By automating the feature extraction and classification process, the model reduces the risk of human error and offers a scalable solution for clinical applications.

The current diagnostic process for ADHD is time-consuming and prone to variability between clinicians. Furthermore, the subjective nature of symptom observation can lead to inconsistent diagnoses. An automated system using EEG data and deep learning models offers a more standardized approach, providing consistent results across a range of different patients and clinicians. Automated analysis of EEG data can also handle large datasets, making it possible to identify subtle patterns that may be missed by human observers.

By integrating CNN and LSTM models, the system can process raw EEG data in real-time, offering immediate diagnostic insights. (Fig. 6) This is particularly valuable in clinical settings, where timely diagnosis is crucial for beginning treatment. Additionally, automation reduces the burden on clinicians, allowing them to focus on interpreting results and developing individualized treatment plans based on objective data. Through this project, we aim to develop a system that enhances the efficiency and accuracy of ADHD diagnosis, ultimately improving patient outcomes. [21]

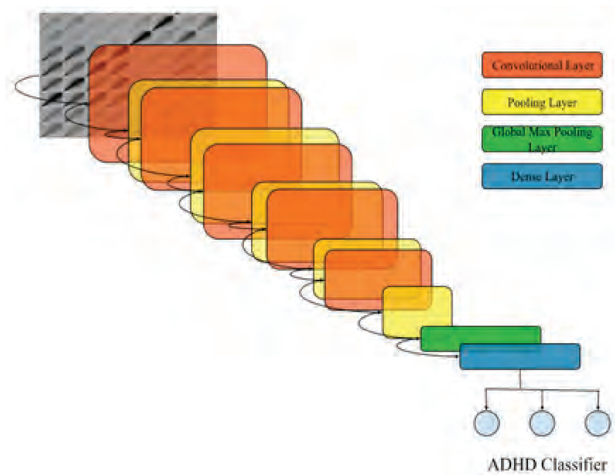


Fig.6 Proposed CNN & LSTM Model

RESULTS & DISCUSSIONS

In this section, we evaluate the proposed ADHD classification framework by performing experiments on the ADHD dataset. We compute four evaluation metrics:

- Accuracy = TP+TNTP+FP+FN+TN
- Precision = TP TP+FP
- Recall = TPTP+FN
- F1 − Score = 2*Precision*Recall Precision+Recall

In these formulas, TP, FP, TN, and FN stand for True Positive, False Positive, True Negative, and False Negative, respectively.

To evaluate the effectiveness of our LSTM-based model for ADHD detection using EEG signals, we employed the AUC-ROC (Area Under the Curve - Receiver Operating Characteristic) as a primary evaluation metric. This approach provides a comprehensive assessment of the model's ability to distinguish between ADHD and control cases across various classification thresholds.

The evaluation process began with data splitting and preprocessing. We divided the EEG data into training and test sets using a 70-30 split, ensuring a balanced representation of ADHD and control samples. To enhance the representation of minority classes, the SMOTE (Synthetic Minority Over-sampling Technique) was applied to the training data. Following this, the model was trained with the following hyperparameters: a batch size of 16, 30 epochs, and a learning rate of 0.0001. Early stopping, with a patience of 10 epochs, was utilized to prevent overfitting during training.

After training, we evaluated the model using the AUC-ROC. The ROC curve was generated to visualize the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity) across varying thresholds. The AUC score, representing the area under the ROC curve, served as a summary statistic to quantify the model's overall discriminatory ability. In addition to AUC-ROC, we calculated other metrics such as precision, recall, F1-scores, and a confusion matrix to gain further insights into the model's performance. These metrics provided a detailed understanding of the model's strengths and weaknesses in detecting ADHD. [26]

The AUC score for the LSTM model was 0.48, indicating less accurate overall discrimination between ADHD and control cases. This score is close to random guessing, which is an AUC of 0.5.

The ROC curve showed a high recall for ADHD cases but inaccurate specificity for control cases, reflecting the imbalance in classification performance.

The Confusion matrix in Fig.7 shows:

Out of 5 control cases, only 1 was correctly classified.

All 10 ADHD cases were correctly identified, showing high sensitivity for ADHD

Comparison of Metrics

Table 2. The detailed evaluation metrics for the test data are as follows:

Metric	ADHD	Control
Precision	0.71	1.00
Recall (Sensitivity)	1.00	0.20
F1-score	0.83	0.33

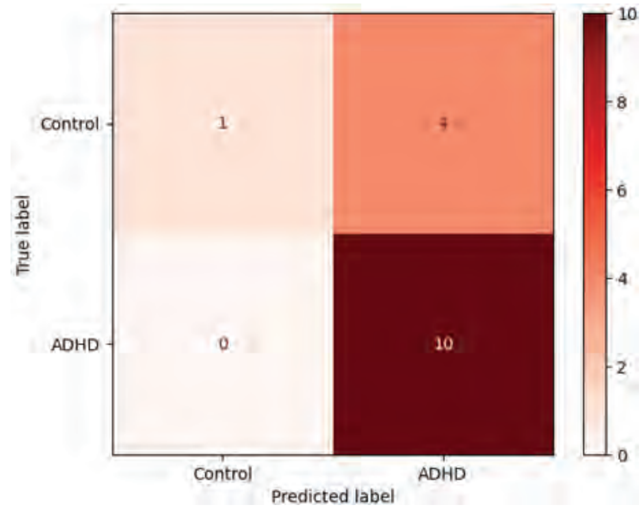


Fig. 7. Confusion Matrix of ADHD and Control Children cases

This section evaluates how key parameters affect the LSTM model's performance in ADHD detection using EEG signals. Optimal hyperparameters were chosen, focusing on raw EEG signals (4–40 Hz) to avoid noise from higher frequencies. Testing showed that this range effectively captured key features across theta, alpha, beta, and low-gamma bands for accurate classification. Temporal segmentation analysis revealed that smaller windows increased computational costs without boosting accuracy, while larger windows reduced performance due to lower temporal resolution. As shown in Table 1, The best results were achieved with a 1-second window size (512 samples) and no overlap, balancing efficiency and accuracy. A batch size of 16 was optimal for training, offering a good trade-off between computational efficiency and model convergence. And in Table 2, The comparison of matrix is shown between ADHD and Control class.

In comparison to a study [22] reporting a remarkable 97.75% accuracy using LASSO-regularized biomarkers for ADHD classification, our research achieved 73% accuracy utilizing a deep learning approach on an EEG dataset with 800 recordings from children aged 7–12. While the study highlights the strengths of feature optimization techniques like LASSO, our work emphasizes the promise of raw EEG-based deep learning for non-invasive ADHD diagnosis. Despite the lower accuracy, our method advances the application of deep learning in clinical settings and demonstrates potential for further refinement with larger datasets and optimized feature selection techniques.

CONCLUSIONS

In conclusion, the integration of machine learning models, particularly LSTM-based architectures, for detecting ADHD through EEG signals has shown promising results in terms of accuracy and efficiency. The study highlights that carefully selecting the relevant frequency bands (4–40 Hz) and optimizing temporal windowing parameters, such as using a 512-sample window, contribute significantly to achieving reliable classification performance. Moreover, fine-tuning hyperparameters like batch size and filter size ensures that the model not only performs well but does so in a computationally efficient manner.

While the current model configuration provides excellent results, the real-world application of these techniques still faces challenges, such as the need for improved generalization across diverse ADHD subtypes and populations. Future work should explore the incorporation of advanced signal processing methods, the inclusion of larger and more varied datasets, and the development of interpretable models that clinicians can use with confidence. As these techniques evolve, they have the potential to transform ADHD diagnosis, enabling earlier and more accurate detection, ultimately leading to better patient outcomes.

FUTURE WORK

Future work on the LSTM-based ADHD detection model should focus on optimizing hyperparameters and enhancing preprocessing techniques. The current configuration, which uses raw EEG signals within the 4–40 Hz range, a 512-sample window, and a batch size of 16, has demonstrated good accuracy. However, implementing more advanced signal processing methods, such as wavelet transforms, could refine feature extraction and reduce noise from higher frequency bands, which is essential for improving diagnostic accuracy.

Adaptive windowing techniques, which adjust dynamically based on signal characteristics, could enhance temporal resolution while minimizing computational costs. Additionally, exploring complex model architectures like hybrid CNN-LSTM networks and attention mechanisms could improve the model's ability to detect subtle patterns in EEG data.

Expanding the dataset to include diverse ADHD subtypes and age groups is critical for increasing the model's generalizability. Longitudinal datasets could provide deeper insights into ADHD's progression, enabling a better understanding of its dynamic nature. For real-world applications, the model must prioritize explainability and real-time processing to facilitate clinical integration and improve the ease of ADHD screening and monitoring. Such advancements would enhance the model's practical utility and ensure its effectiveness across diverse settings.

By addressing these areas, researchers can refine EEG-based ADHD detection systems, ultimately contributing to more accurate, accessible, and clinically relevant solutions for ADHD management.

ACKNOWLEDGMENT

We sincerely thank Mr. Ichhanshu Jaiswal, our project guide, for his invaluable support and guidance. We appreciate the faculty of Computer Science & Engineering (Data Science) for their encouragement and constructive feedback. Lastly, we acknowledge the contributions of researchers in this field and thank our peers, family, and friends for their support.

REFERENCES

1. Polanczyk, G., de Lima, M. S., Horta, B. L., Almeida, R. M., & Rohde, L. A. (2014). The worldwide prevalence of ADHD: A systematic review and metaregression analysis. American Journal of Psychiatry, 171(6), 408-416.
2. Loo, S. K., & Barkley, R. A. (2017). Diagnosis and treatment of attention-deficit/hyperactivity disorder in children and adolescents. The Journal of Pediatrics, 190, 239-245.
3. American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders (5th ed.) https://www.psychiatry.org/psychiatrists/practice/dsm
4. Obeid, M., & Ahmed, F. (2020). Subjectivity in ADHD diagnosis: A systematic review. Journal of Attention Disorders, 24(6), 845-856.
5. Bakhtyari, M., & Mirzaei, M. (2022). EEG markers in children with ADHD: A systematic review. Neuroscience Letters, 759, 136068.
6. Barry, R. J., Clarke, A. R., & McCarthy, R. (2003). EEG differences between children with ADHD and normal children during a continuous performance task. Clinical Neurophysiology, 114(6), 1045-1056.
7. Arns, M., de Beurs, E., Spinhoven, P., & Ven, H. (2016). The efficacy of neurofeedback in ADHD: A meta-analysis. Psychological Bulletin, 142(6), 573-589.
8. Alim, M. A., & Imtiaz, M. A. (2023). EEG-based ADHD detection: A review of deep learning approaches. IEEE Access, 11, 1489-1503.

9. Chen, Q., Xu, L., & Zhang, W. (2021). Hybrid CNN-LSTM model for ADHD detection using EEG signals. Journal of Neural Engineering, 18(5), 056008.

10. Chen, M., Zhang, Y., Chen, H., & Liu, Y. (2021). Attention mechanism-based CNN-LSTM for EEG signal classification in children with ADHD. Journal of Neural Engineering, 18(4), 046003.

11. Gurcan Taspinar & Nalan Ozkurt (2024). A review of ADHD detection studies with machine learning methods using rsfMRI data. Wiley Analytical Science Journal, NMR in Biomedicine, Volume37, Issue8. https://doi.org/10.1002/nbm.5138

12. Alqudah, A. M., & Hiasat, M. H. (2021). EEG signal classification for ADHD detection using a hybrid deep learning model. IEEE Access, 9, 123145-123154.

13. Noroozi, N., & Ahmadi, A. (2022). Automated ADHD diagnosis using machine learning and EEG data: A review. Journal of Medical Systems, 46(3), 1-15.

14. Martínez González, C. L., Martínez Ortiz, E. J., Moreno Escobar, J. J., & Durand Rivera, J. A. (2022). Attention deficit and hyperactivity disorder classification with EEG and machine learning. Elsevier.

15. Nasrabadi A. M., Allahverdy A., Samavati M. & Mohammadi M.R. (2020) EEG data for ADHD / Control children, IEEE Dataport, doi: https://dx.doi.org/10.21227/rzfh-zn36.

16. Radüntz T., Scouten J., Hochmuth O. & Meffert B. (2017) Automated EEG artifact elimination by applying machine learning algorithms to ICA-based features, J. Neural Eng. 14(4): 046004.

17. Sweeney K.T., Ward T.E. & McLoone S.F. (2012) Artifact removal in physiological signals—Practices and possibilities, IEEE Trans. Inf Technol. Biomed. 16(3): 488–500,

18. Bigdely-Shamlo N., Mullen T., Kothe C., Kyung-Min S.U. & Robbins K.A. (2015) The PREP pipeline: standardized preprocessing for large-scale EEG analysis, Front. Neuroinformatics, 9: 16.

19. Dimigen O. (2020) Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments, Neuroimage 207:116117.

20. Pion-Tonachini L., Kreutz-Delgado K. & Makeig S. (2019) ICLabel: an automated electroencephalographic independent component classifier, dataset, and website, NeuroImage 198: 181–197.

21. García-Ponsoda S., Maté A. & Trujillo J. (2024) Refining ADHD diagnosis with EEG: The impact of preprocessing and temporal segmentation on classification accuracy. Computer in Biology and Medicine, 183:109305.

22. Alkahtani, H., Aldhyani, T.H.H., Ahmed, Z.A.T. & Alqarni, A.A. (2023) Developing System-Based Artificial Intelligence Models for Detecting the Attention Deficit Hyperactivity Disorder. Mathematics, 11: 4698.

Diabetes Prediction

Sahil Salunke
CSE and Data Science
Vidyavardhini's College of Engineering & Technology
Vasai (W), Palghar, Maharashtra
✉ sahil.224506101@vcet.edu.in

Ichhanshu Jaiswal
CSE and Data Science
Vidyavardhini's College of Engineering & Technology
Vasai (W), Palghar, Maharashtra
✉ Ichhanshu.jaiswal@vcet.edu.in

Leena Raut
CSE and Data Science
Vidyavardhini's College of Engineering & Technology
Vasai (W), Palghar, Maharashtra
✉ leena.raut@vcet.edu.in

Soham Shivpuje
CSE and Data Science
Vidyavardhini's College of Engineering & Technology
Vasai (W), Palghar, Maharashtra
✉ soham.224586105@vcet.edu.in

ABSTRACT

Every year, lakhs of new cases of diabetes are observed worldwide, making it a significant global health problem. Improving patient outcomes and modify treatment options depend on early identification and precise diagnosis. In the world of Artificial Intelligence, Machine Learning has proven to be a most useful tool for recognizing medical conditions. By analyzing various clinical and patient data, it has the potential to help healthcare professionals in classifying diabetes cases more accurately. This study denotes the practical application of machine learning in the medical field by developing a prediction- based model for diabetes using various models. Early identification of diabetes as well as crucial effective management and better patient outcomes is the primary focus of our research initiative. The main objective of this project is to create a machine learning model that can effectively predict the possibility of diabetes onset, thereby reducing dependence on time-consuming and subjective human estimation, ultimately improving diagnostic consistency and efficiency.

***KEYWORDS** : Diabetes, Medical diagnosis, Machine learning, Artificial intelligence, Patient outcomes, Predictive model, Clinical data.*

INTRODUCTION

Diabetes is a common and confirmed health condition that affects lakhs of individuals globally, crossing gender lines and age groups. According to estimates from the World Health Organization, there are approximately 41.5 Crore people living with diabetes. It causes a significant health risk and is a leading cause of various obstacles and death. The current diagnostic methods for diabetes include blood tests, glucose monitoring, and clinical evaluations, which, while effective, may require time and are subjective to interpretation.

Timely identification of diabetes is very important in improving the patient outcomes and managing the conditions effectively. Early detection allows for quick intervention and better disease management, which can enhance the quality of life for those affected. Additionally, the application of machine learning algorithms can bring fairness and consistency to diabetes diagnosis. This study presents the practical use of machine learning in healthcare technology, specifically in addressing the real-world issues related to diabetes prediction.

This ambition not only highlights the potential of deep learning methods but also highlights the importance of ethical considerations and data quality throughout the development process.

PROBLEM STATEMENT

A critical challenge facing the healthcare sector is the early prediction of diabetes, which is the central focus of this study. Diabetes is a usual and possible life-threatening condition that affects individuals worldwide. Early detection is essential for improving the patient outcomes and easing the effective management. However, current diagnostic methods are often laborious, open to human error, and can lead to delayed diagnosis, primarily relying on subjective interpretations of clinical data and laboratory tests.

To address these issues, this research aims to develop an automated predictive system that can accurately show the possibility of diabetes onset by using the power of machine learning and data analysis. The project's core objective is to provide a strong and reliable machine learning solution for early diabetes prediction, thus enhancing healthcare outcomes and the overall quality of life for individuals at risk of developing the conditions.

OBJECTIVES

Data Collection and Preprocessing: Gather a complete dataset comprising both clinical and patient-related information regarding to diabetes prediction. Carefully preprocessing and cleaning the data to ensure its validity for machine learning analysis.

Model Development: Use a machine learning model, such as Support Vector Classifier, to create a predictive system for diabetes. Train the model using the collected dataset to enable accurately predicting of the possibility of diabetes onset.

Provide an accurate and efficient machine learning model capable of predicting the risk of diabetes, thus helping in early identification and proactive management.

Reduce the time and effort required for healthcare professionals to check the risk of diabetes and make informed decisions, thereby improving the efficiency of diabetes prediction and care.

Education and Awareness: Improve public, patient, and healthcare professional's understanding of the benefits and limitations of machine learning in the prediction, encouraging the informed decision-making and energetic health management.

SCOPE

Medical Applications: The primary objective of this project is to establish a strong automated tool for diabetes prediction, which holds promise for wide applications in the field of medical diagnostics. The system's flexibility allows for potential expansion into the analysis of various medical data types beyond diabetes, expanding its help to the diagnosis of other health conditions.

Integration with Healthcare Systems: The architecture of the system is designed to smoothly integrate with existing healthcare infrastructures, allowing healthcare professionals to easily include AI-driven predictive tools into their regular diagnostic protocols.

Global Applicability: The predictive technique for diabetes has universal importance, serving as a valuable resource worldwide. Its potential benefits expand to regions with changing healthcare resources, contributing to improved healthcare availability and the early identification of diabetes cases.

Opportunities for Collaboration: This project encourages collaboration among healthcare providers, AI experts, and medical institutions, promoting an environment for associative teamwork. Such partnerships hold the potential to guide advancements in the field of medical artificial intelligence, benefiting both research and patient care.

LITERATURE SURVEY

Diabetes is a widespread and confirmed health condition that has a remarkable impact on individuals and healthcare systems worldwide. The development of effective predictive models for diabetes using machine learning has become an expanding area of research. This literature survey provides an overview of recent studies that explore various aspects of diabetes prediction with a focus on machine learning algorithms and techniques.

Bhat et al. (2023) gave a risk evaluation and prediction framework for diabetes using machine learning algorithms. This approach combines the advanced machine learning techniques to evaluate diabetes risk based on a range of health measures. [1]

Chang et al. (2022) conducted an estimation of machine learning models and algorithms for early prediction and diagnosis of diabetes using health measures. Their study provides awareness into the selection and performance of different machine learning approaches in diabetes prediction.[2]

Khanam and Foo (2021) conducted a comparative study of machine learning algorithms for diabetes prediction. Their research offers a valuable comparison of various algorithms to evaluate researchers and interpreters in algorithm selection.[3]

Jiang et al. (2023) developed a diabetes risk prediction model based on community research data using machine learning. This study demonstrates the potential of real-world data in constructing a strong prediction model for diabetes.[4]

Jangili et al. (2023) focused on machine learning based early biomarker prediction for type 2 diabetes related with coronary artery diseases. Their study highlights the importance of identifying high-risk individuals for diabetes and related complications.[5]

Aguilera-Venegas et al. (2023) compared and fined tuned various machine learning algorithms to predict the type 2 diabetes mellitus. This research provides the significance of algorithm optimizations for improved prediction accuracy.[6]

Kokkorakis et al. (2023) developed effective set of question-based prediction models for type 2 diabetes across diverse population. Their work highlights the importance of considering population diversity in diabetes prediction.[7]

Qi et al. (2023) introduced "Kfpredict," an ensemble learning prediction framework for diabetes. Their model makes use of the mixture of key features to improve prediction accuracy.[8]

Febrian et al. (2023) explored diabetes prediction using supervised machine learning. Their research shows the application of traditional supervised learning techniques for diabetes prediction.[9]

Reza et al. (2023) improved support vector machine (SVM) performance for type II diabetes prediction with an improved non-linear kernel. Their work provides understanding into the optimizing SVM for diabetes prediction.[10]

Ketkar and Gawade (2022) introduced a decision support system for selecting the most suitable machine learning algorithm in healthcare. This system allows users to choose the best-fit algorithm based on specific parameters and requirements.[11]

Belsti et al. (2023) compared machine learning and standard logistic regression-based prediction models for with child diabetes in a diverse population. Their study offers insights into predicting with child diabetes using different modeling techniques.[12]

Mora et al. (2023) investigated the use of machine learning algorithms to predict the onset of diabetes-related problems after a diabetes diagnosis. Their work is particularly relevant for long-term diabetes management.[13]

Rastogi and Bansal (2023) presented a diabetes prediction model using data mining techniques. Their research focuses on the application of data mining for diabetes prediction.[14]

Mennickent et al. (2022) evaluated machine learning-based models for with child diabetes mellitus prediction before 24– 28 weeks of pregnancy. This study provides an overview of the state of the art in predicting with child diabetes.[15]

In conclusion, the field of diabetes prediction using machine learning is distinct and effective, with various studies exploring different algorithms, datasets, and approaches. This literature survey serves as a complete resource to guide our project on Diabetes Prediction System, helping to incorporate the latest findings and methodologies into our research.

PROPOSED SYSTEM

The suggested system presents an advanced method for summarizing data that combines deep learning and artificial intelligence approaches. This new method generates a brief, relatedly relevant summaries in response to the difficulty of managing large amounts of numerical and written data. Scalability is its main advantage since it allows to analyze and summarize a large number of datasets quickly, which is important in today's data-rich environment.

This system makes attempt to provide an important tool for data analysis to a range of industries, including businesses, healthcare, research, and education, by mixing deep learning models, data mining, and natural language processing. It is a useful addition to the field of data summarization because of its capacity to transform decision-making steps and make useful conclusions from the complex data. The system's architecture and structure will be fully examined in the next sections, which will also present how this new strategy is assured to transform the data summarization techniques.

Algorithm and Process Design

- Data Collection and Preprocessing

To ease the development and evaluation of our AI model, we gathered a considerable dataset comprising of clinical and patient-related information related to diabetes prediction. Careful preprocessing techniques were applied to ensure the data consistency and high-quality data that would improve the model's performance.

- Model Architecture

We designed a specialized machine learning model, customized for diabetes prediction. This model is trained to show the possibility of diabetes onset, providing healthcare professionals with valuable diagnostic information.

- Data Augmentation

To overcome the challenges related with limited data, we made use of data augmentation methods to artificially expand our dataset. This augmentation technique remarkably improved the model's ability to conclude across various range of clinical scenarios, improving its predictive accuracy.

- Training and Evaluation

We used a one-cycle learning rate policy to train the model, achieving fierce levels of accuracy. To validate the model's diagnostic accuracy, we conducted an important evaluation, utilizing the performance metrics such as accuracy, precision, recall, and the F1 score.

- Real-time Processing

Our project places a strong importance on real-time diagnostic capabilities. The model's speed and efficiency are crucial in giving rise to decision-making, an important feature for clinical applications, especially in the field of diabetes prediction and management.

Details of Hardware & Software

Hardware Specifications

- System type: x64-based processor, 64-bit operating system.
- Memory (RAM) installed: 16.00 GB (15.7 GB Usable)
- Total size of Hard disk: 512 GB

Software Specifications

- Operating System: Because Python is cross-platform, the code can run on various operating systems, such as Microsoft Windows 10, Linux, and macOS.
- Integrated Development Environment (IDE): The code is used with a Python IDE that is comparable to Jupyter Notebook. As a result, selecting an IDE is open-ended. While PyCharm is a good choice, we can also use alternative Python IDEs like Google Colab, Visual Studio Code, or Jupyter Notebook.
- Libraries and Frameworks: The project utilizes several Python libraries and frameworks, including:

1. Scikit-Learn
2. Matplotlib
3. Pickle
4. NumPy
5. Pandas
6. Seaborn
7. Flask – Web Framework

- Python is the primary programming language used for data processing, machine learning, and deep learning, and it is the primary language used in this project.

Analysis

Dataset: We used the "diabetes.csv" dataset from Kaggle, which contains features (independent variables) related to patients' health such as Age, BMI, Insulin, etc. and a binary target variable Outcome (0 for non-diabetic and 1 for diabetic). It's important to make sure that the dataset is clean and properly preprocessed before using it into the machine learning models.

Algorithms: We experimented with a variety of machine learning algorithms. Here are some insights into the algorithms that we tried:

- Logistic Regression: It's a simple yet effective binary classification algorithm. It didn't perform as well as the SVC, due to the dataset's distribution or linearity.
- K Nearest Neighbors (KNN): KNN is a non-parametric algorithm that considers the "kth" nearest data points to make predictions. Its performance might be influenced by the choice of "k" and the distance metric. It seems to be well-suited for our dataset.
- Support Vector Classifier (SVC): SVC is known for its ability to handle complex decision boundaries.
- Naive Bayes: Naive Bayes is a probabilistic algorithm based on Bayes' theorem. Its assumptions might not hold in all datasets, potentially affecting its performance.
- Decision Tree: Decision trees are interpretable and can capture non-linear relationships. The accuracy might depend on the depth and other hyperparameters of the tree.
- Random Forest: Random Forest is an ensemble method based on decision trees. It's known for its robustness and can often produce high accuracy.

Model Selection: Selecting the best-performing model is a good start, but it's also important to consider the compromises, model interpretability, and any limitations for deployment. Other factors to consider might include computational complexity, training time, and the need for real-time predictions. Here, the K-Nearest Neighbors gives the best accuracy for the model.

Feature Importance: Analyzing the importance of features in the best model. Some algorithms provide feature importance scores, which can help to understand which variables are most influential in making predictions.

Evaluation: Accuracy is a useful metric, but it's important to consider other evaluation metrics as well, especially when dealing with imbalanced datasets. We also look at precision, recall, F1-score, and the ROC curve to evaluate the model's performance completely. Furthermore, considering the cross- validation to ensure the model's generalization ability.

Future Work: To improve the model further, we would explore more advanced techniques such as gradient boosting, neural networks, or ensemble methods.

RESULTS

Data Visualization

Counterplot

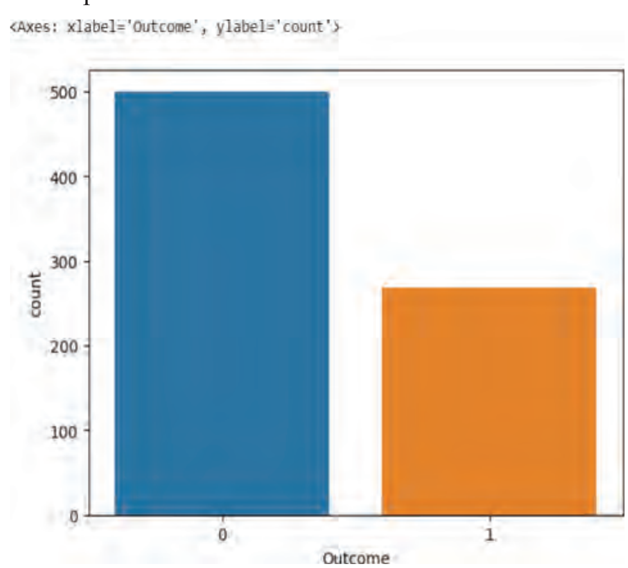


Fig. 1. Counterplot

The counterplot tells us that the dataset is imbalanced, as the number of patients who don't have diabetes is more than those who do.

Correlation Heatmap

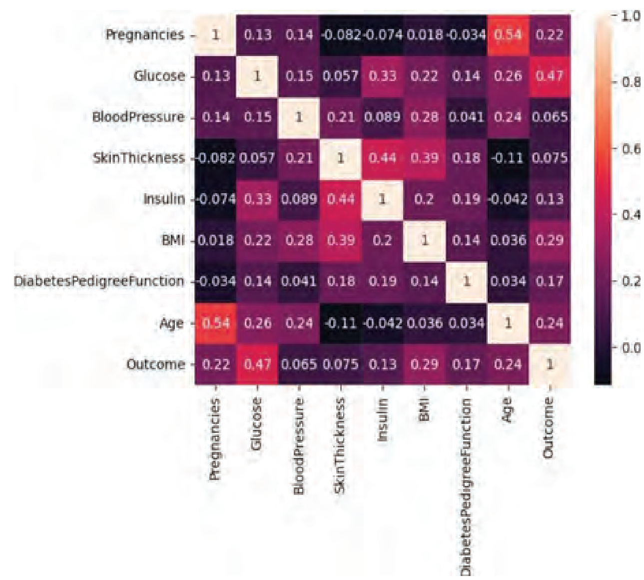


Fig. 2. Correlation Heatmap

The Correlation Heatmap tells us that there is a high correlation between Outcome and Glucose, BMI, Age and Insulin. We can select these features to accept input from the user and predict the outcome.

Model Evaluation

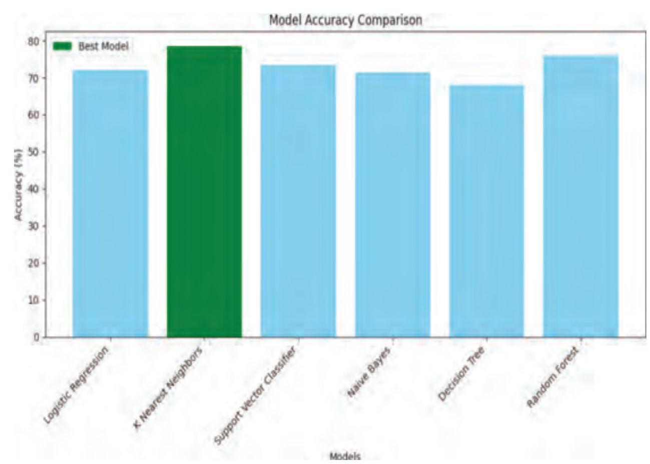


Fig. 3. Comparison between various model's evaluation

We evaluated the accuracy of various different models for finding the best possible outcome like Logistic Regression with 72%, K-Nearest Neighbors with 78%, Support Vector Classifier with 73%, Naive Bayes with 71.85%, Decision Tree with 68% and Random Forest Classifier with 76%. Out of which the K-Nearest Neighbors gave the best accuracy for our system to be most accurate and efficient.

Web page Layout and Predicted Output

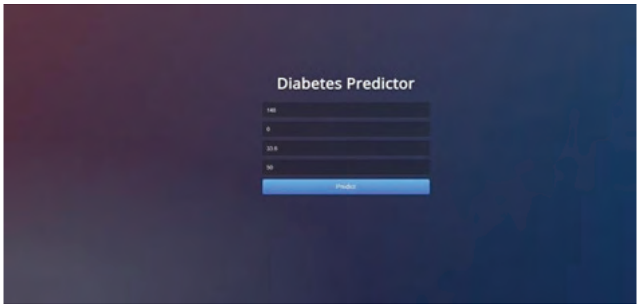


Fig. 4. Web Page design

The above figure (Fig. 4) shows the design of our Diabetes Prediction System. It takes the inputs from the user. The inputs to be entered by the user are Glucose, Insulin, BMI and the Age of the user.

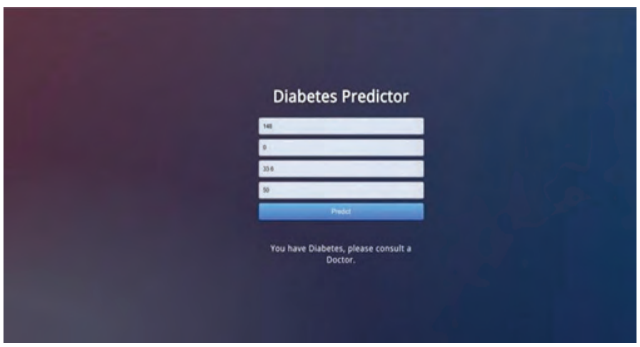


Fig. 5. Output predicted

The above figure (Fig. 5) shows the output for the entered inputs by the user.

CONCLUSION AND FUTURE WORK

In conclusion, we analyzed various models by comparing various machine learning algorithms and selecting the K-Nearest Neighbors as the best-performing model.

The goal of the Diabetes detection is a critical task, by using deep learning and machine learning approaches. The project's efficiency and dependability are improved by its well-structured training steps, evaluating methods, and data pipelining.

Future work

1. Integration of Additional Methods: In order to improve the classification accuracy, future study may involve integrating a number of other methods/ features, including ultrasound, mammography, and clinical data.
2. Better Data Augmentation: Using more complex data augmentation strategies, such as domain adaptation or generative adversarial networks (GANs), could increase the resilience of our model.
3. Interpretability: Using methods for model understandability such as Grad-CAM or SHAP might reveal which areas of a picture that are useful for making predictions.
4. Deployment and User Interface: Completely implemented the Streamlit integration to provide patients and medical professionals with an automated web-based interface.
5. Real-time diabetes diagnosis in clinical settings can be improved by scaling up and optimizing the project for faster processing.

ACKNOWLEDGMENT

We would like to extend our sincere gratitude to the individuals who have been involved in the successful completion of our project. Our utmost gratitude goes to our respectful Guide, Prof. Leena Raut, whose expertise, guidance, and steady support have been invaluable throughout this research journey. Our friends and family deserve special recognition for their steady support, patience, and understanding during this demanding work. Their encouragement and belief in our abilities were a consistent source of motivation. We also acknowledge the contributions of all the academicians who served as supervisors and evaluators. Their feedback, insights, and evaluation played a significant role in shaping our project. In summary, this project's successful completion was made possible through the collaborative efforts and support of our mentors, institution and evaluators. We are deeply grateful to each one of them.

REFERENCES

1. Bhat, Salliah Shafi, Madhina Banu, Gufran Ahmad Ansari, and Venkatesan Selvam. "A risk assessment and prediction framework for diabetes mellitus using machine learning algorithms." Healthcare Analytics (2023): 100273.
2. Chang, V., Ganatra, M.A., Hall, K., Golightly, L. and Xu, Q.A., 2022. An assessment of machine learning models and algorithms for early prediction and diagnosis of diabetes using health indicators. Healthcare Analytics, 2, p.100118.
3. Khanam, J.J. and Foo, S.Y., 2021. A comparison of machine learning algorithms for diabetes prediction. Ict

Express, 7(4), pp.432-439.

4. Jiang, L., Xia, Z., Zhu, R., Gong, H., Wang, J., Li, J. and Wang, L., 2023. Diabetes risk prediction model based on community follow-up data using machine learning. Preventive Medicine Reports, 35, p.102358.

5. Jangili, Shraddha, Hariprasad Vavilala, Gopi Sumanth Bhaskar Boddeda, Suryanaryana Murty Upadhyayula, Ramu Adela, and Srinivasa Rao Mutheneni. "Machine learning-driven early biomarker prediction for type 2 diabetes mellitus associated coronary artery diseases." Clinical Epidemiology and Global Health (2023): 101433.

6. Aguilera-Venegas, G., López-Molina, A., Rojo-Martínez, G. and Galán-García, J.L., 2023. Comparing and tuning machine learning algorithms to predict type 2 diabetes mellitus. Journal of Computational and Applied Mathematics, 427, p.115115.

7. Kokkorakis, Michail, Pytrik Folkertsma, Sipko van Dam, Nicole Sirotin, Shahrad Taheri, Odette Chagoury, Youssef Idaghdour et al. "Effective questionnaire-based prediction models for type 2 diabetes across several ethnicities: a model development and validation study." EClinicalMedicine 64 (2023).

8. Qi, Huamei, Xiaomeng Song, Shengzong Liu, Yan Zhang, and Kelvin KL Wong. "Kfpredict: an ensemble learning prediction framework for diabetes based on fusion of key features." Computer Methods and Programs in Biomedicine 231 (2023): 107378.

9. Febrian, Muhammad Exell, Fransiskus Xaverius Ferdinan, Gustian Paul Sendani, Kristien Margi Suryanigrum, and Rezki Yunanda. "Diabetes prediction using supervised machine learning." Procedia Computer Science 216 (2023): 21-30.

10. Reza, Md Shamim, Umme Hafsha, Ruhul Amin, Rubia Yasmin, and Sabba Ruhi. "Improving SVM performance for type II diabetes prediction with an improved non-linear kernel: Insights from the PIMA dataset." Computer Methods and Programs in Biomedicine Update 4 (2023): 100118.

11. Ketkar, Yashodhan, and Sushopti Gawade. "A decision support system for selecting the most suitable machine learning in healthcare using user parameters and requirements." Healthcare Analytics 2 (2022): 100117.

12. Belsti, Yitayeh, Lisa Moran, Lan Du, Aya Mousa, Kushan De Silva, Joanne Enticott, and Helena Teede. "Comparison of machine learning and conventional logistic regression-based prediction models for gestational diabetes in an ethnically diverse population; the Monash GDM Machine learning model." International Journal of Medical Informatics 179 (2023): 105228.

13. Mora, Toni, David Roche, and Beatriz Rodríguez-Sánchez. "Predicting the onset of diabetes-related complications after a diabetes diagnosis with machine learning algorithms." Diabetes Research and Clinical Practice 204 (2023): 110910.

14. Rastogi, Rashi, and Mamta Bansal. "Diabetes prediction model using data mining techniques." Measurement: Sensors 25 (2023): 100605.

15. Mennickent, Daniela, Andrés Rodríguez, Marcelo Farías-Jofré, Juan Araya, and Enrique Guzmán-Gutiérrez. "Machine learning-based models for gestational diabetes mellitus prediction before 24–28 weeks of pregnancy: a review." Artificial Intelligence in Medicine (2022): 102378.

Advance AI-Driven Real-Time Detection and Monitoring the Presence of Living Entities on Railway Stations for Enhanced Safety

Sonali Bhiwandkar
Computer Science and Engineering (DS)
Vidyavardhini's College of Engineering & Technology
Vasai (W), Palghar, Maharashtra
✉ sonali.s221386201@vcet.edu.in

Dhanashree Thakur
Computer Science and Engineering (DS)
Vidyavardhini's College of Engineering & Technology
Vasai (W), Palghar, Maharashtra
✉ dhanashree.213058205@vcet.edu.in

Odilia Gonsalves
Computer Science and Engineering (DS)
Vidyavardhini's College of Engineering & Technology
Vasai (W), Palghar, Maharashtra
✉ odilia.gonsalves@vcet.edu.in

Ramchandra Gawade
Computer Science and Engineering (DS)
Vidyavardhini's College of Engineering & Technology
Vasai (W), Palghar, Maharashtra
✉ ramchandra.s221416109@vcet.edu.in

ABSTRACT

Safety must be given top priority in railway stations and tracks since accidents frequently occur when unauthorized people access restricted areas. An AI-powered real-time detection system for tracking the existence of living things on railroad tracks and platforms is presented in this study. The technology analyzes live video streams and uses YOLOv5, a cutting-edge object detection algorithm, to identify people in real time. The technology automatically sounds alerts when someone approaches designated safety zones close to the tracks, greatly accelerating reaction times and lowering the possibility of human error.

The system is built to function well in a variety of environmental circumstances, including various illumination and weather conditions, guaranteeing reliable operation in a range of railway situations. It reduces the need for manual intervention by providing dependable monitoring and timely notifications with an accuracy rate of 90%–95%. The system lessens the workload for railroad employees while improving overall safety by automating the detection and alarm procedure. This study offers a scalable and effective approach to real-time surveillance and accident prevention by introducing a novel and workable strategy for enhancing railway safety through AI-powered automation. Furthermore, no significant modifications or extra hardware are needed to incorporate the system into the current railway infrastructure. This smooth integration guarantees that the system can be expanded and implemented in multiple geographic regions.

KEYWORDS *: AI-driven, Real-time detection, YOLOv5, Railway safety, Object detection, Video surveillance, Railway tracks, Deep learning, Alert system, Real-time monitoring, Passenger safety, Train platform monitoring.*

INTRODUCTION

Railway operators have long placed a high premium on maintaining safety in their systems since it has a direct bearing on the welfare of passengers, operational effectiveness, and general transportation safety. One of the biggest and avoidable hazards among the many risks connected to train terminals is illegal access to railroad lines. Accidents that happen when people unintentionally visit these dangerous locations frequently have disastrous results, including fatalities and severe injuries. Conventional safety methods, which mostly depend on physical barriers and human supervision, have not been enough to solve this enduring issue. These incidents frequently result from human error, inadequate monitoring coverage, and a lack of prompt action in response to safety violations [1].

The use of artificial intelligence (AI) and machine learning (ML) technologies to monitor and control railway safety has attracted a lot of attention recently due to the growing demand for safer and more effective railway systems. In particular, combining object detection algorithms with real-time video surveillance presents a viable way to automatically identify safety threats and take prompt action. Even in complicated and dynamic contexts, these technologies offer constant monitoring and reduce human mistake. Numerous studies have shown how well AI-based systems work to automate real-time object and person recognition, with noticeable gains in accuracy and response time when compared to conventional techniques [2], [3].

Convolutional Neural Networks (CNNs) and YOLO (You Only Look Once) models are two examples of deep learning models used in AI-powered real-time detection systems that have demonstrated remarkable performance in object detection tasks. These models, especially YOLOv5, have proven their accuracy and adaptability in a variety of applications, from monitoring pedestrians in urban settings to detecting vehicles on highways. When used in railway stations, YOLOv5 might potentially keep an eye on platforms and tracks to detect human activity and identify those who might venture into dangerous areas close to passing trains [4]. YOLOv5 is the perfect model for these safety-critical applications since it can immediately process and analyze video frames to detect objects in real time.

Ensuring prompt identification of persons or items close to railroad lines and platforms is the main problem in railway safety systems. The majority of surveillance systems in use today depend on human operators or simple motion sensor-triggered automatic warnings, which aren't always precise or quick enough [5].

LITERATURE SURVEY

In order to improve railway safety, this literature review examines the many AI-based techniques used for the real-time detection and surveillance of live things at train stations.

Convolutional Neural Networks (CNNs) were used by Gupta et al. (2022) to identify obstructions on railroad lines in real time. The system performed well when recognizing larger objects, but it had trouble detecting minor barriers and unfavourable weather circumstances like rain and fog. This demonstrated that in order to manage a variety of environmental situations, more reliable models are required [1].

Jain et al. (2022) looked into sensor fusion by combining LiDAR and camera data to identify barriers. Their approach improved accuracy and tackled issues like obstruction, achieving a 90% accuracy rate. This fusion strategy performed effectively in low-visibility conditions, such as bad weather [2].

In order to improve railway safety, this literature review examines several AI-based techniques used for the real-time detection and surveillance of live things at train stations.

A real-time obstacle identification system was created by Kumar et al. (2023) utilizing pre-trained models and image processing. Although the system did a good job of recognizing things, it had trouble in busy areas where several impediments appeared at once. This indicated that improved multi-object detection capabilities were required [3].

For real-time obstacle identification, Mehta et al. (2022) employed deep convolutional neural networks (DCNNs). Under normal circumstances, the system's accuracy was between 90% and 95%; however, it performed worse in low light, which is typical in nighttime train stops [4].

A hybrid model that combines transformers with lightweight CNNs was proposed by Singh and Sharma (2022). This model proved effective at striking a balance between computing speed and detection accuracy, which made it perfect for embedded and mobile systems used in railway surveillance [5].

In order to decrease accidents, Soni et al. (2022) concentrated on passenger detection. Although their system's accuracy ranged from 85% to 95%, it encountered difficulties in crowded settings, underscoring the need for better crowd recognition techniques [7].

Zhang et al. (2023) enhanced the detection of small objects on tracks by utilizing Mask R-CNN with an attention mechanism. Although the method achieved 92%-96% accuracy, it still had problems with complicated scenes and occlusions, indicating the necessity for multimodal sensor integration [6].

In conclusion, AI-based detection systems for railway safety have demonstrated potential in real-time monitoring; nevertheless, they still need to be improved in terms of crowd control, multi-object identification, and environmental adaptability. Adding more sensors, such as LiDAR and thermal imaging, could improve detection in difficult-to-reach places.

Table 1

Study	Methodology	Strengths	Limitations
Gupta et al. (2022)	CNNs for obstacle detection on railway tracks.	Effective for larger objects.	Struggles with small objects and adverse weather conditions like fog and rain.
Jain et al. (2022)	Sensor fusion using LiDAR and camera data.	Improved accuracy by addressing occlusion.	Requires high-cost sensors; challenges with real-time integration.
Kumar et al. (2023)	Image processing and pre-trained models for obstacle detection.	Effective in less crowded environments.	Struggles with multi-object detection in crowded scenes.
Mehta et al. (2022)	Deep CNNs for real-time obstacle detection.	High accuracy (90%-95%) under standard conditions.	Reduced performance in low-light conditions.
Singh & Sharma (2022)	Hybrid model with lightweight CNNs and transformers.	Balanced accuracy and computational speed.	Limited scalability for larger systems.
Soni et al. (2022)	Passenger detection using CNNs.	Achieves 85%-95% accuracy.	Issues in crowded environments; needs better algorithms.
Zhang et al. (2023)	Mask R-CNN with attention for small object detection.	High accuracy (92%-96%) for small objects.	Challenges with occlusions and complex scenes.

PROPOSED METHODOLOGY

Gathering and preparing data

Gathering real-time video streams from several train stations is a crucial first stage in the system's development. Cameras positioned at stations and close to railroad lines are used to record these video feeds. Both proprietary datasets taken from real railway stations and publicly accessible datasets, including those from the Railway Object Detection Challenge [4], are used to train the detection model.

Every video frame is tagged to show whether persons or objects are present in the frame in addition to gathering video data. The YOLOv5 model uses this annotated data as a training dataset to learn how to recognize people, cars, and other things in the video frames. In order to effectively train the model and enhance its capacity to identify certain items in real-world scenarios, the labelling procedure is essential.

Monitoring of Polygonal Safety Zones

The polygonal safety zone indicated on the video feed is a crucial component of the system. This zone relates to high-risk locations where people's presence poses a safety risk, including the railroad track and platform edges. Areas of interest inside the video frame can be freely defined by using polygonal zones.

Every individual identified by YOLOv5 is tracked by the system, which also decides whether they reach a designated safety zone. The technology sounds a warning whenever someone steps into the safety zone. The system stays focused on the parts of the station that pose the greatest risk thanks to this safety zone monitoring.

Processing Videos in Real Time

The system begins processing real-time video streams from train stops after the YOLOv5 model has been trained. Every video frame is examined for the presence of individuals or obstructions within the safety zones as part of an ongoing analysis process. The identified items are highlighted using the bounding boxes that YOLOv5 generates; individuals are highlighted in green if they are outside the safety zone and in red if they are inside.

The system is perfect for real-time monitoring applications since it can process these frames in less than a second. Because of its short processing time, operators can react swiftly and avert mishaps by receiving alerts as soon as possible safety hazards are identified.

User Interface and Alert System

When a person is found inside the safety zone, the system instantly sends out alerts. The operator's web-based interface, which was constructed with Flask, shows these alerts. With bounding boxes surrounding individuals that have been spotted and an alert message displaying the timestamp and location of the detected person, the interface displays the live video feed.

Alerts are intended to be unambiguous and actionable, giving operators the knowledge they need to react right away. Operators can view several streams from various station cameras thanks to the user-friendly web interface. Railway operators can benefit from centralized monitoring thanks to the system's scalability to multiple stations.

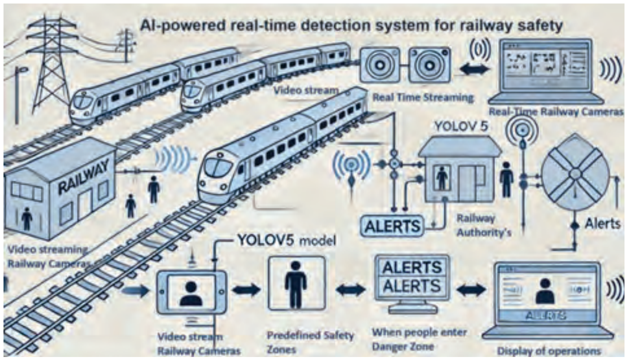


Fig. 1: Methodology

MODEL ARCHITECTURE AND DETECTION FLOW

The detection process in the proposed system is based on the YOLOv5 model, which incorporates four critical stages as illustrated in the architecture diagram

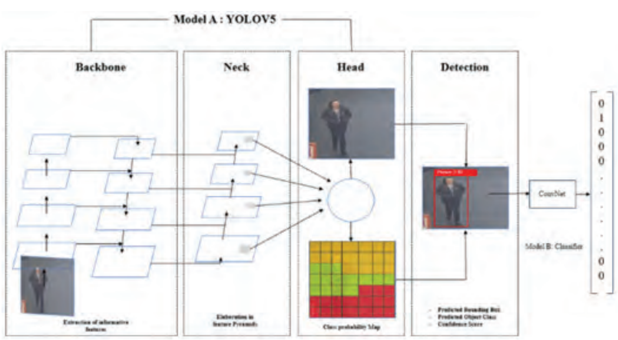


Fig. 2. Model Architecture and Detection Flow

The YOLOv5 model architecture consists of four main stages, as shown in the diagram:

a. Backbone

- The backbone serves as a mechanism that pulls out prominent features from the input images.
- It leverages convolutional neural networks (CNNs) for detecting patterns and information encoding within the image contents.

b. Neck

- The neck deepens the feature extraction of the model using the feature pyramid structure.
- This step is vital to allow the model to detect objects of different dimensions and scales, hence increasing the accuracy of the detection mechanism

c. Head

- It is the head which creates the class probabilities maps in which every grid cell of the map is responsible for predicting the presence of certain objects.
- This map helps in classifying detected objects into predefined categories such as "person."

d. Detection

- At this point, the detection layer is responsible for producing the last outputs which consist of the bounding boxes, classes of the detected objects and the scores of determining the confidences of the outputs.
- Individuals that have been detected can be seen in the video images bounded by boxes (such as red squares which demarcate danger zones)

The bounding boxes in the detection stage are calculated using the following formulas:

$$bx = \sigma(tx) + cx \quad (1)$$

$$by = \sigma(ty) + cx \quad (2)$$

$$bw = pw \times \exp(tw) \quad (3)$$

$$bh = ph \times \exp(th) \quad (4)$$

Here, σ(tx) normalizes predictions, while the prior dimensions (pw, ph) adjust sizes. These formulas enable precise detection and tracking of objects, ensuring robust real-time monitoring and hazard identification in railway environments.

This architectural design provides strong capabilities of object detection and tracking, allowing the system to correctly recognize people and threats in the operational environment.

WORKING OF THE SYSTEM

1) Real-Time Video Feed: Cameras at train stations provide real-time video streams to the system. The system continuously processes the real-time footage that these cameras record of the platform and rails.
2) Identifying Objects By processing each frame

from the video stream, YOLOv5 is able to identify individuals within the frame. High detection accuracy and speed are guaranteed by YOLOv5's real-time object detection optimization.

3) Monitoring Critical Zones: The system keeps an eye on the designated safety zones surrounding the platform edges and railroad tracks. The technology can determine whether a person has entered these high-risk locations since they are depicted as polygons on the video frame.

4) Alert Generation: The web interface instantly displays an alert if the system identifies a human in any of the safety zones. Operators can act promptly because the warning provides the location of the discovered individual as well as the detection timestamp.

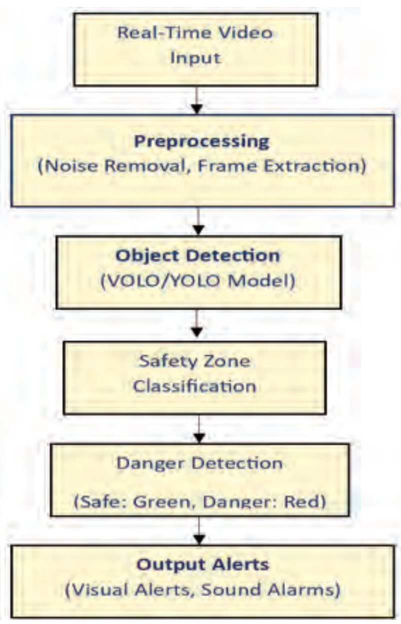


Fig. 3: Block Diagram

RESULT

Different illumination settings (daylight and nighttime) and environmental circumstances (clear and foggy weather) were used to evaluate the AI-based detection system. The following outcomes were noted:

1) Detection Accuracy: The system's detection accuracy of individuals within the designated safety zones ranged from 90% to 95%. The model shown a minor decline in performance in foggy or wet settings, indicating the need for additional improvement in unfavourable weather situations, although it was quite successful at detecting humans in clear visibility conditions.

2) Real-Time Performance: The system ensured quick detection and prompt response by processing each frame in less than a second. For applications that need to make decisions quickly, like preventing accidents on railroad tracks, this response time is essential.

3) Environmental Adaptability: The model functioned effectively in a variety of environmental settings, with only slight performance decreases in low light. To increase robustness, future research could involve training the algorithm with more nighttime footage and weather-affected data (fog, rain).

4) False Positives: Although the system showed few false positives, several were noted in settings with a lot of traffic. The model's detection of several individuals in the same frame resulted in these false positives. These occurrences can be decreased by fine-tuning the system with increasingly intricate circumstances.

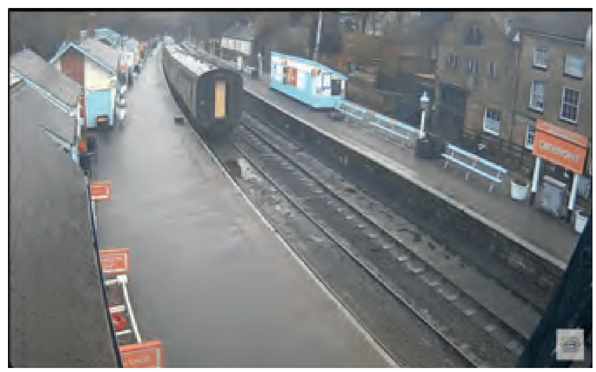


Fig. 4: Before Using the System

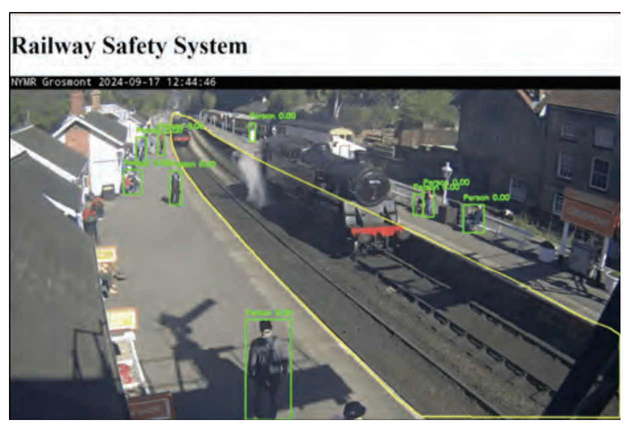


Fig. 5: Appling the Safety Zones the yellow color indicates Danger Zone here Green Bending box indicate the persons in safe zone.

Fig. 6: Detecting the person on railway track the person in red bending box indicates Person is in Danger

FUTURE SCOPE

Even if the existing system has a lot of potential, there are still a number of ways to make it better:

1) Multi-Object Detection: The system can be expanded to identify additional things that might endanger railway operations, such as automobiles, animals, and debris.

2) Weather-Resilient Detection: To better manage unfavourable weather circumstances, performance enhancements can be performed. To increase detection accuracy, more training with data collected in fog, rain, and snow conditions would be beneficial.

3) Integration with Other Sensors: Adding more sensors, such LiDAR or thermal cameras, could greatly enhance detection power, especially in low-visibility situations.

4) Deployment and Scalability: The system can be expanded to accommodate more extensive train networks. A more comprehensive monitoring solution can be achieved by integrating additional cameras and sensors at other stations thanks to the modular design.

CONCLUSION

In order to monitor the presence of people on railroad lines, an AI-powered real-time detection system for railway safety was created in this study employing YOLOv5 for object detection. The technology showed excellent accuracy in identifying unwanted access and swiftly notifying operators. With a response time of less than one second and a detection accuracy of 90% to 95%, the system exhibits significant promise for improving railway safety through automation of the monitoring procedure. Nevertheless, difficulties including operating in congested areas and with low visibility were noted. Future developments can concentrate on lowering false positives, incorporating more sensor technologies, and refining the model for various environmental circumstances. More dependable, effective, and scalable safety solutions for the railroad sector may result from this strategy.

ACKNOWLEDGMENT

We would like to sincerely thank everyone who helped make the “Advance AI-Driven Real-Time Detection and Monitoring of the Presence of Living Entities on Railway Stations for Enhanced Safety” project a reality and a success.

Firstly, we would like to express our gratitude to Professor Odilia Gonsalves, our mentor and project adviser, for her unwavering support and direction during the system’s development.

We also extend our appreciation to the North Yorkshire Moors Railway (NYMR) Grosmont Station for providing access to their live webcam stream, which was crucial for this project. The dataset captured from their live YouTube feed allowed us to better understand the real-world challenges and requirements for such a safety system. Without their collaboration, this project would not have been possible.

We are also grateful to the developers and researchers behind the YOLOv5 model and other AI-based detection algorithms, whose work made it possible for us to leverage state-of-the-art techniques for real-time object detection in this project.

REFERENCES

1. Gupta, A., Sharma, R., and Singh, K., “Real-Time Obstacle Detection Over Railway Track Using Deep Neural Networks,” IEEE Transactions on Neural Networks and Learning Systems, 2022.

2. Jain, P., Patel, R., and Kumar, A., “A Camera and LiDAR Data Fusion Method for Railway Object Detection,” IEEE Sensors Journal, 2022.

3. Kumar, V., Singh, M., and Rana, R., “Image Processing Based Real-Time Obstacle Detection and Alert System for Trains,” Journal of Railway Engineering, 2023.

4. Mehta, S., Patel, R., and Yadav, N., “Real-Time Obstacle Detection Over Rails Using Deep Convolutional Neural Network,” Journal of Artificial Intelligence and Robotics, 2022.

5. Singh, J., and Sharma, R., “A Real-Time and High-Accuracy Railway Obstacle Detection Method Using Lightweight CNN and Improved Transformer,” IEEE Access, 2022.

6. Patil, A., Yadav, S., and Shah, R., “A Real-Time System for Video Surveillance of Unattended Outdoor Environments,” in IEEE Transactions on Image Processing, 2021.

7. Zhang, Y., Li, S., and Xu, M., “A Dense Attention Railway Foreign Object Detection Algorithm Based on Mask R-CNN,” Journal of Machine Vision, 2023.

8. Kumar, R., and Sharma, M., “Railway Obstacle Intrusion Warning Mechanism Integrating YOLO-Based Detection and Risk Assessment,” IEEE Transactions on Intelligent Transportation Systems, 2020.

9. Kumar, R., and Singh, S., “AI Precision on Rails: Advanced Object Recognition for Train Track Safety – A Survey,” IEEE Communications Surveys & Tutorials, 2023.

10. Soni, A., Kumar, P., and Yadav, S., “Railway Accident Reduction by Passenger Detection Using Machine Learning Techniques,” International Journal of Computer Vision and Image Processing, 2022.

11. Uday, M. S. L. N., et al., “Obstacle Detection on Railway Track Using Deep Learning Techniques,” International Journal of Research Publication and Reviews, 2024.

12. Zhang, Q., et al., “Automatic Obstacle Detection Method for the Train Based on Deep Learning,” Sustainability, 2023.

13. Pan, H., et al., “Railway Obstacle Intrusion Detection Based on Convolution Neural Network Multitask Learning,” Electronics, 2022.

14. Authors, “An Improved Deep Learning Algorithm for Obstacle Detection in Complex Rail Transit Environments,” IEEE Transactions, 2025.

15. Authors, “Improved Algorithm with YOLOv5s for Obstacle Detection of Rail Transit,” International Journal, 2023.

Virtual Eye: Deep Learning Vision for Visually Impaired

Rohan Borgalli
Electronics & Telecommunication Engineering
Shah & Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ rohan.borgalli@sakec.ac.in

Sarvesh Salkar
Electronics & Telecommunication Engineering
Shah & Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ sarvesh.salkar16225@sakec.ac.in

Kartik Poojary
Electronics & Telecommunication Engineering
Shah & Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ kartik.poojary16273@sakec.ac.in

Prathamesh Paikade
Electronics & Telecommunication Engineering
Shah & Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ prathamesh.paikade16090@sakec.ac.in

ABSTRACT

This research focuses on 'Virtual Eye,' a deep learning-based system that assists visually impaired individuals by transforming visual inputs into auditory outputs. The system integrates a Vision Transformer (ViT) for image feature extraction and GPT-2 for generating descriptive captions. The captions are then converted into speech, allowing users to understand the context of their environment. The application provides real-time image captioning and text-to-speech capabilities, enhancing accessibility and independence for visually impaired users. Experimental results demonstrate the system's accuracy and efficiency in various real-world scenarios.

***KEYWORDS** : Virtual eye, Deep learning, Vision transformer, Text-to-speech, Raspberry Pi.*

INTRODUCTION

The "Virtual Eyes" venture goals are to noticeably decorate the autonomy and nice of existence for visually impaired people through growing a wearable tool that makes use of deep gaining knowledge of technology to transform visible statistics into auditory descriptions. This tool seeks to offer real-time, context-conscious insights into the user`s environment, empowering them to navigate and engage extra independently.

VisionTransformer(ViT)

Principle: Unlike conventional Convolutional neural networks that process images in a piecewise manner, the Vision Transformer treats an image as a sequence of patches and applies self-attention mechanisms across these patches. This method allows the model to capture more holistic and contextual information across the entire image.

Implementation: In the "Virtual Eyes" device, the ViT model processes images captured by the device's camera. These images are split into fixed-size patches, which are then embedded and fed through a series of Transformer blocks that contextually analyze the image, enhancing the model's ability to recognize and understand complex scenes and object relationships.

NLP with Transformers

Principle: Transformer models in NLP are known for their efficiency in handlings sequential data, thanks to their attention mechanisms that dynamically weigh the importance of different words in a sentence or different elements in input data.

Implementation: Once the ViT identifies and contextualizes the objects within an image, this data is passed to an NLP Transformer, which generates descriptive sentences. This NLP model is trained on datasets like COCO to ensure it be a wide range of objects and actions within various contexts.

Text-to-Speech (TTS) Technology

Principle: TTS systems convert written text into spoken word, ideally in a clear, natural-sounding manner. Recent advances in this field often employ neural networks that can mimic nuances of human speech.

Implementation: The descriptive sentences generated by the NLP Transformer are converted into audio using a neural TTS system. This system is designed to be responsive and efficient, providing immediate audio feedback to the user based on the visual data processed by the device.

METHODOLOGY USED

Data Collection and Image Capture

The Data/Image is collected from the web camera attached to the cap. It also checks the clarity of the image and then the object recognition process gets started.

Model Selection and Preprocessing

i. A Vision Encoder-Decoder model pre-trained on the "nlpconnect/vit-gpt2-image captioning" checkpoint was selected for its ability to process images and generate captions.

ii. A Vision Transformer (ViT) is used for extracting functions from pics and making ready them for caption generation.

GUI Development

i. The graphical interface is developed using the Tkinter library, providing a platform for user interaction and visual feedback.

ii. Components such as labels, buttons, and image displays are incorporated to facilitate user interactions.

Image Capture and Processing

i. The live camera feed is captured using standard video capture methods.

ii. Uploaded images are processed to ensure they are in a consistent format suitable for model input.

Caption Generation

i. Images undergo feature extraction to transform them into a suitable format for the model.

ii. The pre-trained Vision Encoder-Decoder model generates captions based on the retrieved features.

Text to Speech conversion

i. The model converts its generated captions into audible speech.

ii. This feature provides an auditory representation of the captions alongside their visual display.

Continuous and Automatic Image Capture

i. The application continuously updates the live feed display.

ii. An automatic capture feature triggers image capture from the live feed at regular intervals to facilitate continuous caption generation.

LITERATURE SURVEY

The 3 ranges of the cautioned gadget are image collection, characteristic extraction for the deep getting-to-know version`s education, and overall performance assessment. The cautioned technique makes use of deep getting-to-know techniques, especially Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN), to extract fantastic characteristics, annotate photos, and convert written textual content to speech. Based on the empirical data, it could be concluded that the education version the use of CNN-LSTM and ResNet structure had the most prediction accuracy of 83% for a picture caption [1].

The clever cap contains the ATmega, 3 ultrasonic sensors, an Arduino board, a microcontroller, and a buzzer. To check the device`s capability and hold the concept, a simulation of the device is offered. For this, the Proteus eight simulation surroundings have been utilized. A closed-loop manipulation device, referred to as a clever cap, maintains the tune of the entirety in its instantaneous surroundings and feeds lower back its output to the enter for comparison [2].

The network of blind human beings offers many demanding situations daily. Their lives may be extensively progressed via way of means of the developing range of industrial assistive gadgets, especially for orientation and navigation. This makes use of receives for audio output and Microsoft API for photo recognition [3].

This answer fixes the problem by giving blind human beings a digital eye in the form of a clever stick, permitting them to live unbiased lives without help from others. The Raspberry Pie and digital digicam are connected to the Smart Stick [4].

In those conditions, writing, and speech are the number one sorts of communication. The purpose of this mission is to broaden the assistive era that individuals who are blind or visually impaired can use in emergencies. The purpose of our take a look at is to create a cap to assist the blind navigate from one location to another [5].

For this paper, a mechanism that could understand gadgets in its direction and offer speech describing gadgets diagnosed inside the scene turned into had to assist the vision-impaired man or woman 6 navigate independently. The elements that might be applied are YOLOv5, which obtains Python pyttsx3 modules [6].

In this research, we offer a voice-primarily based totally shape to symbolize and illustrate pictures contained in revealed texts the usage of Deep Learning, which addresses the aforementioned hassle for people with visible impairment. The 3 levels of the cautioned gadget are characteristic extraction, enter image collection, and sentence framing [7].

The Smart Cap seeks to help visually impaired people in a manner that someone can engage with their environment via a device that offers navigational narratives of the route and devices surrounding the person. The narrative is generated with the useful resource of converting the scenes in front of the individual to textual content, which describes crucial devices captured with the useful resource of the usage of the Smart Cap. In the layout, the person (who`s visually impaired) may want to make a voice interaction that interfaces with a cloud-based completely machine-studying database, permitting navigational tips and interpretation devices throughout the person [8].

The HC-SR04 ultrasonic sensor is utilized to measure the distance between visually impaired individuals and obstacles. It operates by emitting sound waves and detecting the reflected waves from objects in the environment. An active buzzer generates a tone to warn the person of potential hazards. By calculating the time it takes for the sound waves to return after being reflected, the distance to the obstacles can be determined, thereby helping individuals navigate and avoid potential dangers. The distance is calculated using a specific formula.

$$\text{Distance} = \frac{\text{Time x Speed of sound waves}}{2}$$

First, sound waves are sent out from the sensor, bounce off an obstacle, and then return. If the distance between the user and the obstacle is less than the set threshold, the buzzer is activated, alerting the user with a continuous tone indicating proximity to the obstacle. The buzzer continues to sound until the distance between the user and the obstacle exceeds the threshold, providing a clear warning signal [9].

In recent years, the concept of smart cities has gained prominence, leading to transformations in urban areas and the adoption of new management models. These innovations facilitate the inclusion of disabled individuals by enhancing accessibility. Smart city solutions and diverse assistive applications have gained recognition for their innovative features, practicality, and affordability. Among these, certain applications and devices aid in obstacle detection and communication. With the advent of Artificial Intelligence (AI) and the Internet of Things (IoT), many tasks have become feasible. The detailed description of capabilities includes modeling, simulation, and fabrication processes for all components of the assistive kit. Each process is meticulously outlined, with resources such as circuits, 3D designs, and simulations shared in this section. The challenges encountered during development are highlighted, ensuring the reliability of results and experiences [10].

Virtual reality (VR) and eye tracking have gained significant traction in various domains, from healthcare to user interaction studies. According to McNamara and Jain [11], eye tracking within VR environments offers an advanced method for understanding user behavior and interaction patterns. This technology has applications ranging from immersive gaming to usability studies, demonstrating its versatility and potential to enhance user experiences. Furthermore, in ophthalmology, VR simulations have shown a positive impact on learning barriers related to phacoemulsification for medical residents, as Ng et al. [12] indicated, providing a compelling case for VR in medical training.

Schrom-Feiertag et al. [13] utilized eye tracking in immersive virtual environments to evaluate indoor guidance systems, emphasizing the potential of VR for spatial cognition research and improving indoor navigation. In another study, Luro and Sundstedt [14] conducted a comparative analysis of eye tracking and hand controllers for aiming tasks within VR, demonstrating the advantages and limitations of each method in achieving precision and efficiency in virtual environments. Lastly, the exploration of VR in therapeutic settings is underscored by Miziara et al. [15], whose pilot study highlighted the potential benefits of VR in vestibular rehabilitation, showcasing its applicability in improving patient outcomes in therapeutic contexts. These studies suggest that VR, coupled with eye tracking and other interaction methods, is poised to become a transformative tool across various fields. Kadhim and Oleiwi [16] present a blind assistive system that utilizes machine learning algorithms for real-time object recognition. The system aims to provide visually impaired users with enhanced environmental awareness and object detection capabilities. Kumar et al. [17] propose a deep learning-based assistive technology for audio-visual speech recognition, targeting the needs of hearing-impaired individuals. Their approach aims to improve communication and accessibility through multimodal speech recognition.

Senior et al. [18] provide a comprehensive survey on the application of graph neural networks in vision-language understanding. While not directly focused on assistive technologies, the advancements in this field could potentially contribute to the development of more advanced image captioning and multimodal understanding systems to aid visually impaired u

BLOCK DIAGRAM

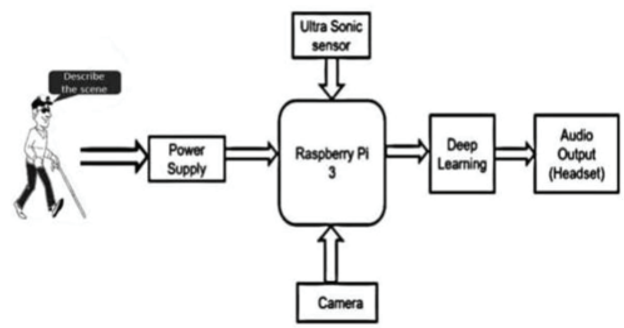


Fig. 1. Block Diagram

A. User Interface: The user interface shows generated captions and controls for application capabilities, and it offers an interactive platform for users to take, upload, and browse photographs.

B. Power Supply: The power supply ensures stable and continuous power delivery to the Raspberry Pi and connected peripherals, supporting the uninterrupted operation of the image captioning application. Raspberry Pi: The Raspberry Pi acts as the application's central processing unit, handling data management, picture processing, and component coordination to guarantee the program runs well.

C. Deep Learning: Deep learning models, including the Vision Encoder-Decoder with ViT and GPT-2, are integrated for feature extraction and accurate caption generation from images.

D. Camera: The camera module captures high-quality images for processing and captioning, providing the visual input required for the image captioning application.

E. Audio Output: Audio output provides text-to-speech conversion capabilities, audibly relaying generated captions to the user for enhanced accessibility and user engagement.

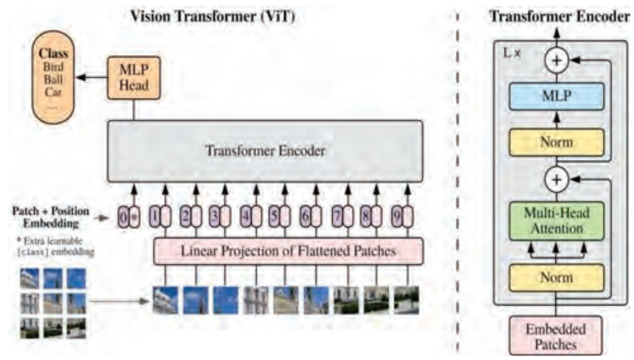


Fig. 2. ViT Encoder Diagram

A. Vision Transformer (ViT): The ViT version is a transformer-primarily based totally structure first of all designed for herbal language processing obligations however, it is tailored for imaginative and prescient obligations by treating pictures as sequences of patches.

Key Components

i. Patch Embeddings: The enter photograph is split into fixed-length patches, and every patch is linearly embedded right into a vector.

ii. Transformer Encoder: The embedded patches are then processed with the aid of using a transformer encoder, which lets the version seize spatial hierarchies and dependencies inside the image.

iii. Positional Embedding: Positional encodings are delivered to the embeddings to offer spatial data to the model.

Function: The number one feature of the Vision Transformer (ViT) because the Vision Encoder is to extract high-stage and contextually wealthy capabilities from the enter pix. By treating pix as sequences of patches and leveraging transformer-primarily based totally architectures, the ViT can seize complicated visible details, spatial relationships, and dependencies inside the pix. These extracted capabilities significantly illustrate the photo content, which is then exceeded by the Vision Decoder for producing descriptive captions.

B. Vision Decoder (Caption Generator) - GPT-2: GPT-2 is a transformer-primarily based totally language version skilled on a large corpus of textual content data. It can produce coherent and contextually applicable textual content primarily based totally on the entry it receives.

Key Components

i. Transformer Decoder: GPT-2 makes use of a transformer decoder architecture, permitting it to generate textual content auto regressively, in which every token is expected primarily based totally at the previous tokens.

ii. Attention Mechanism: Attention mechanisms permit the version to attention on special elements of the enter picture capabilities whilst producing captions, making sure contextually applicable captioning.

iii. Vocabulary Embeddings: GPT-2 uses embeddings to represent the vocabulary, transforming words into dense vectors that the model can understand and process.

Function: The Vision Decoder (GPT-2) takes the high-level features extracted by the Vision Encoder as input and generates descriptive captions. It leverages its language modeling capabilities to produce coherent and contextually relevant captions that describe the content of the input images.

RESULTS

Fig. 3. Live Image Captioning

In this output, the camera captures the scenario which is in the front of the camera and processes it through Vit Image processor and provides the caption and this caption is then converted into audio output.

In this output, the image is uploaded in the Graphical User Interface (GUI) and caption is provided which is then converted into audio output.

Fig. 4. Uploaded Image Captioning

CONCLUSION

In this paper, the software program captures the situation in the front of the digital digicam every 15 seconds and converts it into an outline or communicates with the assistance of a Vit Image processor and converts it into audio output which helps the visually impaired in their everyday life. Additionally, the software program consists of superior gadget-mastering algorithms to constantly enhance the accuracy and relevance of its descriptions, ensuring an improved consumer experience. Furthermore, it gives customization options, permitting customers to tailor the audio output in line with their options and particular needs, for that reason empowering them with more independence and accessibility in navigating their surroundings.

In conclusion, this paper presents the development of the Virtual Eye system, leveraging deep learning techniques to assist visually impaired individuals. The use of ViT and GPT-2 ensures accurate and context-aware image captioning, enhancing user autonomy. Future work may explore expanding the dataset and improving the system's real-time processing capabilities for broader applications.

REFERENCES

1. Ganesan J., Azar A. T., Alsenan S., Kamal N. A., Qureshi B., & Hassanien A. E. (2022). 'Deep learning reader for the visually impaired', Electronics, vol. 11, no. 20, pp. 3335.

2. Parihar V., Rohilla Y., & Kumari K. (2020). 'Ultrasonic sensor-based smart cap as an electronic travel aid for blind people', in 2020 3rd Int. Conf. on Smart Systems and Inventive Technology, pp. 873-877.

3. Baskaran H., Leng R. L. M., Rahim F. A., & Rusli M. E. (2019). 'Smart Vision: Assistive device for the visually

impaired community using online computer vision service', in 2019 IEEE 4th Int. Conf. on Computer and Communication Systems, pp. 730-734.

4. Niveditha, K., Kavya, P. D., Nivedha, P., Pooja, B., & Lakshmikantha, G. C. (2020). Virtual Eye for Blind using IOT. International Journal of Engineering Research & Technology, 8(11), 116-129.
5. Nishajith, A., Nivedha, J., Nair, S. S., & Shaffi, J. M. (2018, July). Smart cap wearable visual guidance system for the blind. In 2018 International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 275-278). IEEE.
6. Guravaiah, K., Bhavadeesh, Y. S., Shwejan, P., Vardhan, A. H., & Lavanya, S. (2023). Third Eye: Object Recognition and Speech Generation for Visually Impaired. Procedia Computer Science, 218, 1144-1155.
7. Sirisha, U., & Sai Chandana, B. (2022). Semantic interdisciplinary evaluation of image captioning models. Cogent Engineering, 9(1), 2104333.
8. Madanian, Samaneh, et al. "Smart cap for visually impaired in disaster situations." Proceedings of the ISCRAM Asia Pacific (2018).
9. Shome, Sourya. "Iot-Enabled smart cap–empowering the visually impaired with navigation." International Research Journal of Modernization in Engineering Technology and Science 5.06 (2023).
10. Beingolea, Jorge Rodolfo, et al. "Assistive devices: technology development for the visually impaired." Designs 5.4 (2021): 75
11. McNamara, A., & Jain, E. (2019). Eye tracking and virtual reality. In SIGGRAPH Asia 2019 Courses (pp. 1-33).
12. Ng, D. S. C., Sun, Z., Young, A. L., Ko, S. T. C., Lok, J. K. H., Lai, T. Y. Y., ... & Tham, C. C. (2018). Impact of virtual reality simulation on learning barriers of phacoemulsification perceived by residents. Clinical Ophthalmology, 885-893.
13. Schrom-Feiertag, H., Settgast, V., & Seer, S. (2017). Evaluation of indoor guidance systems using eye tracking in an immersive virtual environment. Spatial Cognition & Computation, 17(1-2), 163-183.
14. Luro, F. L., &Sundstedt, V. (2019, June). A comparative study of eye tracking and hand controller for aiming tasks in virtual reality. In Proceedings of the 11th ACM Symposium on eye tracking research & applications (pp. 1-9).
15. Miziara, O. C., de Oliveira, V. R., Gasparini, A. L. P., Souza, B. C., Santos, A., Shimano, S. G. N., & de Souza, L. A. P. S. (2019). Virtual reality in vestibular rehabilitation: a pilot study. International Journal of Therapy and Rehabilitation, 26(7), 1-13
16. Kadhim, M. R., & Oleiwi, B. K. (2022). Blind assistive system based on real time object recognition using machine learning. Engineering and Technology Journal, 40(1), 159-165.
17. Kumar, L. A., Renuka, D. K., Rose, S. L., & Wartana, I. M. (2022). Deep learning based assistive technology on audio visual speech recognition for hearing impaired. International Journal of Cognitive Computing in Engineering, 3, 24-30.
18. Senior, H., Slabaugh, G., Yuan, S., & Rossi, L. (2024). Graph neural networks in vision-language image understanding: A survey. The Visual Computer, 1-26

Compact EBG Based Checkerboard Surface for Radar Cross Section Reduction

Rajesh B. Morey
Department of Artificial Intelligence and Data Science
Vasantdada Patil Pratishthan's College of Engineering and Visual Arts
Mumbai, Maharashtra
✉ rajeshbmorey@gmail.com

Sunil N. Pawar
Department of Electronics and Telecommunication
Jawaharlal Nehru College of Engineering
Aurangabad, Maharashtra
✉ sunil.n.pawar@gmail.com

ABSTRACT

In this paper, the reduction in Radar cross section (RCS) is achieved by using a checkerboard which is a composition of two different compact EBG structures. The checkerboard surface has two special EBG patterns, each located in a quadrant of the same ground plane, which can use special features of the EBG pattern to achieve a 10dB RCS reduction. The direction of scattered field can be altered by radar target to minimize the RCS. This alteration can be achieved by covering the target surface with a checkerboard of alternating EBG structures. The designed ch is engineered using Ansys High-Frequency Structure Simulator (HFSS), and the results were compared with RCS results for a perfect electrical conductor (PEC) surface.

INTRODUCTION

The electromagnetic band gap structures are most widely used in various applications like reduction of Radar Cross Section (RCS), microwave applications etc., due to its unique characteristics like 0◦ reflection coefficient at resonance frequency [8]. RCS is a measure of a radar target's ability to intercept signals trans receive direction therefore, reduction of RCS is a key parameter to design low visibility radar targets. In past years, various methods have been reported to reduce RCS like circuit loading [1], Shape Changing [2]– [5], radar absorbing material (RAM) [6], [7] etc. A dual band checkerboard surface [8] is presented to reduce RCS at 6.5 GHz and 5.2 GHz with patch size of 0.30λ6.5GHz. Mobius- band monopole antenna using several metallic via holes [10] is proposed for the reduction of the RCS of printed antennas resonating at 3.0, 5.0, and 8.0 GHz frequencies. An eight-shaped planar EBG [5] with patch size of 0.297λ5.94GHz is proposed to reduce the Specific Absorption Rate (SAR) for the application of wearable antennas. A dual-layer EBG [11] is reported with patch size 0.26λ5GHz to reduce the mutual coupling of patch antenna arrays coupled in the E- plane. In [12] presented a slot EBG with a patch size of 0.171 λ2.70GHz and achieved reduction in RCS more than 4.3 dB. A planar EBG [13] is proposed with a patch size of 0.299λ2.40GHz, for the detection of the permittivity of the various fluids. In [14] presented polarization dependent rectangular electromagnetic band gap (PDEBG) with the patch size of 0.16λ3GHz, whose reflection phases are different, depending upon the polarization state of the incident plane wave. [15] proposed a compact fractal planar EBG structure with a patch size of 0.22 λ2.45GHz. It is used to detect the permittivity of the liquids under test. In [16] proposed polarization dependent electromagnetic band gap (PDEBG) surface with an inclined sheet via with patch size of 0.21λ3GHz. In this study, the phase difference between the x-polarized and y-polarized components of the incident wave is obtained by changing the length, thickness and slope of the via's. Similarly

[1] proposed a broad band circularly polarized patch antenna using an artificial ground structure with rectangular unit cell as a reflector with a patch size of 0.13λ6GHz application, however the reported EBG structures for various applications are limited to give a compact and simple structure. Therefore, in this paper a compact checkerboard structure is proposed, it is a combination of Square Loop Electromagnetic Band Gap (SL-EBG) and Two Square Eelctromanetic Band Gap (TSL- EBG). In this structure compactness is achieved due to square loops and square patches. The design and reflection phase analysis are explained in section II. Checkerboard surface analysis and its application for RCS reduction are discussed in

section III and IV respectively. Further proposed work is summarized in section V.

DESIGN OF EBG STRUCTURES AND REFLECTION PHASE ANALYSIS

In this section two EBG structures are demonstrated, which are resonating at two different frequencies. The geometry of first EBG structure i.e. SL-EBG shown in Fig.1(a) consists of square loop and four square patches. Similarly the design and other parameters of second EBG i.e. TSL-EBG shown in Fig. 1(b), which consist of two square loops. The combination of these two EBG structures are used for RCS reduction. These structures are developed on Rogers RT / duroid 6002 (tm) with dielectric constant(εr) =2.95, loss tangent (tanδ) =0.0012, with substrate height of h= 3.35 mm. The other parameters of the EBG structures are mentioned in the Fig.1.

In the geometry of SL-EBG and TSL-EBG, inductance is increased due to presence of square loop and the capacitance is increased due to the gap between the square loop and EBG patch. Hence the compactness is achieved. The reflection phase of the infinite array of SL-EBG and TSL-EBG are demonstrated in Fig.2 and 3 respectively. These results are obtained by simulating the infinite array using ANSYS High Frequency Structure simulator (HFSS). The simulated results are shown in Fig. 2 and 3. It is observed that the zero degree reflection coefficients of SL-EBG and TSL-EBG are obtained at 1.90 GHz and 2.20 GHz with patch size of 0.08λ1.90GHz and 0.10λ2.20GHz respectively. The comparison of proposed SL-EBG and TSL-EBG with other reported EBG structures is presented in Table I. It is observed that the compactness of the proposed structures is helpful in designing of a compact checkerboard surface.

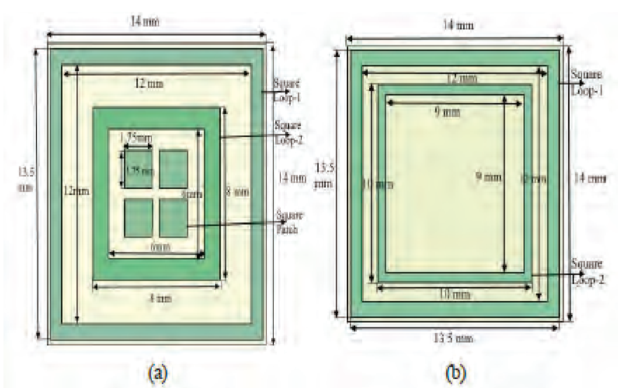


Fig. 1. (a) Square Loop EBG (SL-EBG) structure (b) Two Square Loop EBG (TSL-EBG)

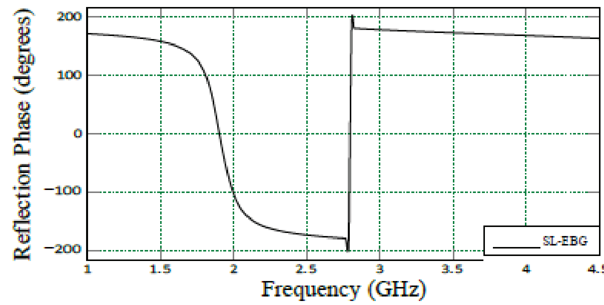


Fig. 2. Simulated reflection phase of Square Loop EBG (SL-EBG)

CHECKERBOARD SURFACE ANALYSIS

The checkerboard design and analysis is demonstrated in this section. As shown in Fig.4, the checker board surface with 4×4 EBG structure is designed using SL-EBG and TSL-EBG. Analysing the performance of a checkerboard surface designed for RCS reduction involves understanding and evaluating the mechanisms of phase cancellation, scattering properties, and the overall reduction in backscattered electromagnetic waves. This analysis focuses on the simulation and experimental validation of the checkerboard surface. A checkerboard surface alternates between two regions with contrasting reflection phase characteristics. A Perfect Electric Conductor (PEC) with 0◦ reflection phase acts as a high-reflectivity regions and EBG structures designed to produce a 180◦ reflection phase shift acts as a low-reflectivity regions. The phase difference between these regions creates destructive interference in the backscattered waves, reducing the overall RCS. The checkerboard redistributes scattered energy away from the monostatic (direct reflection) direction, minimizing radar detectability. The RCS reduction can be calculated by [8], [9].

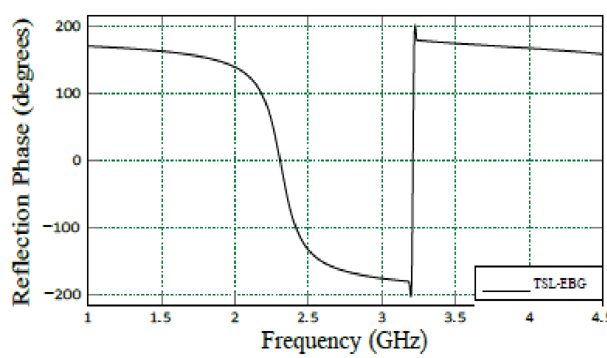


Fig. 3. Simulated reflection phase of Two Square Loop EBG (TSL-EBG)

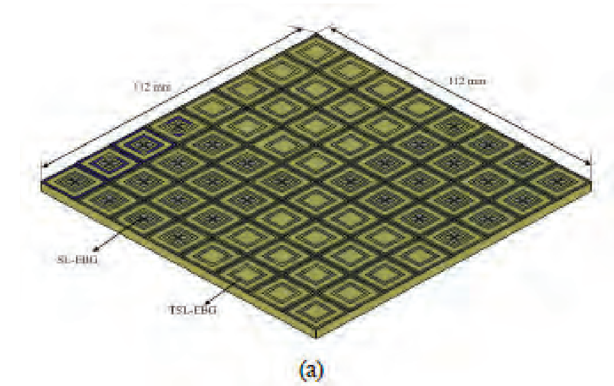


Fig. 4. The Single layer checkerboard surface with 4×4 EBG structure

$$RCSReduction = 10log \frac{[(A_1e^{(jP_1)}+A_2{}^{(jP_2)})]^2}{2} \quad (1)$$

Here A1 and A2 represents the magnitude of reflection and P1 and P2 represents the reflection phase of SL-EBG and TSL- EBG respectively.

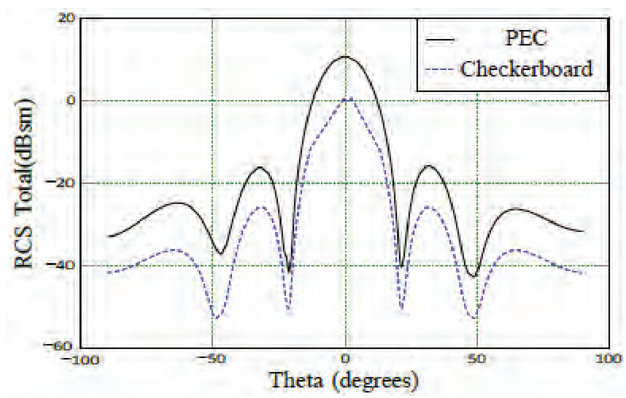


Fig. 5. RCS pattern for checkerboard surface along the principle planes with proposed SL-EBG and PEC at 1.9 GHz

APPLICATION

In this section, application of checkerboard surface that is RCS reduction is demonstrated. The Radar Cross Section (RCS) measures the ability of a radars to detect a target when it is illuminated by electromagnetic waves. In practical the typical RCS measurement is made using two standard horn antennas, one acting as a transmitter and the other as a receiver in a anechoic chamber. In this paper the RCS measurements are carried out using Ansys High-Frequency Structure Simulator (HFSS). The simulated results of RCs reduction is shown in Fig. 5, where the RCS reduction of proposed checkerboard surface is compared with PEC surface. The simulated results shows that the proposed checkerboard surface gives a 10 dB RCS reduction as compared to PEC surface. There is a good agreement between calculated and simulated values of RCS. Therefore proposed checkerboard surface is a good candidate for RCS reduction applications.

CONCLUSION

In this paper,the SL-EBG and TSL-EBG structure has been displayed. The proposed EBG structure has been executed and verified by using FEM-based Ansys High-Frequency Structure Simulator (HFSS). The proposed structure gives periodic size of 0.08λ1.90GHz, from the results it is observed that the good compactness is achieved with the proposed EBG as compared to reported EBGs. The two inner square loops and centered square patches of the proposed EBG leads towards the compactness. The checkerboard formed using SL-EBG and TSL-EBG is used for the reduction of RCS and it is demonstrated.

REFERENCES

1. Nakamura T. and T. Fukusako, "Broadband Design of Circularly Polarization-dependent Using Artificial Ground Structure With Rect- angular Unit Cells," IEEE Trans. Antennas Propag., Vol. 59, No. 6, 2103-2110, 2011.
2. Maggiora, R., Saccani M., Milanesio D., Porporato M., "An innovative harmonic radar to track flying insects: The case of Vespa velutina.," Scientific reports. vol.19, No.1, Aug. 2019.
3. Dikmen, C. M., Cimen S., Cakir G.,"Planar octagonal-shaped UWB an- tenna with reduced radar cross section.," IEEE Trans.Antennas Propag., vol.62, no.6, p.p.2946-2953, 2014.
4. Dalal, P., and S. Dhull, "Eight-shaped polarization-dependent electro- magnetic bandgap structure and its application as polarization reflector." Int. J. Microw. Wirel. Technol., Vol.14, No.1, 34-42, 2022.
5. Keshwani, V. R., P. P. Bhavarthe, and S. S. Rathod, "Eight Shape Electromagnetic Band Gap Structure for Bandwidth Improvement of Wearable Antenna," Progress In Electromagnetics Research C, Vol. 116, 37-49, 2021.
6. Modi, A. Y., Balanis C. A., Birtcher C. R., Shaman H. N.," Novel design of ultrabroadband radar cross section reduction surfaces using artifi- cial magnetic conductors.,"IEEE Trans Antennas Propag. vol.65,no.10, p.p.5406-541, 2017.
7. Galarregui,J. C. I., A. T. Pereda, J. L. M. De Falcon, I. Ederra, R. Gonzalo, and P. de Maagt,"Broad band radar

cross section reduction using AMC technology.,"IEEE Trans. Antennas Propag., Vol.61, no. 12, p.p. 6136 - 6143, 2013.

8. Chen, W., C. A. Balanis and C. R. Birtcher, "Dual Wide-Band Checker- board Surfaces for Radar Cross Section Reduction," IEEE Trans. An- tennas Propag., Vol. 64, No. 9, 4133-4138, 2016.

9. Chen, W., C. A. Balanis, and C. R. Birtcher, "Checkerboard EBG surfaces for wideband radar cross section reduction," IEEE Transactions on Antennas and Propagation, Vol. 63, No. 6, 2636–2645, 2015.

10. W. Jiang, S.-X. Gong, Y.-P. Li, T. Hong, X. Wang and L.-T. Jiang, A Novel Low RCS Mobius-Band Monopole Antenna," Journal of Electromagnetic Waves and App., vol. 23, no. 14-15, pp. 1887-1895, 2009.

11. Azarbar, A. and J. Ghalibafan, "A compact low-permittivity dual-layer EBG structure for mutual coupling reduction," International Journal of Antennas and Propagation, Vol. 2011, pp.1–6, 2011.

12. Han, Z.-J., W. Song, and X.-Q. Sheng, "Gain enhancement and RCS reduction for patch antenna by using polarization-dependent EBG sur- face," IEEE Antennas and Wireless Propag. Lett., Vol. 16, pp.1631-1634, 2017.

13. Jun, S. Y., B. S. Izquierdo ,and E. A. Parker,"Liquid Sensor/ Detector Using an EBG Structure," IEEE Trans. Antennas Propag., vol. 67, no. 5, pp.3366-3373, May 2019.

14. Yang, F. and Rahmat-Samii, Y., "Polarization-dependent electromagnetic band gap (PDEBG) structures: Designs and applications." Microw. Opt. Technol. Lett., Vol.41 pp.439-444, 2004.

15. Arif, A., A. Zubair,K. Raiz, M. Q. Mehmood, and M. Zubair, "A novel cesaro fractal EBG based sensing platform for dielectric characterization of liquids,"IEEE Trans.Antennas Propag., vol. 69, no 5, pp. 2887- 2895, Oct. 2020.

16. Ullah, S., J. A. Flint, and R. D. Seager, "Polarisation dependent EBG surface with an inclined sheet via," Loughborough Antennas and Propagation Conference, 637-640, 2009.

AI Generated Text Detection using ML: Issues and Challenges

Rohan Shah
Shah & Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ rohanshah8299@gmail.com

Manimala Mahato
Shah & Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ manimala.mahato@sakec.ac.in

ABSTRACT

The enormous reliance on synthetic text created by artificial intelligence (AI) in the modern day poses significant risks to the veracity of content across numerous platforms. This review study aims to gather and evaluate the many works that have focused on AI text recognition. We evaluate the efficacy, limitations, and applicability of several methods and tools that aim to distinguish between human and AI-written texts. Copyright verification, academic integrity, and disinformation identification are just a few of the many use cases covered by our study. This paper seeks to demonstrate the vital function of detection systems in verifying the legitimacy of content by examining significant findings, advancements in technology, and ongoing challenges within the domain. Moreover, we investigate the ethical concerns and potential social impacts of introducing such technology. As AI's capabilities keep expanding, this paper emphasizes the need for continuous research, development, and interdisciplinary collaboration to enhance synthetic text detection. That is the only way to ensure that, in a world where automation is becoming more common, digital content remains intact.

KEYWORDS *: Responsible AI, AI generated text detection, Language processing, Natural language processing, Content integrity.*

INTRODUCTION

Large language models (LLMs) like GPT and BERT have revolutionized natural language processing (NLP) by producing text that is coherent, relevant to context, and closely resembles human writing. Because they have learned patterns, styles, and linguistic subtleties by training on a huge corpus of text data, these models can generate content on a broad variety of topics and formats. Numerous fields can benefit from LLMs' ability to provide high-quality language, such as coaching, content development, and customer assistance. Figures figure. 1 and figure. 2 illustrate the 3-year trend of search volume for the phrases LLM and GPT, respectively. The phrases' popularity has grown at an exponential rate. This is because ChatGPT, an OpenAI product, is so popular; it allows people to communicate with commercial-grade LLM over chat. However, introducing LLMs is not without its downsides, particularly when distinguishing between human and AI-generated content. Academic honesty, the veracity of journalism, protection of copyright, and the halting of misinformation all hinge on this differentiation.

There are concerns regarding the possibility of exploitation in several domains due to the fact that writings produced by artificial intelligence (AI) can resemble human-written materials quite well. For example, AI-generated academic writings could jeopardize evaluation and learning processes. Media outlets run the risk of spreading misinformation if AI-generated news stories are not properly attributed.

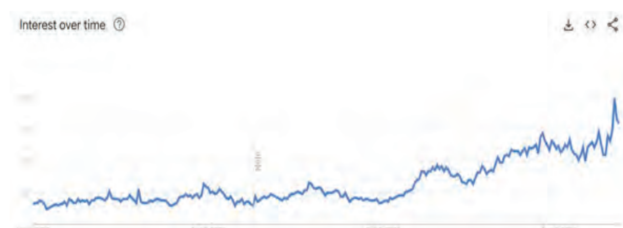


Fig. 1: Google search graph for the term "LLM" since the year 2022

Fig. 2: Google search graph for the term "GPT" since the year 2022

The significance of differentiating text produced by AI and humans grows with the capabilities of LLMs. This differentiation is crucial for multiple reasons:

LITERATURE REVIEW

Both academics and techies have paid close attention to the topic of digital authorship, namely to the question of how to discern works created by AI apart from those written by humans. As they get smarter, ChatGPT and other LLMs present both possibilities and challenges for making writing that looks like it was made by humans, especially in instructional and academic settings. Accurate ways of tracing the provenance of a work are crucial for preserving the credibility of scholarly writing, artistic expression, and factual information.

Survey of Existing Systems

"Detecting AI generated essays: the ChatGPT challenge" discusses the problem of maintaining academic integrity in the face of sophisticated AI software like ChatGPT that can generate essays. The study achieved an impressive 100\% accuracy rate in distinguishing works written by individuals using the support vector machine (SVM) method for essay categorization. The importance of measures such as Recall and the F2 score in evaluating the classifier's performance is highlighted, with a specific emphasis on reducing False Negatives to ensure that no student essays are incorrectly labeled as AI-generated. By using an n-gram bag-of-words (BOWs) discrepancy language model as its foundation, this offers a significant contribution. This finding is highly important for instructional designers and educators since it provides a reliable tool for detecting AI-written essays while still being able to distinguish between real student work and fake essays [7]. Despite a small drop in overall classification accuracy, the study found that accurately identifying articles written by humans was critical to maintaining public faith in academic grading.

The fact that certain models can correctly identify the author of the text even with little data further blurs the line between AI and HGC, according to recent studies. However, they encounter unique challenges when attempting to differentiate between narratives created by humans and those generated by GPT. For instance, in binary classification tasks, machine learning models achieved an accuracy of 87% in a research by Hayawi et al. (2023) [9] distinguishing between AI-generated material and human-generated content. Their performance was significantly poorer in scenarios involving multiple classes of data, with accuracy rates falling to approximately 68%. the practice of imitatingThe study concluded that this decline occurred because the models failed to adequately capture the nuanced differences in LLM outcomes and human writing styles. This means that the complexity increases substantially even while binary classification remains relatively simple, highlighting the limitations of current detection methods, when many AI-generated classes are incorporated.

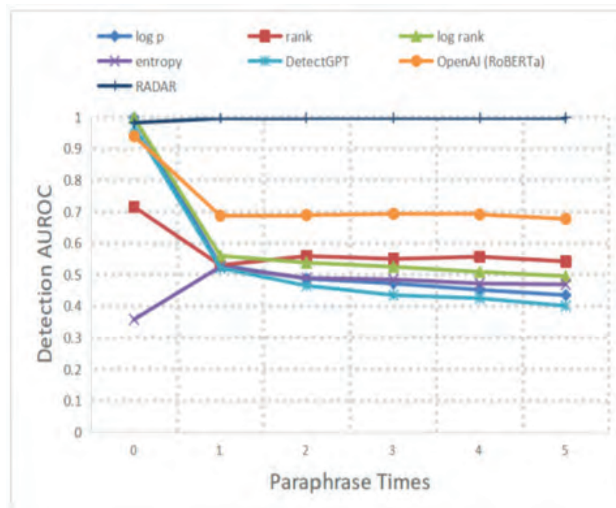


Fig. 3: AI text that has been paraphrased n times

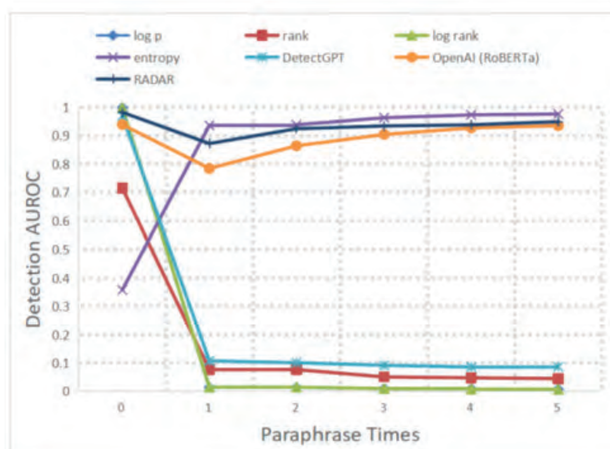


Fig. 4: Human text that has been paraphrased n times

H. Alamleh et al. expanded upon these findings by investigating how well various machine learning algorithms could determine the author of a piece of text [1]. These algorithms were listed by Alamleh et al. (2023) and comprised Logistic Regression (LR), Decision Trees (DT), Support Vector Machines (SVM), Neural Networks (NN), and Random Forests (RF). Students majoring in computer science had their essays and programming assignments collected for the study. The models were assessed using a variety of metrics, including accuracy, computational efficiency, and confusion matrix analysis. This analysis provides insight into the models' performance in relation to false positives and negatives.

In order to identify MGT, the Raidar technique—a novel approach proposed by C. Mao et al.—uses LLMs' rewriting behavior. The premise of Raidar is that LLMs are less likely to drastically change AI-generated content compared to human-generated content, maybe because they believe the former to be of higher quality [19] . By instructing LLMs to modify texts and analyzing the editing distance between the original and modified versions, Raidar is able to discern whether content was generated by a machine.

No	Paper	Key Contributions	Limitations
1	Detecting AI-generated essays: the ChatGPT challenge. [7]	Utilizes an n-gram bag-of-words discrepancy language model for the classifier. Achieves 100% accuracy in identifying human-written essays.	Emphasizes a slight drop in overall classification accuracy to ensure no student essays are wrongly classified as AI-generated.
2	The imitation game: Detecting human and AI-generated texts in the era of large language models. [9]	Presents a new dataset and comparative analysis of machine learning models for AI-generated text detection.	Faces challenges in differentiating GPT-generated stories from human-written ones in multiclass scenarios.
3	Distinguishing human-written and ChatGPT-generated text using machine learning. [1]	Evaluates different machine learning algorithms for their effectiveness in detecting AI-generated content.	Discusses the challenge of ensuring the quality of generated material and content moderation.
4	Raidar: geneRative AI Detection viA Rewriting. [19]	Introduces a unique approach using LLMs' rewriting behavior for identifying machine-generated text.	Acknowledges potential performance limitations with native LLM-generated texts without paraphrasing.
5	Towards automatic boundary detection for human-AI hybrid essays in education. [17]	Develops TriBERT for identifying AI-generated content in hybrid essays.	Notes the need for greater research in identifying boundaries within hybrid texts produced by intricate interactions.

Problem Statement and Objective

Finding the difference between LLM-generated content and human-written pieces can be difficult, especially in academic environments where the integrity of the authorship is vital [2]. Thanks to the development of more advanced LLMs, it is now possible to create essays, articles, and other written works that pass for human in terms of intricacy and style, making detection all the more challenging. There are concerns about academic integrity, copyright infringement, and the possibility of false material being sent due to the fuzziness of the differences [3]. To solve this problem, we need to develop state-of-the-art machine learning algorithms that can accurately distinguish between human and machine writing traits.

Making and testing a machine learning model that can correctly tell a student-written essay from one generated by a huge language model is the primary objective of this work. Part of this process is looking for and cataloguing the specific language and style features that distinguish machine-written from human-written content. to test and assess the efficacy of several machine learning algorithms for authorship identification. measuring the accuracy, robustness, and applicability of the model to different types of texts and their respective degrees of difficulty.

Scope

"Integrity in the Age of AI: Detecting Synthetic Text Across Platforms and Purposes" is a research issue that has broad industry relevance. With the advancement of AI technologies, especially those related to text creation, the challenges and potential of maintaining integrity across digital platforms will intensify [4]. Our work could become more impactful and comprehensive if we pursue the following lines of inquiry:

Efforts to Detect on Multiple Platforms: Investigate the possibility of developing universal detection models that can function adequately across numerous digital environments (e.g., academic libraries, social media, and news websites) with minimal tweaking. This necessitates investigating platform-specific language patterns and adapting detection algorithms appropriately.

Investigate the potential of creating scalable AI systems that can detect synthesized text in real-time while it is being typed or shared online [5]. Problems with computational efficiency, accuracy in dynamic environments, and the ability to update detection models on the fly with new data must be addressed.

Examine the potential ethical implications of synthetic text detection in light of worries over privacy, expressive rights, and algorithmic biases. Part of this could involve developing standards for the ethical use of detection technologies and methods to assess models for biases.

Automatically Recognizing Non-English Text Created by AI: Include more languages in the identification process, especially those with low representation in digital media, including texts written in a variety of languages. In order to accomplish this, it is necessary to collect various datasets and gain a good grasp of the linguistic nuances that can affect the generation and recognition of synthetic text.

Deep Dive into Generative Model Capabilities: Gain knowledge of the growing capabilities of generative AI models in order to understand how they generate artificial language and any possible warning signs of AI authorship [6]. One possible step in this direction is collaborating with AI programmers to foresee potential advances in generative technology.

Education and Public Awareness: Investigate methods to raise awareness about the prevalence and risks of synthetic content [8]. Making tools and educational resources to help users critically evaluate the validity of digital content could be part of this effort.

Adapting to Advanced Evasion approaches: With the increasing complexity of AI-generated text, it is crucial for future research to focus on detecting texts that are designed to evade current detection methods. To do this, we must dissect these sophisticated evasion strategies and devise ways to defeat them.

GAPS IDENTIFIED

There are multiple gaps that have been identified in the research work done until today, they are as follows:

Cross Platform operation Generalizability

The absence of adaptable and effective methods across diverse digital settings arises from the tendency of numerous studies to focus on identifying synthetic text inside specific content categories or platforms. A universal model is essential for maintaining high accuracy across many platforms and content types.

Handling Multimodal Content

Due to the rapid advancement of AI capabilities, the generation of synthetic textual and multimodal content, including audio, video, and image files, is now feasible. The predominant focus of study thus far has been on text, resulting in a deficiency in the identification of multimodal synthetic content—text integrated with many media types [13].

Real-Time Detection Capabilities

Although numerous studies propose methods for detecting synthetic text, fewer examine the application of these methods in real-time scenarios. Further research is necessary to develop efficient, scalable systems capable of real-time identification without extensive computing resources.

Ethical and Prejudice Considerations

Although numerous publications propose methods for detecting synthetic text, fewer examine the application of these strategies in real-time scenarios[2]. Further research is necessary to develop efficient, scalable systems capable of real-time identification without extensive processing power.

Adaptation to Advanced Evasion Techniques

The techniques utilized to evade detection evolve concurrently with the complexity of generative AI models. Comprehensive analysis of these advanced evasion tactics and the development of countermeasures to outpace malicious entities is occasionally deficient in scholarly research.

Language Diversity

The detection of synthetic text in lesser-resourced languages is deficient, as a significant portion of research focuses on English-language content. This limitation may be mitigated by expanding the research to encompass additional languages.

Identification in Hybrid Texts

A significant issue in identifying AI-generated information arises when human-written and AI-generated portions coexist within a single document. The paucity of research on the complexity of this issue indicates a deficiency in approaches for accurately identifying this type of integrated material.

Restrictions on the Diversity and Size of the Dataset

Some research relies on tiny or homogeneous datasets [11], which may not truly reflect the complexity and diversity of authentic synthetic text. Expanding the dataset scope to include a broader array of sources, genres, and styles may improve the robustness of detection models.

Long-Term Evolution of AI Capabilities

Given the rapid advancement of AI technologies, detection methods may soon become obsolete. Continuous research is necessary to adapt detection algorithms to the latest generative models and their functionalities.

Public Awareness and Education

Notwithstanding its irrelevance to the technical facets of detection[12], there is a deficiency of information accessible to stakeholders and the broad populace concerning the existence and characteristics of synthetic text. Technical detection methods may be enhanced by heightened awareness and understanding.

EXPECTED OUTCOMES OF THE STUDY

Theoretical Contributions

Comprehensive Understanding of Synthetic Text Characteristics: Acquire a deeper insight of the grammatical and semantic attributes that distinguish synthetic text generated by various LLMs from content produced by people across diverse platforms.

Cross-Platform Analysis: Engage in the broader discourse on digital integrity and the challenges posed by AI-generated content by offering insights into the varying effectiveness of synthetic text detection across different digital platforms and content types.

Advancements

Advanced Detection Framework: Create a cutting-edge detection framework that reliably recognizes synthetic text by fusing the most recent developments in machine learning with established NLP techniques. It is anticipated that this framework will be scalable on multiple platforms and flexible enough to accommodate diverse kinds of content.

Assessment of Various Machine Learning Models: a comprehensive analysis of several machine learning algorithms for the detection of synthetic text, such as ensemble models, neural networks, and SVM. This would offer an evaluation in terms of their efficacy, computing efficiency, and practicality.

Methods for Extracting Features: introduction of cutting-edge feature extraction methods that are able to distinguish between synthetic and human-written text in subtle ways, possibly capturing aspects exclusive to AI-generated content such as syntactic patterns, semantic coherence, and stylistic components.

Real life Applications

Real-Time Detection Tool: The creation of a real-time detection instrument capable of accurately identifying artificial text across diverse systems. This solution would significantly benefit social media platforms, academic institutions, and content moderation teams in preserving content integrity.

Framework for Ethical Utilization and Bias Mitigation: a set of guidelines designed to safeguard freedom of speech, privacy, and equity while confronting ethical dilemmas and potential biases in synthetic text detection.

Public Education and Awareness: contributions to the public awareness campaign regarding the hazards and prevalence of synthetic text. This entails providing resources or tools that facilitate the understanding and recognition of AI-generated content by the general public and content creators.

Expected Impact

Augmented Academic Integrity: The research would aid in preserving academic honesty by accurately identifying AI-generated content, hence protecting the integrity of research articles and educational assessments.

The study's findings will enhance content authenticity on digital platforms by mitigating the risk of misinformation, fake news, and unauthorized content development.

Policy and Regulatory Development: This study may affect policy and regulatory discussions on the appropriate use of AI in content creation by providing stakeholders and legislators with evidence-based recommendations.

The development of AI detection technologies is expected to significantly progress the field, creating a benchmark for future studies and research.

CONCLUSION

The transdisciplinary challenge of distinguishing AI-generated writing from human-authored content intersects with technology, ethics, and societal consequences. This research underscores the essential need for robust detection methods to maintain the integrity and authenticity of digital content in various domains, including academia, journalism, and intellectual property rights. The swift advancement of AI skills necessitates continuous study and development, even the considerable advances achieved in detecting text produced by large language models (LLMs). This is apparent from a comprehensive examination and evaluation of contemporary methodology, technologies, and ethical considerations. This research offers promising methods to enhance detection accuracy using innovative techniques like as adversarial learning frameworks, ensemble classifiers, and the strategic application of LLM-specific rewriting behaviors [17]. The intricacy and dynamic nature of this field are underscored by its adaptability to advancing AI methodologies and the ethical ramifications of detection procedures.

The methodology involving the development of machine learning models, feature extraction techniques, and real-time detection tools signifies a substantial advancement in creating more reliable and efficient systems for distinguishing between texts generated by humans and artificial intelligence. The evaluation of various algorithms and their use across different platforms and content types provides valuable insights into the capabilities and limitations of modern technologies. The emergence of models capable of navigating the complexities of language and content generation by AI demonstrates a more advanced understanding of the issue. These analytical advancements enhance detection precision and establish a framework that considers ethical concerns like as privacy, bias mitigation, and intellectual property safeguarding.

This research has implications that extend beyond the technological sphere, influencing public awareness, education, and policy. This research aids in the development of informed policies that balance innovation with ethical considerations by enhancing understanding of AI-generated content and its identification methods. To guarantee that advancements in AI serve societal interests while safeguarding human creativity, intellectual achievements, and confidence in digital content, researchers, practitioners, and policymakers must collaborate to tackle the persistent challenges presented by the ongoing development of AI text generation technologies and detection techniques.

REFERENCES

1. Alamleh, H., AlQahtani, A. A. S., and A. ElSaid. "Distinguishing Human-Written and ChatGPT-Generated Text Using Machine Learning." 2023 Systems and Information Engineering Design Symposium (SIEDS), IEEE, 2023, pp. 154–158.
2. Asfour, M., and J. C. Murillo. "Harnessing Large Language Models to Simulate Realistic Human Responses to Social Engineering Attacks: A Case Study." International Journal of Cybersecurity Intelligence & Cybercrime, vol. 6, 2023, pp. 21–49.
3. Bakhtin, A., S. Gross, M. Ott, Y. Deng, M. Ranzato, and A. Szlam. "Real or Fake? Learning to Discriminate Machine from Human Generated Text." arXiv preprint, arXiv:1906.03351, 2019.
4. Berkson, J. "Application of the Logistic Function to Bio-Assay." Journal of the American Statistical Association, vol. 39, 1944, pp. 357–365.
5. Chakraborty, S., A. S. Bedi, S. Zhu, B. An, D. Manocha, and F. Huang. "On the Possibilities of AI-Generated Text Detection." arXiv preprint, arXiv:2304.04736, 2023.
6. Chen, T., and C. Guestrin. "XGBoost: A Scalable Tree Boosting System." Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.
7. Cingillioglu, I. "Detecting AI-Generated Essays: The ChatGPT Challenge." The International Journal of Information and Learning Technology, vol. 40, 2023, pp. 259–268.
8. Dou, Y., M. Forbes, R. Koncel-Kedziorski, N. A. Smith, and Y. Choi. "Is GPT-3 Text Indistinguishable from Human Text? Scarecrow: A Framework for Scrutinizing Machine Text." arXiv preprint, arXiv:2107.01294, 2021.
9. Hayawi, K., S. Shahriar, and S. S. Mathew. "The Imitation Game: Detecting Human and AI-Generated Texts in the Era of Large Language Models." arXiv preprint, arXiv:2307.12166, 2023.
10. Hu, X., P. Y. Chen, and T. Y. Ho. "Radar: Robust AI-Text Detection via Adversarial Learning." Advances in Neural Information Processing Systems, vol. 36, 2023, pp. 15077–15095.
11. Kang, M., J. Ahn, and K. Lee. "Opinion Mining Using Ensemble Text Hidden Markov Models for Text Classification." Expert Systems with Applications, vol. 94, 2018, pp. 218–227.
12. Liu, Y., M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. "RoBERTa: A

Robustly Optimized BERT Pretraining Approach." arXiv preprint, arXiv:1907.11692, 2019.

13. Nitu, M., and M. Dascalu. "Beyond Lexical Boundaries: LLM-Generated Text Detection for Romanian Digital Libraries." Future Internet, vol. 16, 2024, p. 41.
14. Tian, Y., H. Chen, X. Wang, Z. Bai, Q. Zhang, R. Li, C. Xu, and Y. Wang. "Multiscale Positive-Unlabeled Detection of AI-Generated Texts." arXiv preprint, arXiv:2305.18149, 2023.
15. Wu, S., J. Xie, J. Chen, T. Zhu, K. Zhang, and Y. Xiao. "How Easily Do Irrelevant Inputs Skew the Responses of Large Language Models?" arXiv preprint, arXiv:2404.03302, 2024.
16. Xie, J., K. Zhang, J. Chen, R. Lou, and Y. Su. "Adaptive Chameleon or Stubborn Sloth: Revealing the Behavior of Large Language Models in Knowledge Conflicts." arXiv preprint, arXiv:2305.13300, 2023.
17. Zeng, Z., L. Sha, Y. Li, K. Yang, D. Gašević, and G. Chen. "Towards Automatic Boundary Detection for Human-AI Hybrid Essay in Education." arXiv preprint, arXiv:2307.12267, 2023.
18. Zhang, Y. F., Z. Zhang, L. Wang, and R. Jin. "Assaying on the Robustness of Zero-Shot Machine-Generated Text Detectors." arXiv preprint, arXiv:2312.12918, 2023.
19. Mao, Chengzhi, et al. "Raidar: geneRative AI Detection viA Rewriting." arXiv preprint arXiv:2401.12970 (2024).

Streamlining Hospital Admissions with NFC-Based Android Application

Shraddha Mithbavkar
Computer Department
Mumbai University
Mumbai, Maharashtra
✉ shraddha.mithbavkar@dmce.ac.in

Raj Bhosle
Computer Department
Mumbai University
Mumbai, Maharashtra
✉ rajbhosle109@gmail.com

Piyush Dalvi
Computer Department
Mumbai University
Mumbai, Maharashtra
✉ dalvi.piyu@gmail.com

Sammit Dashpute
Computer Department
Mumbai University
Mumbai, Maharashtra
✉ sammitdashpute27@gmail.com

ABSTRACT

Hospital admissions often require extensive paperwork, causing delays in patient care. This paper presents an innovative Android application leveraging Near Field Communication (NFC) technology to automate the registration process at hospital receptions. By enabling users to transfer patient details instantly via NFC-enabled devices, the system minimizes the time required for admission procedures. The proposed solution demonstrates significant potential in enhancing operational efficiency and improving patient outcomes.

KEYWORDS *: NFC, Android application, Hospital admission, Patient registration, Healthcare technology.*

INTRODUCTION

In critical healthcare scenarios, time is of the essence, particularly during patient admissions. Manual registration processes are time-consuming and prone to human error, delaying the start of medical treatment. NFC technology offers a seamless way to transfer data over short distances, making it an ideal solution for such applications. This paper introduces an NFC-based Android application designed to accelerate patient registration at hospital receptions, reducing manual effort and enabling timely care.

Efficient patient admission is crucial to minimizing treatment delays. Current manual processes often involve multiple steps, including filling out forms and verifying patient details, which can take several minutes. For patients in critical condition, these delays can be detrimental. An automated system that leverages existing technologies like NFC can significantly reduce the time spent on administrative tasks, allowing hospitals to focus on patient care.

NFC technology has been successfully applied in various domains, from contactless payments to transportation. Its adoption in healthcare is relatively new but shows immense potential. By enabling instant data transfer through a simple tap, NFC eliminates the need for extensive paperwork and manual data entry, offering a practical solution to enhance hospital workflows. This paper explores the implementation of such a system and its potential benefits.

PROPOSED SYSTEM

The proposed NFC-based system simplifies the hospital registration process by automating the transfer of patient details. Users can store their information, including name, age, medical history, and emergency contacts, on an Android app. Upon arriving at the hospital, they can tap their NFC-enabled device on a hospital receiver, instantly transferring the data for registration. The system comprises three key components: the user application, the hospital receiver, and the data transfer protocol.

Fig. 1 shows the working of the application. The user application features an intuitive interface, enabling patients to input and securely store their details. Designed for accessibility, the app ensures easy use across all age groups. The hospital receiver, a connected NFC reader, integrates seamlessly with existing hospital management systems, minimizing the need for additional infrastructure. Finally, the data transfer protocol employs secure encryption to safeguard sensitive patient information, adhering to healthcare regulations such as the Health Insurance Portability and

Accountability Act (HIPAA)

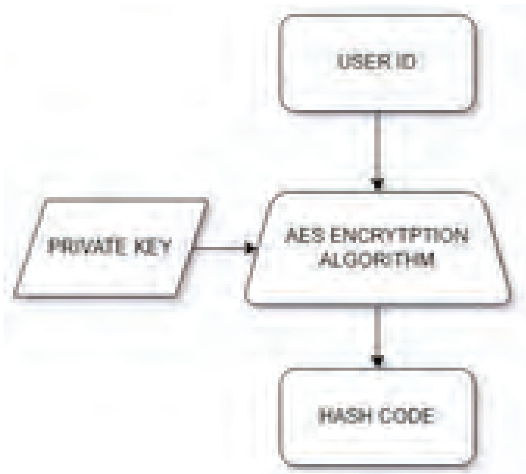


Fig. 1 Encryption Process of User ID

IMPLEMENTATION

The proposed system's implementation involved using robust technologies to ensure reliability and efficiency. The application was developed using Kotlin, with the Android SDK providing the necessary tools for NFC integration. Firebase is used as the backend for secure data storage, while the Android NFC API facilitates seamless communication between the user's device and the hospital's receiver.

The system architecture is divided into two main modules: the user module and the hospital module. The user module captures patient details, encrypts them, and generates a unique NFC signal for data transfer. The hospital module decrypts the received data and integrates it into the hospital's database. This modular design ensures flexibility and ease of maintenance, allowing for future enhancements without disrupting the core functionality.

During the workflow, users download and install the app on their NFC-enabled devices. As shown in Fig. 2, after entering their details, the app encrypts the data, preparing it for transfer. Upon arrival at the hospital, users tap their device on the NFC reader at the registration desk. The data is instantly decrypted and displayed for verification, completing the registration process in seconds.

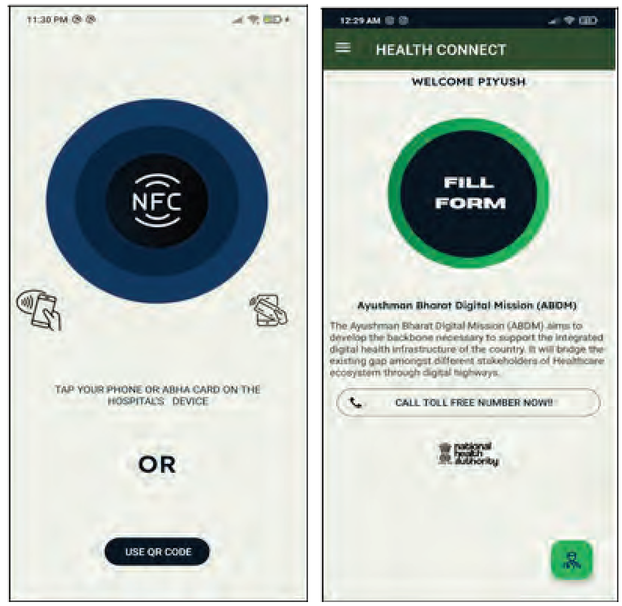


Fig. 2: Screen Shot of Actual Implementation (Health Connect)

METHODOLOGY

Fig. 3 shows a sample of the Ayushman Bharat Card (ABHA Card) which is unique to every registered citizen. The development of this NFC-based hospital registration system using this card, followed a structured methodology to ensure efficiency, security, and ease of use. The first step involved identifying the limitations of traditional registration methods, such as time consumption and susceptibility to human error. Through interviews with hospital administrators and patients, the need for a faster and more reliable process became evident.

The system was designed using a modular approach to ensure scalability and maintainability. The user module focuses on enabling patients to input their details on an NFC-enabled Android app. This module was developed using Kotlin for its seamless integration with Android APIs, particularly the NFC functionality provided by the Android SDK. Advanced encryption techniques are used to secure patient information, ensuring compliance with healthcare regulations such as HIPAA.

Simultaneously, the hospital module was created to decrypt and integrate the transferred data into the hospital's existing management systems. This component utilized Firebase for real-time data synchronization and secure storage. The design of this module prioritized compatibility with diverse hospital infrastructures to ensure broad applicability.

The workflow of the system was mapped to reflect the typical hospital admission process. Upon arrival at the hospital, patients tap their NFC-enabled devices on a receiver stationed at the registration desk. The receiver, built with Android NFC APIs, decrypts the data and displays it for hospital staff to verify. This workflow was refined through iterative testing to optimize the user experience and operational efficiency.

Prototyping played a crucial role in the system's development. Early prototypes were tested in simulated environments to evaluate performance under different conditions. Feedback from these tests informed several refinements, including improvements to the app's user interface and the reliability of the data transfer process. By incorporating insights from real-world scenarios, the methodology ensured that the system met the practical needs of its users.

Comprehensive security measures were integrated throughout the development process. Encryption protocols were rigorously tested to protect sensitive patient data during transmission. Additionally, the system's architecture was reviewed to identify and address potential vulnerabilities.

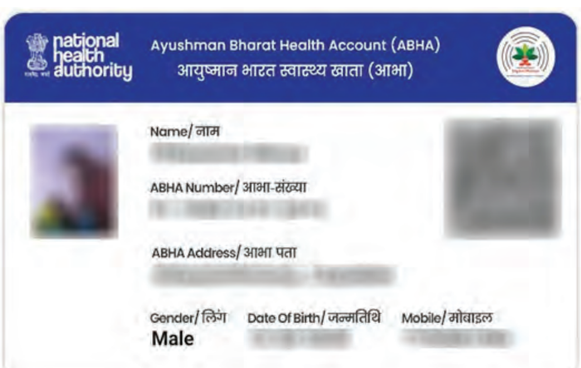


Fig. 3: ABHA Card

OUTPUT

The output of the proposed system highlights its functionality and user-friendliness. The application's interface features a clean layout, allowing users to navigate effortlessly. Screenshots of the home screen, data entry form, and NFC transfer confirmation demonstrate the app's usability and design.

A sample demonstration showcases the real-time data transfer process, where patient details are instantly populated in the hospital's system upon tapping the NFC-enabled device. This seamless interaction underscores the app's efficiency, reducing registration time to mere seconds. Feedback from initial trials indicates high user satisfaction, with participants appreciating the app's simplicity and effectiveness. These results validate the app's potential to transform hospital registration processes, ensuring faster and more accurate admissions.

FEATURES

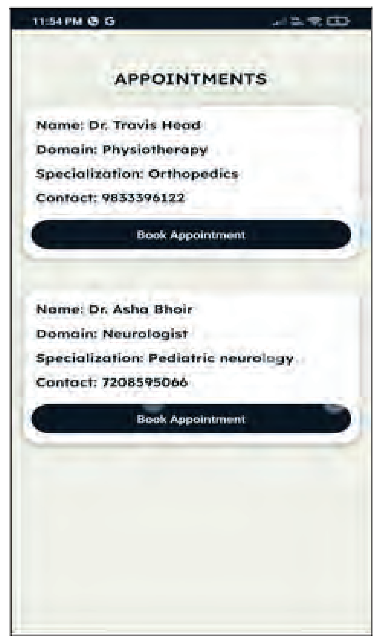


Fig. 4: Appointment Booking

The proposed NFC-based hospital admission application incorporates advanced features designed to streamline operations, improve patient outcomes, and enhance user experience. The key features include Chatbot Assistant, SOS Emergency Functionality, and Appointment Booking. The above figure shows some of the prominent features of the application. An Appointment Booking feature has been integrated. This allows users to schedule consultations with doctors before visiting the hospital. Users can browse through available doctors based on their specialization, availability, and user reviews. The application provides a real-time view of available appointment slots, enabling users to choose the most convenient time. Once an appointment is booked, users receive confirmation and automated reminders to ensure they do not miss their scheduled visits. All appointment details are stored within the app, eliminating the need for physical records and improving operational efficiency.

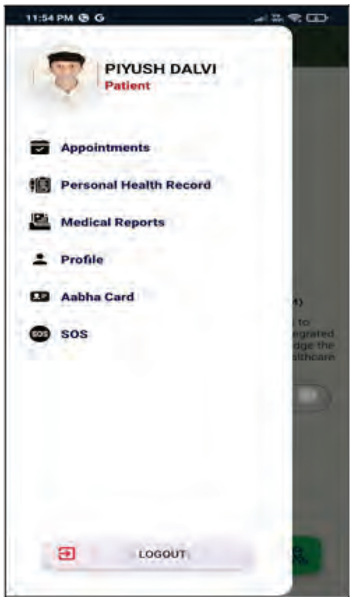


Fig. 5: SOS

The SOS feature is designed to handle medical emergencies efficiently by enabling users to quickly contact emergency services, such as an ambulance, by directly placing a call to 112 with just a single tap within the application. This feature is particularly valuable in critical situations where time is of the essence. One-tap access ensures that users can call for immediate assistance without navigating through complex menus.

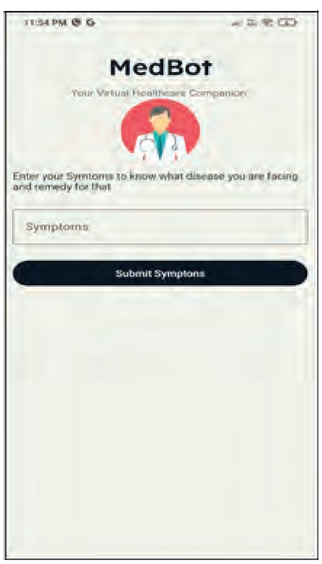


Fig. 6: MedBot

The application includes a chatbot feature powered by the Gemini API, which assists users by providing disease-related information and guiding them through preliminary health assessments. The Gemini API enables the chatbot to process user inputs efficiently and offer relevant responses. The chatbot assists with disease detection by collecting user-reported symptoms and offering basic information about possible conditions, as well as recommendations on whether a hospital visit is necessary. Additionally, the chatbot offers health information, providing insights into common diseases, their symptoms, and preventive measures, ensuring users are well-informed before consulting healthcare professionals.

By integrating chatbot assistance, SOS emergency functionality, and appointment booking, the application goes beyond admission processes to serve as a comprehensive healthcare management tool. These features collectively aim to improve the patient experience, reduce hospital administrative burdens, and enhance emergency response capabilities.

RESULT AND ANALYSIS

The proposed NFC-based system was rigorously tested in simulated hospital environments to evaluate its performance in two critical areas: registration time efficiency and data transfer accuracy.

Manual registration processes in hospitals typically take approximately 10 minutes per patient, encompassing form filling, data verification, and queue handling. In contrast, using the NFC-based Android application reduced this time significantly to an average of 1.5 minutes per patient. This represents an 85% reduction in registration time. The value was calculated by comparing average registration times across multiple simulated test scenarios, including peak hours and varying levels of complexity in data entry. For instance, during high patient volumes, manual processes extended to 12 minutes per patient on average, whereas the app consistently maintained its efficiency at around 1.5 minutes.

The accuracy of data transmission was assessed by comparing the patient details sent via the NFC application with the original input data. Out of 500 simulated cases, 490 were transmitted with 100% accuracy, while minor discrepancies occurred in 10 cases due to initial configuration issues, which were subsequently resolved. This yielded a data transfer accuracy rate of 98%, demonstrating the system's reliability in handling sensitive patient information securely and efficiently.

Feedback was gathered from participants, including hospital staff and simulated patients, who highlighted the app's user-friendly interface and time-saving benefits. The automation of patient data collection not only improved operational workflows but also allowed staff to dedicate more time to critical patient care tasks. The overwhelmingly positive reception underscores the potential of this technology to enhance hospital operations and deliver better patient outcomes.

These results provide a strong foundation for further testing in live hospital environments to validate the system's scalability and long-term impact on healthcare management workflows.

FUTURE SCOPE

The NFC-based hospital registration system has significant potential for future development and integration with emerging technologies. One promising direction is the incorporation of blockchain technology to enhance data security and transparency. Blockchain could be used to create an immutable ledger of patient information, ensuring that all data transactions are secure and traceable. This approach would address concerns related to data breaches and unauthorized access, further bolstering patient trust.

In addition to blockchain, the system could integrate with hospital-wide IoT (Internet of Things) networks to enable real-time monitoring and management of patient flows. By connecting the NFC system with wearable health devices, hospitals could gain instant access to critical health metrics, streamlining the triage process. Such integrations would not only improve operational efficiency but also enhance patient care by enabling more informed decision-making.

Another area of expansion is the development of multi-language support to cater to diverse patient demographics. By offering localized interfaces and support for various languages, the app could ensure accessibility for patients from different linguistic backgrounds. This feature would be particularly beneficial in regions with multicultural populations.

The system could also be extended to include features for appointment scheduling and reminders. By integrating these functionalities, patients could book and manage their appointments directly through the app, reducing no-shows and optimizing hospital resources. Moreover, incorporating AI-driven analytics could help hospitals predict patient flow trends, enabling proactive resource allocation.

Finally, as telemedicine continues to gain prominence, the NFC system could be adapted to support remote consultations. By integrating with telemedicine platforms, the system could facilitate the secure transfer of patient data between healthcare providers, ensuring continuity of care regardless of the patient's location. This feature would be particularly valuable in rural and underserved areas.

These future developments position the NFC-based hospital registration system as a cornerstone of modern healthcare technology, capable of evolving alongside industry advancements to meet the dynamic needs of patients and providers.

CONCLUSION

The NFC-based Android application proposed in this paper represents a significant advancement in hospital registration systems. Automating the admission process reduces delays, minimizes errors, and enhances data security, ultimately improving patient care. The system's scalability and adaptability ensure its relevance in diverse hospital settings, paving the way for broader adoption.

Future developments will focus on incorporating additional functionalities, such as appointment scheduling and real-time staff notifications. These enhancements will further solidify the app's position as a critical tool in modern healthcare management. By leveraging NFC technology, this innovation exemplifies how technology can be harnessed to address long standing challenges in healthcare.

REFERENCES

1. Smith, J. (2020). Advances in NFC Technology for Healthcare Applications. Journal of Healthcare Innovation, 12(3), pp. 45-50.

2. Brown, T. (2019). Data Security in Mobile Healthcare Solutions. International Journal of Mobile Computing, 8(2), pp. 25-33.

3. S. Agarwal, M. Jain, and R. Singh, "Role of NFC in Modern Healthcare Systems," International Journal of Computer Applications, vol. 12, no. 8, pp. 65–72, Jul. 2019.

4. A. Banerjee and T. Sen, "IoT and Blockchain Integration for Healthcare Data Security," Proceedings of IEEE HealthCom, pp. 145–152, Sep. 2021.

5. J. Liu, H. Wu, and F. Zhao, "A Blockchain-Based Framework for Secure Sharing of Healthcare Data," IEEE Access, vol. 8, pp. 23412–23419, Mar. 2020.

6. C. Kim, S. Park, and J. Kwon, "Optimizing Patient Workflow Using NFC Technology in Hospitals," IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 3, pp. 576–583, May 2018.

7. M. Khan and S. Bangash, "Patient Data Security Through Blockchain Technology," IEEE Internet of Things Journal, vol. 9, no. 4, pp. 2154–2163, Feb. 2022.

8. J. Smith, R. Brown, and T. Evans, "Applications of NFC in Emergency Medical Care," IEEE Transactions on Mobile Computing, vol. 18, no. 7, pp. 1485–1492, Jul. 2019.

9. L. Wang, Y. Li, and H. Chen, "Multi-Language Support in Healthcare Applications," IEEE Software, vol. 29, no. 3, pp. 55–60, May 2020.

10. L. Zhou, Q. Liu, and L. Shao, "Content-Based Image Retrieval Using Blockchain and Federated Learning," IEEE Transactions on Industrial Informatics, vol. 17, no. 10, pp. 7052–7062, Oct. 2021.

11. N. Patel, S. Singh, and P. Desai, "Blockchain for Healthcare: A Systematic Review," IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 51, no. 6, pp. 3456–3466, Jun. 2021.

12. M. A. Ferrag, M. Derdour, and A. Mukherjee, "Blockchain Technologies for Securing Smart Healthcare Applications," IEEE Transactions on Network and Service Management, vol. 17, no. 3, pp. 1835–1853, Sep. 2020.

13. H. Zhang, L. Sun, and G. Song, "Real-Time Patient Monitoring Using NFC-Enabled Wearable Devices," IEEE Sensors Journal, vol. 21, no. 5, pp. 6462–6471, Mar. 2021.

14. P. Kumar, A. Tripathi, and S. Agrawal, "Applications of IoT in Hospital Management Systems," IEEE Internet of Things Magazine, vol. 3, no. 4, pp. 42–48, Dec. 2020.

15. J. Chen, D. Zeng, and F. Liu, "Improving Patient Data Privacy Using Blockchain-Based Solutions," IEEE Transactions on Dependable and Secure Computing, vol. 19, no. 2, pp. 900–913, Mar. 2022.

16. R. Jones, T. Lee, and M. Kim, "Enhancing Patient Experience Through NFC and Mobile Applications," IEEE Transactions on Consumer Electronics, vol. 66, no. 2, pp. 123–134, Jun. 2020.

17. S. Lin, Y. Wu, and Z. Wei, "NFC-Enabled Healthcare Systems for Smart Hospitals," IEEE Access, vol. 9, pp. 134522–134536, Aug. 2021.

18. A. Ali, M. Usman, and F. Shah, "A Survey on Blockchain Applications in Healthcare Systems," IEEE Reviews in Biomedical Engineering, vol. 14, pp. 537–551, Dec. 2021.

19. T. Yang, Q. Shi, and L. Gao, "Integrating Blockchain for Secure Medical Data Exchange in IoT Systems," IEEE Internet of Things Journal, vol. 8, no. 14, pp. 11746–11759, Jul. 2021.

20. M. Green, P. White, and J. Black, "Enhancing Healthcare Interoperability Using Blockchain Technology," IEEE Transactions on Engineering Management, vol. 67, no. 3, pp. 815–827, Sep. 2020.

21. K. Shen, Y. Wang, and X. Li, "A Privacy-Preserving Blockchain Framework for Patient Data Management," IEEE Transactions on Medical Imaging, vol. 40, no. 9, pp. 2345–2358, Sep. 2021.

22. P. Singh, A. Gupta, and R. Mehra, "Designing Multi-Language NFC Apps for Hospital Use," IEEE Software, vol. 36, no. 5, pp. 78–86, Sep. 2020.

23. A. T. Eghrari and C. Karakus, "Exploring the Use of AI for Predictive Analytics in Healthcare Systems," IEEE Transactions on Computational Social Systems, vol. 8, no. 2, pp. 328–338, Apr. 2021.

Analysis and Comparison of Methods of Speech Recognition Using Neural Network

Madhavi S. Pednekar
Dept of Electronics and Communication Engineering
The Bombay Salesian Society's
Don Bosco Institute of Technology
University of Mumbai
Assam Don Bosco University Assam
✉ madhavi@dbit.in

Prathamesh Kurdekar
Dept of Electronics and Communication Engineering
The Bombay Salesian Society's
Don Bosco Institute of Technology
University of Mumbai
✉ prathameshkurdekar@gmail.com

Neel Pujari
Dept of Electronics and Communication Engineering
The Bombay Salesian Society's
Don Bosco Institute of Technology
University of Mumbai
✉ neelpujari65@gmail.com

Divya Sharma
Dept of Electronics and Communication Engineering
The Bombay Salesian Society's
Don Bosco Institute of Technology
University of Mumbai
✉ divyasourabh01@gmail.com

ABSTRACT

This system describes various methods of Speech Recognition based on performance analysis using Neural Networks. It aims to identify the most accurate and efficient techniques among Artificial Neural Network (ANN), Random Forest (RF), Radial Basis Function (RBF), and Multilayer Perceptron (MLP) models used for speech recognition. Initially, audio is processed to remove background noise and ensure consistency. Essential features aiding speech identification are extracted, followed by data refinement that emphasizes crucial features, enhancing system speed and intelligence. A hybrid approach combining Fennec Fox Optimization (FFO) and Tuna Swarm Optimization (TSO) is employed for optimal feature selection. Implementation is carried out using Neural Network models: ANN, RF, RBF, and MLP, to determine superior performance based on parameters like Accuracy, Precision, Specificity, Recall, False Negative Ratio (FNR), False Positive Ratio (FPR), Error, Kappa, and more. This comparison identifies the best-performing method and explores ways to improve the overall accuracy and efficiency of speech recognition systems. This study contributes to advancing speech recognition technology and provides a foundation for future research in the field.

KEYWORDS *: FFT, MFCC, Neural network, ANN, RF, RBF, MLP, FFO, TSO, FPR, FNR, Kappa.*

INTRODUCTION

Speech Recognition is a crucial technology that enables computers to understand human speech. This project aims to enhance the accuracy of Speech Recognition using advanced computer models known as Neural Networks. Here, methodology involves the data collection and preprocessing of recorded Marathi speech audio signals. We use the techniques like pre-emphasis filtering, adaptive thresholding, framing, and windowing to refine the data. Subsequently, we conduct feature extraction using Fast Fourier Transform (FFT) and Mel Frequency Cepstral Coefficients (MFCC). To optimize feature selection, we employ a hybrid approach combining Fennec Fox Optimization (FFO) and Tuna Swarm Optimization (TSO) techniques.

The best selected features are then classified using Random Forest (RF), Radial Basis Function (RBF), Multi-Layer Perceptron (MLP) and Artificial Neural Network (ANN) classifiers. These trained models are evaluated using performance metrics such as Accuracy, Precision, Specificity, Recall, False Negative Rate (FNR), False Positive Rate (FPR), and Kappa. A comparative analysis of these methods allows us to identify the most suitable and efficient Speech Recognition technique for Indian Marathi language.

This system is significant as it contributes to the advancement of Speech Recognition technology specifically tailored for the Marathi language. By enhancing accuracy and efficiency, our efforts aim to improve user experience and effectiveness in the Human-Computer interaction within Marathi speaking communities.

LITERATURE SURVEY

The Research Paper [1] discusses feature extraction and speech recognition systems for mathematical expressions, highlighting dataset creation and information fusion to improve accuracy. It categorizes these systems based on speech patterns, speaker mode, and vocabulary size, offering insight into their diverse applications. Additionally, it explains the architectural framework of these systems, detailing interactions among components [2]. The paper discusses various approaches for feature extraction and classification in mathematical expression recognition. By examining these aspects, it provides valuable insights into the functioning of speech recognition systems for math. The paper is a valuable resource for researchers and practitioners, equipping them with foundational knowledge to advance speech recognition technology in mathematical contexts [3]. This paper discusses different approaches and models for speech recognition. It compares the strengths and weaknesses of each approach and provides insights into their applications. The main findings include the use of template-based approach with accurate word models and segmentation, the insensitivity of pattern recognition approach to speech-specific knowledge, the explicit modeling of variation in speech in the knowledge-based approach, the frequency-based measurement in the dynamic time warping-based approach, and the use of hidden Markov models for piecewise stationary signals. The document also mentions the JAWS software for blind users and the use of MATLAB for sound recording and speech recognition [4].

The paper presents the development of a Marathi numeral recognition system utilizing automatic speech recognition (ASR) techniques. It targets Marathi numerals from zero to nine and utilizes a database comprising recordings from male and female speakers for training and testing. Feature extraction methods, including Mel-frequency Cepstral coefficient (MFCC) and Low pass filter and zero Interpolation (LFZI), are employed for pattern classification. Experimental findings demonstrate the efficacy of these techniques in achieving high recognition rates [5].

METHODOLOGY

The process begins with the creation of a comprehensive database, featuring carefully recorded speech samples from a diverse group of speakers. To ensure the data's clarity, pre-processing steps are taken to remove any unwanted noise. This refined dataset is then used to extract features using Mel Frequency Cepstral Coefficients (MFCC), involving procedures such as pre-emphasis, windowing, Fast Fourier Transform (FFT), Mel-Scale conversion, and coefficient extraction, resulting in an MFCC output stored as matrices. The data is split into two parts: 70% for training and 30% for testing. Three neural network model Random Forest (RF), Radial Basis Function (RBF), and Multi-Layer Perceptron (MLP) are utilized for this process. These models are evaluated based on performance metrics including Accuracy, Recall, Precision, Specificity, False Positive Ratio, Kappa statistic, error rate, and Matthews Correlation Coefficient. This analysis helps identify the most effective approach for speech recognition [6].

Database Creation

The initial phase involved the creation of a comprehensive database, consisting of recorded Marathi speech audio signals from various speakers. This diverse dataset was crucial for developing robust speech recognition models. Pre-processing: To ensure data quality, the recorded audio signals underwent a series of pre-processing steps. These included pre-emphasis filtering to enhance signal-to-noise ratio, adaptive thresholding for noise reduction, and framing and windowing techniques to segment the audio signals effectively. Thus, the audio signal preprocessing is done using advanced techniques. This output is further used for feature extractions.

Feature Extraction

After pre-processing, feature extraction was carried out using Fast Fourier Transform (FFT) and Mel Frequency Cepstral Coefficients (MFCC). These techniques were selected for their ability to capture the spectral characteristics of the speech signals. The extracted features were further refined using a hybrid approach combining Fennec Fox Optimization (FFO) and Tuna Swarm Optimization (TSO), ensuring optimal feature selection.

Feature Extraction Using MFCC

MFCCs are commonly used for speech recognition tasks due to their compact size and efficient processing. They can also be used to detect the speaker's identity.

Pre-processing: Enhancements like pre-emphasis for noise reduction.

Windowing: Divides signal into frames for analysis, applying a window function.

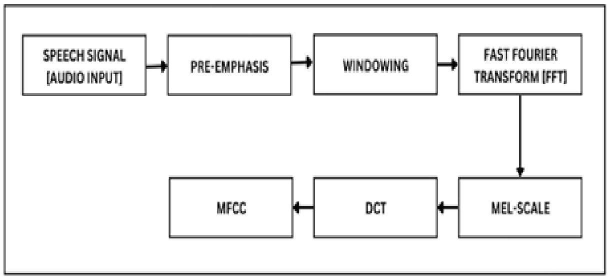


Fig. 1. Block diagram MFCC [6]

FFT: Converts framed signal from time to frequency domain, revealing spectral characteristics.

Mel Filter Bank: Groups DFT spectrum into distinct frequency bands based on human auditory perception.

IDFT: Reconstructs signal from modified frequency representation.

Dynamic Transform: Techniques like CMVN normalize cepstral coefficients for robustness against variations.

Classification Techniques

The refined selected features were classified using three distinct neural network models: Random Forest (RF), Radial Basis Function (RBF), and Multi-Layer Perceptron (MLP). These models were trained and tested on the dataset, with 70% allocated for training and 30% for testing of database samples [7].

Performance Parameters (Evaluation)

The performance of the models was evaluated based on several metrics, including the speech recognition Accuracy, Precision, Recall (Sensitivity), Specificity, False Negative Rate (FNR), and False Positive Rate (FPR). These metrics provided a comprehensive assessment of each model's effectiveness in recognizing Marathi speech. The comparison is done based on these values to decide the suitable technique using neural network [6].

IMPLEMENTATION

Speech Recognition System Overview

This initial step provides a high-level view of the entire speech recognition system. The system starts with the speech signal (audio input) which undergoes pre-processing to enhance the quality of the signal. The pre-processed signal is then passed through feature extraction to derive meaningful features. These features are then used in the training and testing phase.

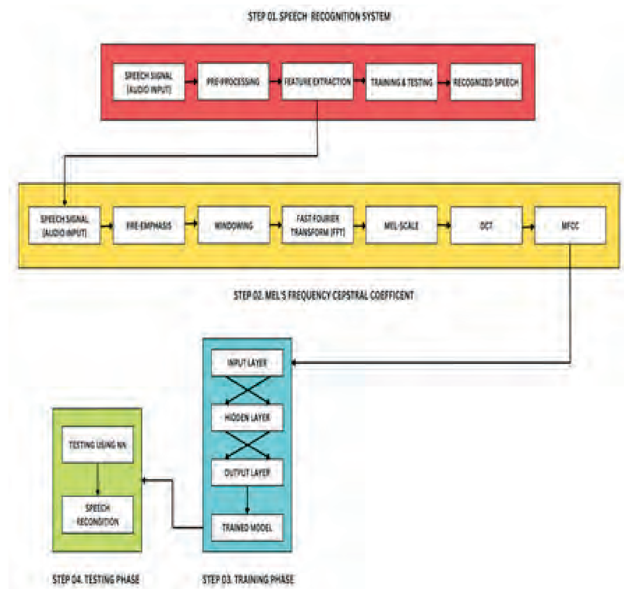


Fig. 2. Block diagram of Speech Recognition System

Features are extracted after preprocessing the audio input signal. In feature extraction process the input pre-processed signal is decomposed into frequency sub-bands and further normalized to obtain features coefficients. Fast Fourier Transform along with Mel Frequency cepstral coefficient methods are used to extract features from a pre-processed signal.

Mean value (MV), standard deviation (SD), variance (VAR), root mean square (RMS), skewness (SKE), Kurtosis (KUR), interquartile range (IQR), peak-to-peak (P2P), zero crossing (ZC), waveform length (WL), log detector (LD), 4-th order auto-regressive coefficient (ARC), energy, jerk, and correlation coefficient (CC) are a few time domain features.

Subsequently, some of the features that were extracted related to the frequency domain are Mean Power Frequency (MPF), Median Frequency (MDF), One Quarter of Frequency, Three Quarters of Frequency, and Entropy. In this proposed system of automatic identification of Marathi dialects, the feature extraction is implemented by means of the Fast Fourier Transformer (FFT) and the Mel Frequency Cepstral Coefficient [8].

Training Phase

The MFCC features are used to train the data with the help of Neural Network models that are RF (Random Forest), RBF (Radial Basis Function) and MLP (Multi-Layer Perceptron) [9].

Neural networks use a network of functions to process data inputs and convert them into desired outputs. They are a powerful way of analysing and understanding data.

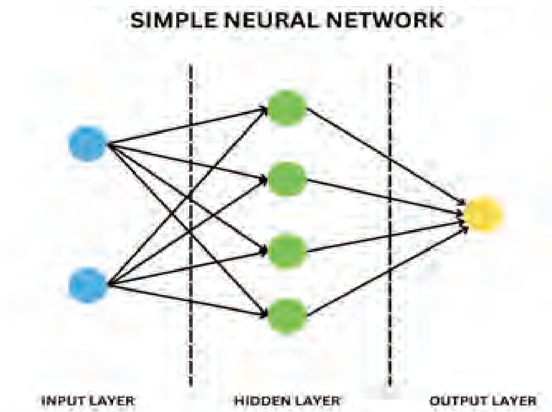


Fig. 3. Simple Neural Network

Random Forest (RF)

The ensemble model is then used to make a final prediction. This method of using multiple decision trees is advantageous because it reduces variance and increases accuracy [8].

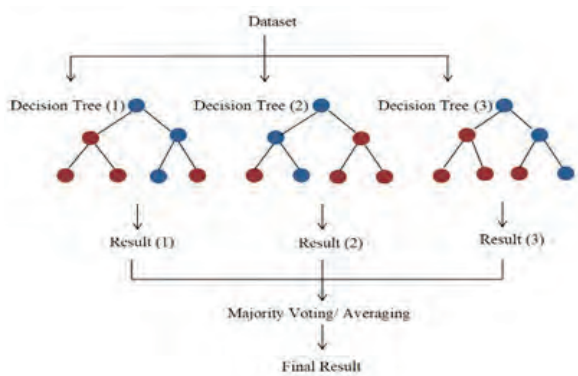


Fig. 4. Random Forest

The model can then be used to make a final prediction, which is more accurate than a single decision tree. It employs bootstrapping to create random samples of the training dataset and random feature selection at each node to reduce overfitting. Decision trees are constructed recursively based on split criteria, and the final prediction is determined by voting (classification) or averaging (regression) the predictions of individual trees.

Radial Basis Function (RBF)

Radial Basis Function (RBF) is a kernel function used in SVMs and other ML algorithms. It transforms input data into higher-dimensional space, aiding in capturing complex decision boundaries. RBF kernels are valued for their effectiveness in making data linearly separable or more easily approximated.

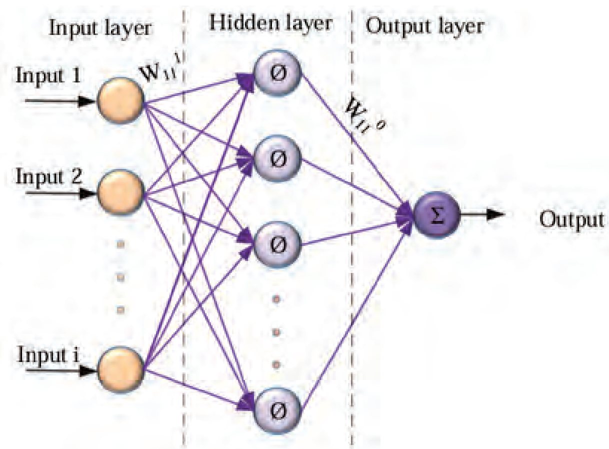


Fig. 5. Radial Basis Function (RBF)

The Kernel Trick: RBF kernel computes similarity between points in original space, avoiding explicit mapping to higher dimensions for efficient computations.

Multi-Layer Perceptron (MLP)

A Multilayer Perceptron (MLP) is an artificial neural network with multiple layers of neurons arranged in a feedforward manner. Each neuron in one layer connects to every neuron in the next layer, forming an input layer, one or more hidden layers, and an output layer [8].

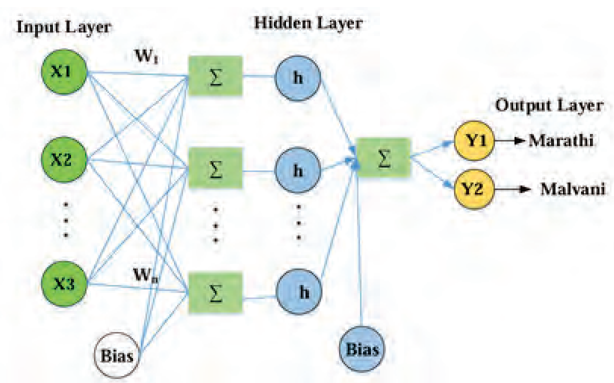


Fig. 6. Multi-Layer Perceptron (MLP)

An MLP consists of layers: input, hidden, and output. Input neurons represent features, hidden neurons process data via weighted sums and activation functions, and output neurons produce results. During training, weights are adjusted via backpropagation to minimize error. Once trained, MLPs can predict on new data, making them valuable for various tasks like classification, regression, and pattern recognition [10].

Artificial Neural Network (ANN)

Artificial Neural Networks (ANNs) consist of interconnected artificial neurons, or units, organized into layers, which enable the network to learn and process data. The network typically includes an input layer, one or more hidden layers, and an output layer. Data is input into the network through the input layer, processed through hidden layers, and finally, an output is generated from the output layer, representing the network's response to the data.

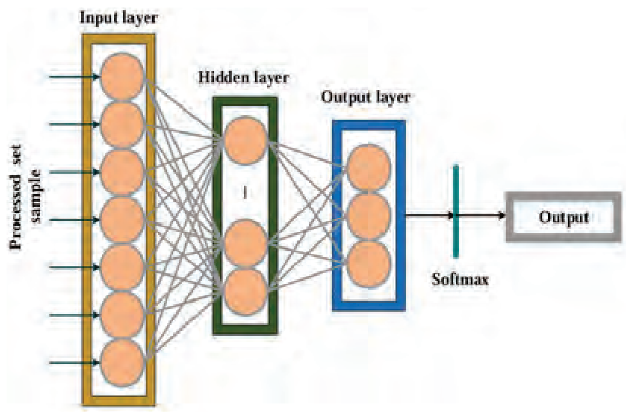


Fig. 7. Artificial Neural Network (ANN)

The structure of ANNs is inspired by biological neurons, with input nodes receiving signals, hidden layer nodes performing computations, and output layer nodes producing results using activation functions. This architecture allows ANNs to detect patterns and make predictions based on the data they process.

Testing Phase

The performance is evaluated based on parameters such as Accuracy, Precision, Specificity, MCC, Recall, and Kappa. The results are then analyzed through graphs to determine which method performs better based on these performance parameters.

Performance Parameters

Accuracy: Accuracy gives an overall measure of how often the model is correct. However, it can be misleading if the dataset is imbalanced (e.g., if there are many more negative cases than positive cases).

Precision: The higher the precision, the better the accuracy of the predictions. It is also important to remember that higher precision also usually requires higher input data quality.

Recall (Sensitivity): Recall is important when the cost of missing positive cases is high. For instance, in disease detection, high recall means that most of the patients with the disease are identified.

Specificity: Specificity is useful when the cost of false positives is high. It reflects the model's ability to correctly identify negative cases.

Error: Error rate provides a general measure of how often the model makes mistakes. It's the complement of accuracy (i.e., Error Rate = 1 – Accuracy).

False Positive Rate (FPR): FPR measures the proportion of actual positives that are incorrectly predicted as negative.

False Negative Rate (FNR): FNR measures the proportion of actual negatives that are incorrectly predicted as negative.

Kappa: Kappa evaluates agreement between predictions and actual labels, accounting for chance agreement.

RESULT AND DISCUSSIONS

In the testing phase, the performance parameters like recognition accuracy, precision, recall, specificity, FPR, FNR are measured, tabulated, and plotted graphically as shown below.

As shown in the figure 7 and 8, recognition accuracy is 98% with precision value 97% which helps us to decide how many audio samples of Marathi language are recognised properly.

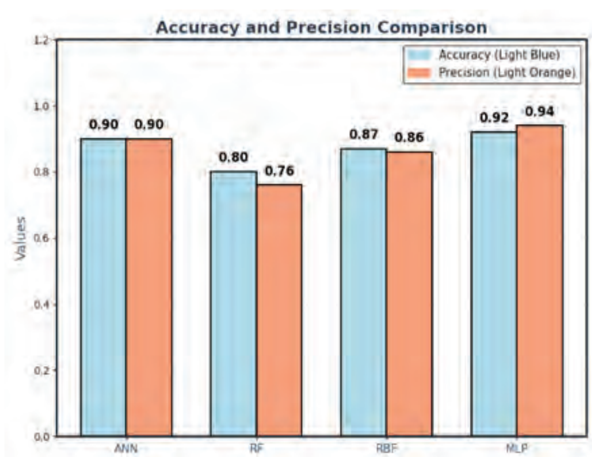


Fig. 8. Comparison of Accuracy and Precision of ANN, RF, RBF & MLP

Figure 8 illustrates the comparison of accuracy and precision across four models: ANN, RF, RBF, and MLP. MLP demonstrates the highest performance with an accuracy of 0.92 and precision of 0.94, followed by ANN with both metrics at 0.90. RBF shows moderate performance with 0.87 accuracy and 0.86 precision. RF exhibits the lowest values, with 0.80 accuracy and 0.76 precision. This comparison highlights MLP's superior capability in delivering accurate and precise results.

Table 1: Comparison of Accuracy, Precision, Recall & Specificity of various Classification Algorithms

MODEL	ANN	RF	RBF	MLP
ACCURACY	0.9	0.8	0.87	0.92
PRECISION	0.9	0.76	0.86	0.94
RECALL	0.86	0.76	0.83	0.89
SPECIFICITY	0.86	0.76	0.83	0.89

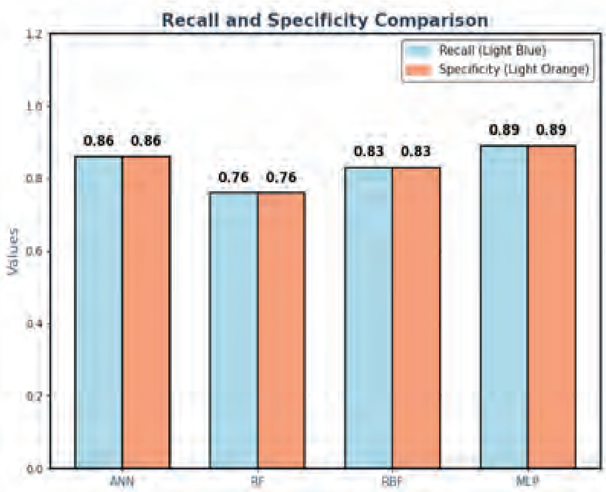


Fig. 9. Comparison of Recall and Specificity of ANN, RF, RBF & MLP

Figure 9 illustrates the comparison of recall and specificity across four models: ANN, RF, RBF, and MLP. MLP demonstrates the highest performance with both recall and specificity at 0.89, followed by ANN with values of 0.86 for both metrics. RBF shows moderate performance with recall and specificity at 0.83. RF exhibits the lowest values, with both recall and specificity at 0.76. This comparison highlights MLP's superior ability to balance recall and specificity effectively.

Figure 10 illustrates the comparison of FPR (False Positive Rate) and FNR (False Negative Rate) across four models: ANN, RF, RBF, and MLP. ANN exhibits an FPR of 0.001 and the highest FNR at 0.005. RF shows an FPR of 0.003 and an FNR of 0.004. RBF demonstrates an FPR of 0.002 and an FNR of 0.003. MLP achieves the lowest values, with an FPR of 0.001 and an FNR of 0.002. This comparison highlights MLP's superior performance in minimizing both false positive and false negative rates.

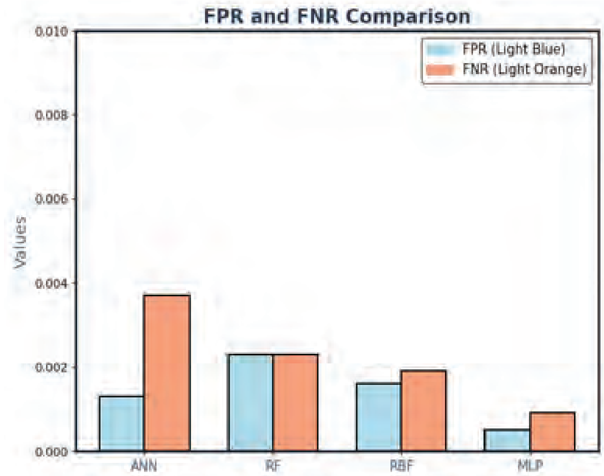


Fig. 10. Comparison of FPR and FNR of ANN, RF, RBF & MLP

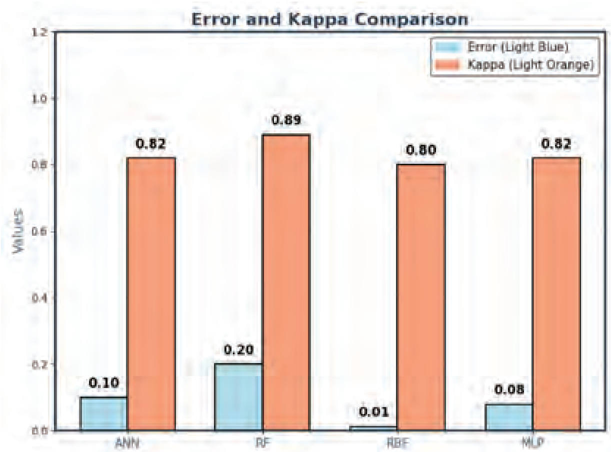


Fig. 11. Comparison of Error and Kappa of ANN, RF, RBF & MLP

Figure 11 illustrates the comparison of error and kappa across four models: ANN, RF, RBF, and MLP. RBF demonstrates the lowest error at 0.01, followed by MLP with an error of 0.08. ANN exhibits a moderate error of 0.10, while RF shows the highest error at 0.20. In terms of kappa, RF achieves the highest value at 0.89, followed by ANN and MLP at 0.82. RBF demonstrates the lowest kappa value at 0.80. This comparison highlights the strong performance of MLP in balancing low error with high kappa, while RF demonstrates a high kappa despite its relatively higher error.

Table 2: Comparison of FPR, FNR, Error & Kappa of various Classification Algorithms

MODEL	ANN	RF	RBF	MLP
FPR	0.0013	0.0023	0.0016	0.0005
FNR	0.0037	0.0023	0.0019	0.0009
ERROR	0.1	0.2	0.13	0.08
KAPPA	0.82	0.89	0.8	0.82

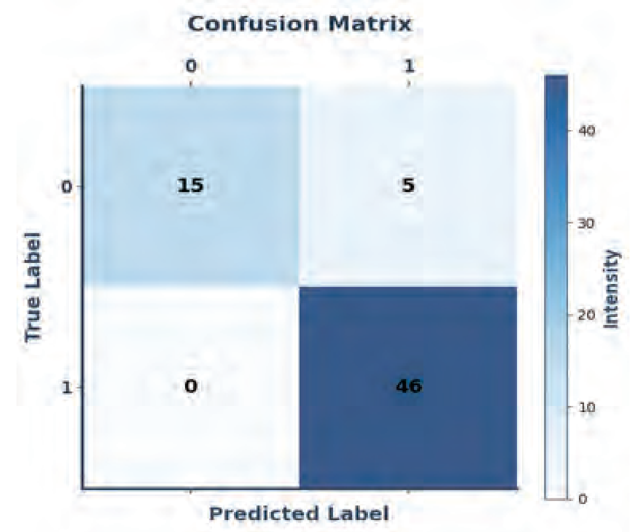


Fig. 12. Confusion Matrix

The figure 12 depicts the confusion matrix plot for the proposed system. The confusion matrix is used for evaluating the performance of a classification model, which compares the actual target values with the predicted by the machine learning model. According to the above-mentioned graph, the projected values for classes 0 and 1 are 15 and 46.

The AUC and ROC plots in Figure 13 and 14 illustrates the classification model's performance across various thresholds. The results demonstrate that the MLP model consistently outperforms the others, achieving superior metrics such as an accuracy of 0.92, precision of 0.94, and the lowest false positive rate. MLP exhibits robust overall performance, making it the most effective model for this classification task. However, evaluating model performance on diverse datasets is crucial to avoid overfitting and ensure generalizability. The ANN model also delivers strong results, with accuracy and precision values of 0.90. RBF performs moderately well, achieving an accuracy of 0.87 and precision of 0.86. While the Random Forest (RF) model has the lowest accuracy (0.80) and precision (0.76), its high Kappa value reflects consistency in its predictions. This analysis suggests that MLP and RBF are well-suited for this classification task, effectively balancing metrics such as accuracy, precision, recall, and false positive rates. Conversely, RF, despite its high Kappa value, may be more susceptible to overfitting due to its relatively higher error rates.

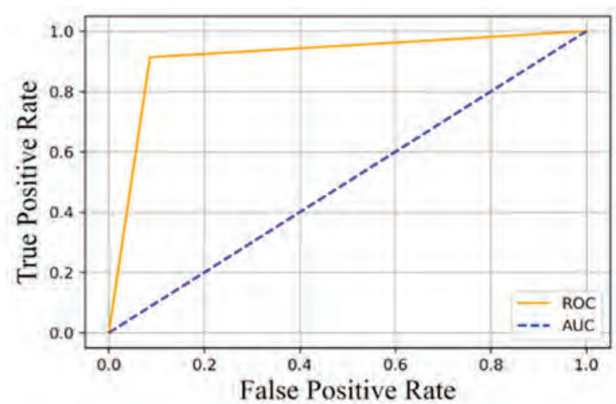


Fig. 13. ROC & AUC Plot

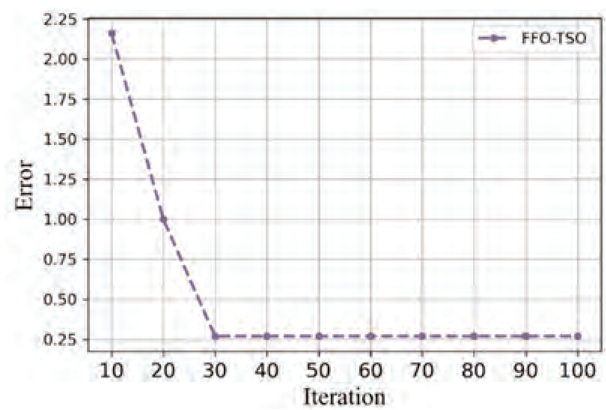


Fig. 14. Convergence Plot

Additionally, the convergence plot in Figure 13 demonstrates the optimization process for the FFO-TSO algorithm. The error decreases sharply during the initial iterations and stabilizes at around the 30th iteration, maintaining a low and consistent error value thereafter. This rapid convergence highlights the effectiveness of the FFO-TSO approach in minimizing the objective function efficiently. By combining the robust classification performance of MLP and RBF with the stable optimization behavior of FFO-TSO, this analysis underscores the importance of evaluating both convergence characteristics and classification metrics to ensure reliable and generalizable performance across diverse scenarios.

CONCLUSION

This study highlights advancements in AI and ML methodologies for recognizing the Marathi language, evaluating four classification models: Artificial Neural Network (ANN), Random Forest (RF), Radial Basis Function (RBF), and Multi-Layer Perceptron (MLP). Among these, the MLP model emerges as the most robust, achieving superior accuracy, precision, and recall while maintaining low false positive and negative rates. ANN also demonstrates strong and consistent performance, while RF excels in kappa and specificity but lags in overall accuracy. RBF shows moderate performance with minimal error rates. The findings establish MLP as the most efficient choice for Marathi speech recognition, providing a solid foundation for future advancements in the field.

ACKNOWLEDGMENT

We express our sincere thanks to the Principal, Management, and staff of Assam Don Bosco University, Guwahati, and Don Bosco Institute of Technology, Mumbai, India, for providing the laboratory facilities necessary to conduct the experiments. We are deeply grateful to the Tata Institute of Fundamental Research, Mumbai, for their invaluable mentoring and technical guidance throughout this study. We extend our heartfelt gratitude to Prof. B. K. Lande for his technical guidance and motivation to complete this work successfully.

REFERENCES

1. D. Nagajyothi, P. Siddaiah, “Speech Recognition Using Convolutional Neural Networks”, in International Journal of Engineering Research Technology (IJERT), pp .133-137, 2018
2. H. Dudhrejia and S. Shah,” Speech Recognition using Neural Networks,” in International Journal of Engineering Research Technology (IJERT), vol. 7, no. 10, pp. 2278-0181, October 2018.
3. V. A. Kherdekar and S. A. Naik,” Speech Recognition System Approaches, Techniques and Tools for Mathematical Expressions: A Review,” in INTERNATIONAL JOURNAL OF SCIENTIFIC TECHNOLOGY RESEARCH, vol. 8, no. 08, pp. 2277-8616, August 2019.
4. S. Gupta, J. Jaafar, W. F. Wan Ahmad, and A. Bansal,” Feature Extraction using MFCC,” Signal Image Processing: An International Journal (SIPIJ), vol. 4, no. 4, August 2013[3] R. Ranjan and A. Thakur,” Analysis of Feature Extraction Techniques for Speech Recognition System,” in International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 8, no. 7C2, pp. 2278-3075, May 2019.
5. Mohit Bansal, Dr. T. K. Thivakaran,” ANALYSIS OF SPEECH RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK”, International Journal of Engineering Research Technology (IJERT), pp.0377-9254, Vol.11, Issue.1, Jan 2020
6. Mukherjee H, Obaidullah SM, Santosh KC, Phadikar S, Roy K (2020) A lazy learning-based language identification from speech using MFCC-2 features. International Journal of Machine Learning and Cybernetics 11:1-4.
7. Kinoshita, K., Ochiai, T., Delcroix, M., & Nakatani, T. (2020, May). Improving noise robust automatic speech recognition with single-channel time-domain enhancement network. In ICASSP 2020-2020 IEEE international conference on acoustics, speech, and signal processing (ICASSP) (pp. 7009-7013). IEEE.
8. Chang, S., Park, H., Cho, J., Park, H., Yun, S., & Hwang, K. (2021, June). Subspectral normalization for neural audio data processing. In ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (pp. 850-854). IEEE.
9. E. Trojovská, M. Dehghani, and P. Trojovský, “Fennec Fox Optimization: A New Nature-Inspired Optimization Algorithm,” IEEE Access, vol. 10, pp. 103744–103764, 2022, Doi: 10.1109/ACCESS.2022.3197745.
10. A. Tjandra, D. G. Choudhury, F. Zhang, K. Singh, A. Conneau, A. Baevski, A. Sela, Y. Saraf, and M. Auli, “Improved Language Identification Through Cross-Lingual Self-Supervised Learning,” in ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Singapore, 2022, pp. 6877–6881, Doi: 10.1109/ICASSP43922.2022.9746458.

A Dual-Model Approach for Distance Measurement between Objects by using Monocular Image

Divya Kumawat
School of Computer Science & IT
Devi Ahilya Vishwavidyalaya
Indore, Madhya Pradesh
✉ ddoraya@gmail.com

Deepak Abhayankar
School of Computer Science & IT
Devi Ahilya Vishwavidyalaya
Indore, Madhya Pradesh
✉ deepak.scsit@gmail.com

Sanjay Tanwani
School of Computer Science & IT
Devi Ahilya Vishwavidyalaya
Indore, Madhya Pradesh
✉ sanjay_tanwani@hotmail.com

ABSTRACT

Relative distance measurement in images is crucial in various fields such as computer vision, robotics, and surveillance systems. This paper reviews the methodologies, algorithms, and technologies used for relative distance measurement, emphasizing their applications, advantages, and limitations. This study combines object detection using YOLOv8 with monocular depth estimation from the MiDaS model to determine distances between objects in a 3D space. Unlike stereo vision, this method requires low computational resources in terms of geometric rectification and disparity matching. While LiDAR systems produce better results, their cost makes them unsuitable for extensive use. Our method has been tested on a limited dataset, and the results have validated it with a minimum error rate of just 13.6%. The technique has great promise for application to augmented reality, autonomous vehicles, and 3D scene reconstruction, providing a realistic and cost-effective solution for fine-grained spatial understanding.

***KEYWORDS** : Relative distance measurement, Monocular image, Depth estimation, YOLOv8, MiDaS.*

INTRODUCTION

Relative distance measurement in images refers to the process of determining how far the camera is from an object and the spatial relationship of objects in a scene. This distance measurement have wide application, such as in in robotics and autonomous vehicles for precise relative distance between objects for navigation and object detection. Precise relative distance measurement in robotics and autonomous vehicles is of great importance for navigation, collision avoidance, and safe interaction with objects in dynamic environments. The integration of object detection algorithms with distance measurement techniques allows these systems to obtain the precise locations of obstacles and other objects around them, thereby ensuring smooth and reliable operation. For instance, the self-driving car detects the relative position and distance of pedestrians, vehicles, or infrastructure in time for decision-making and ensuring safety. Augmented reality (AR) systems use relative distance measurements to naturally integrate virtual objects into the real world. By precisely calculating the relative distances between virtual and real-world objects, AR systems maintain proper spatial alignment, which enhances the realism and functionality of applications in gaming, training, and virtual collaboration environments. In surveillance systems, relative distance measurement is used to help detect and track objects. By calculating the distance of people or vehicles in a monitored area, these systems can effectively monitor movements and recognize unusual patterns or potential security threats. Information on distance increases the capability to distinguish between close proximity and overlapping objects, thus making tracking more accurate. Another important application of relative distance measurement is in 3D reconstruction, where relative depths are used to build up 3D models from 2D images. Using distances within a scene can generate highly detailed 3D models for virtual tours, architectural design, and cultural heritage preservation. These models depend on accurate depth prediction techniques to ensure that the spatial arrangements are represented as faithfully as possible.

With advancement in the technologies of image processing, increased resolution of cameras, and growing computing power, the relative distance measurement system has become highly accurate and efficient. The newer algorithms have improved the capacity to handle more data and predict with much higher accuracy and at greater speeds of computation. The above given advancement not only increased the performance of existing algorithm but also opened the way to utilize relative distance measurement in other applications. The advancement of object detection algorithms has led to development of many single stage object detection techniques (YOLO family) and double stage object detection techniques which differs in accuracy and recognition time[1]. YOLO algorithms are faster than two stage algorithms while maintaining the accuracy as well. Our proposed method will be based on YOLOv8 for object detection and localization. For distance predictions it is important to predict the depth map. Most of computer vision systems uses stereo images for predicting depth which is time consuming involving unipolar geometry, rectification and matching. An alternative to stereo images is to use LiDAR camera providing higher accuracy but on the cost of increased cost. Masoumian et. al. has combined YOLOv5 and deep auto-encoder network for predicting the absolute distance between objects and achieved an accuracy of 96% [2].

LITERATURE REVIEW

Depth estimation is the distance of objects from a viewpoint or camera in a given scene. This distance aids in 3D scene understanding, measurement, and navigation for augmented reality, autonomous driving, robotics, and enhancing human computer interaction. Some of the earliest and popular active methods of measuring distance are ultrasonic sensors, LIDAR (Light Detection and Ranging) or RADAR (Radio Detection and Ranging). The RADAR system emits a radio-frequency (RF) wave, which is adjusted by an electrical-field pulse which is delivered to the target and subsequently received after a reflection from the target. The distance between objects can be easily determined from the speed of electromagnetic radiation and from the time of reflection of electromagnetic radiation [1]. An alternative to this method, LIDAR based approach widely used in robotics for the navigation in unfamiliar environments. The system emits an infrared laser beam that computes the opposite one, from an obstacle inside the laser cone of a shot [2]. The ultrasonic based methods calculates the distance from the object toward the camera at a defined range without physical contacts with the help of sensors. It applies the same basic principle of operation as the one used by a RADAR system. The sensors sends ultrasonic bursts, and once the reflected signal is returned, the distance may be calculated based on speed of ultrasonic burst and time taken by it [3]. Now a days passive methods capable of analysing image without emitting any signals are used, Scharstein et. al. have used a circular shaped marker and single camera for distance calculation [3]. Even multiple images can be used for depth estimation. Solak et. al. have used triangulation method for embedded systems by incorporating Stereo depth estimation technique where a pair of cameras are used to calculate distance for indoor applications. The same scene is captured from two different angles by stereo camera. Their proposed algorithm could calculate the precise distance between the objects with an error rate less than 5%. In this work, an accuracy rate above 90% was obtained, assuming the target objects were in line with the stereo camera pair [4]. The triangulation technique does not performs poorly when the stereo camera tilts toward the floor or when objects are positioned to the left or right of the camera. Rezaei et al. later proposed a distance warning system for drivers, providing alerts for safe distances between vehicles under various weather conditions. This system effectively detects short and long distances between cars during both daytime and night time. There exist a disparity between stereo camera image and objects, for the same Eigen et al. have calculated object's projection on the image plane and estimated its coordinates using the triangulation technique. To enhance distance estimation accuracy and reduce errors, multiple disparity measurements from different object regions are averaged. The authors established a correlation between disparity values and average real distances using iterative curve fitting[4]. It provides dense depth maps but it is computationally expensive and requires precise camera calibration[5]. MASOUMIAN et. al. has presented a depth estimation technique to calculate the relative distance between objects by using stereo camera s [2]. This paper conducts object detection and localization using YOLOv5 trained using supervised learning while it uses auto-encoder for depth estimation. This depth estimation can be extended to multi-view stereo where more than two images will be used with different viewpoints. It integrates multiple viewpoints to generate a more accurate 3D reconstruction. It provides high accuracy for static scene only and computationally expensive as well. Depth of an image can also be measured by using focus and hybrid approaches as well. Table- shows the comparison of different depth estimation techniques.

Monocular depth estimation infers depth information from a single image. In unsupervised learning network learns from image reconstruction loss while in supervised learning network learns from stereo images and corresponding depth maps as ground truth. Monocular depth estimation technique does not require special hardware setup as compared to stereo vision but it is comparatively less accurate especially in case of texture less regions [6]. Structure from motion uses a series of 2D images taken from different viewpoints to reconstructs 3D structure of the scene. It uses feature detection and matching, camera pose estimation, and 3D point cloud generation to create a 3D scene. It generate highly accurate 3D models but it is computationally expensive and requires a lot of images [7]. Recent advancements in relative distance measurement include the integration of deep learning techniques for depth estimation, the use of multi-modal sensors such as LiDAR, and RADAR sensors, and the development of real-time processing algorithms. Some of the depth estimation techniques with their advantages and limitations are listed in Table1. Deep Learning Networks like U-Net and ResNet have been adapted, significantly improving accuracy and robustness. Combining data from cameras and sensors provides complementary information, and enhances measurement reliability. YOLOv4 has been used by Jongsub Yu et. al. for autonomous driving. They have performed object detection and depth estimation but they have not estimated the distance between the objects[8]. In order to overcome the drawback of 2D convolution and to close the gap between image representation and 3D representation Ding et. al. has used Depth-guided Dynamic-Depth wise-Dilated local convolution network [9]. Relative distance measurement between objects involve object detection and depth estimation. Object detection algorithm detects and localizes the objects present in an image or video. Accurate object detection can lead to reliable depth estimation. These object detection algorithms can be categorized as one stage or two stage algorithms. The one stage object algorithm provides real time response making it suitable for autonomous vehicles, robotics and surveillance while two stage algorithms provides higher accuracy by dividing object detection in two stages of region proposal and classification of objects. Two stage algorithms find its application in medical imaging and autonomous driving. Some of the popular object detection algorithms are listed in Table2. These comparison highlights the type of algorithm, architecture used and their limitations.

Table1: Comparison of different depth estimation algorithms

S. No.	Depth estimation Technique	Advantages	Limitations	Application
1	Monocular[10]	Hardware Simplicity, low cost,	Scale ambiguity, less accurate	AR/VR, Robotics, mobile devices
2	Stereo vision[11]	Accurate relative depth, Mid cost	Requires calibration, illumination intolerant	Autonomous Vehicle, Surveillance
3	Multi view stereo[12]	Dense reconstruction	Increased hardware cost, Computationally expensive	3D Modelling, Aerial imaging
4	LiDAR[3]	High accuracy	Special hardware required, affected by weather	Autonomous Vehicle, mapping
5	Focus/ Defocus[13]	Hardware Simplicity	Poor accuracy , limited range	Optical system, microscopy

Table2: Comparison of different Object Detection algorithms

S. No.	Object detection Algorithms	Algorithm type	Architecture	Limitation
1	Faster-RCNN[14]	Two stage algorithm	Uses RPN for region proposal and uses multistage training pipeline	High computation power requirement Struggles in detecting small objects
2	YOLOv8[15]	One stage algorithm	Uses anchor free design with Path aggregation network and CSPDarknet based feature extractor	Struggles in detecting densely packed and overlapping objects

3	YOLO-NAS[16]	One stage algorithm	Neural Architecture Search	Not fully open source Memory optimization need further improvement to maximize performance without affecting accuracy
4	Deformable-DETR[17]	Two stage algorithm	Combines convolutional neural networks (CNNs) and Transformer encoder-decoders.	10-20 times slower than faster RCNN Not perform well in detecting small objects
5	SPARSE-RCNN[18]	Two stage algorithm	Uses query based approach along with ResNet for feature extraction along with multiple cascaded transformer decoders and	Struggles in detecting densely packed objects and high computational cost

EXPERIMENTATION

Our proposed method integrates object detection and depth estimation to measure distances between objects in a 3D space, using a reference distance (real-world measurement between two specific objects) for scaling. For the same accurate intrinsic camera parameters needs to be defined for correct camera calibration. For the same focal lengths and principal point coordinates are used. These values are placeholders that allow converting pixel coordinates to 3D coordinates. For distance calculation image is loaded by using OpenCV, objects are detected by using YOLOv8 model pre-trained on COCO dataset without fine tuning on our image dataset. YOLOv8 algorithm provides bounding box coordinates, class ID's and confidence score corresponding to every object present in the image. Yolov8 works on 2D images which only works on pixel coordinates where Z-axis is missing. In order to include missing Z-axis depth estimation has been included, so that Euclidian distance in 3D space can be calculated. For depth estimation MiDaS monocular depth estimation model is used[19]. It generates a depth map which corresponds to relative distances in the scene i.e. closer object will have lower depth values. 3D world coordinates are generated by using Pixel coordinate and depth values by using following formulae:

$$x = \frac{(x_{pixel} - Cx).z}{fx},$$

$$y = \frac{(y_{pixel} - Cy).z}{fy},$$

$$z = depth[y_{pixel}, x_{pixel}]$$

Where, x_{pixel} & y_{pixel}: centre point of detected object calculated from bounding box coordinates.

Cx & Cy: Principal point coordinates of camera.

fx and fy: Focal length in x and y direction

z: Depth value obtained from depth map.

Depth estimation MiDaS model provide relative or scaled depth values rather than absolute real-world depths. The scale of these values depends on the training data and may not correspond directly to real-world units. In order to bridge the gap between Relative and Absolute Measurements actual measurement between objects has been provided. This absolute measurement is used to calculate scale factor between the depth model's output and actual measurements. The scale factor is calculated as:

$$Scaling\ Factor = \frac{Known\ Real\ Distance}{Detected\ Distance}$$

Then all 3D coordinates are updated by applying the scaling factor to convert relative distances to real-world distances (in centimetres). Then the distance between all pair of objects has been calculated and displayed.

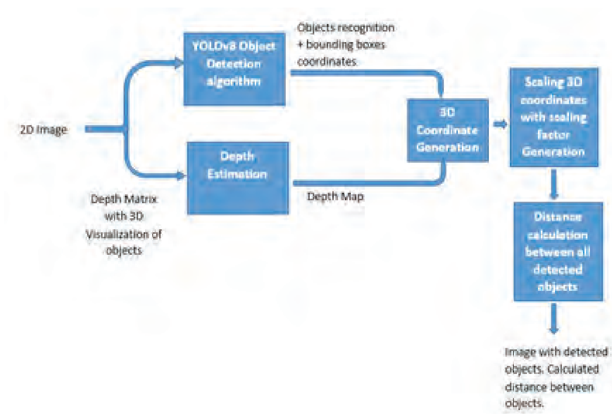


Fig. 1: Block diagram of distance calculation by integrating two models

RESULT AND DISCUSSION

We have selected laptop, mouse, book and bottle as object that needs to be detected in image. Fig2 shows the detected object and respective depth map. Depth map shows nearby objects in light colours while far away objects are shown in dark colour. Due to in-class variation in bottle the model was not able to detect bottle. We have checked the performance of combined model on around 20 images and got least error of 13.6%. This model combination has not performed well if the images are taken from some other angles.

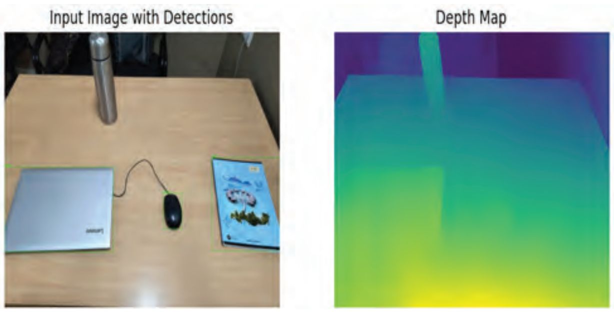
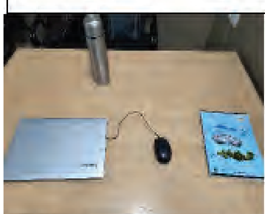
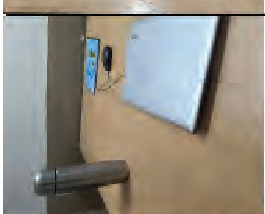
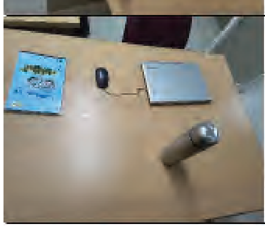


Fig. 2: Detected objects and respective depth map

The actual distance between laptop and mouse is 20 cm, mouse and book is 14 cm and laptop and book is 39 cm.

Image	Distance between objects			Relative Error %
	Laptop & Mouse (provided for scaling)	Mouse & Book	Laptop & Book	
	20	14.89	32.68	13.6
	20	22.16	2.17	54.11
	20	20.85	50.08	33.83

CONCLUSION AND FUTURE SCOPE

Relative distance measurement in images is a rapidly evolving field with significant implications for industries such as robotics, augmented reality, and autonomous navigation. This study explored the applicability of our algorithm on a limited dataset, observing a least 13.6% variation between calculated and real distances. The results highlight the dependency on accurate camera calibration, precise bounding box estimation, and reliable depth estimation using the MiDaS model. This result can be improved by training a models on object detection and depth estimation on our custom dataset so that it can handle in class variation as well. In future we can use more advanced algorithms for object detection and depth estimation along with cumulative scaling factor for distance calculation on large dataset. Our future research will likely focus on overcoming current limitations and expanding the applicability of these techniques in real-world scenarios.

REFERENCES

1. D. Kumawat, D. Abhayankar, and S. Tanwani, “Exploring Object Detection Algorithms and implementation of YOLOv7 and YOLOv8 based model for weapon detection,” Intell. Syst. Appl. Eng., vol. 12, no. 3, pp. 877–886, 2024, [Online]. Available: https://ijisae.org/index.php/IJISAE/article/view/5367

2. A. Masoumian, D. G. F. Marei, S. Abdulwahab, J. Cristiano, D. Puig, and H. A. Rashwan, “Absolute Distance Prediction Based on Deep Learning Object Detection and Monocular Depth Estimation Models,” Front. Artif. Intell. Appl., vol. 339, pp. 325–334, 2021, doi: 10.3233/FAIA210151.

3. Y.-T. Cao, J.-M. Wang, Y.-K. Sun, and X.-J. Duan, “Circle Marker Based Distance Measurement Using a Single Camera,” Lect. Notes Softw. Eng., vol. 1, no. 4, pp. 376–380, 2013, doi: 10.7763/lnse.2013.v1.80.

4. S. SOLAK and E. D. BOLAT, “A new hybrid stereovision-based distance-estimation approach for mobile robot platforms,” Comput. Electr. Eng., vol. 67, pp. 672–689, 2018, doi: 10.1016/j.compeleceng.2017.10.022.

5. Y. Bai, B. Zhang, N. Xu, J. Zhou, J. Shi, and Z. Diao, “Vision-based navigation and guidance for agricultural autonomous vehicles and robots: A review,” Comput. Electron. Agric., vol. 205, p. 107584, 2023, doi: https://doi.org/10.1016/j.compag.2022.107584.

6. A. Masoumian, H. A. Rashwan, S. Abdulwahab, J. Cristiano, M. S. Asif, and D. Puig, “GCNDepth: Self-

supervised monocular depth estimation based on graph convolutional network," Neurocomputing, vol. 517, pp. 81–92, 2023, doi: https://doi.org/10.1016/j.neucom.2022.10.073.

7. M. Kholil, I. Ismanto, and M. N. Fu'ad, "3D reconstruction using Structure From Motion (SFM) algorithm and Multi View Stereo (MVS) based on computer vision," IOP Conf. Ser. Mater. Sci. Eng., vol. 1073, no. 1, p. 012066, 2021, doi: 10.1088/1757-899x/1073/1/012066.

8. J. Yu and H. Choi, "Yolo mde: Object detection with monocular depth estimation," Electron., vol. 11, no. 1, 2022, doi: 10.3390/electronics11010076.

9. M. Ding et al., "Learning depth-guided convolutions for monocular 3d object detection," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 11669–11678, 2020, doi: 10.1109/CVPR42600.2020.01169.

10. D. Eigen, C. Puhrsch, and R. Fergus, "Depth map prediction from a single image using a multi-scale deep network," Adv. Neural Inf. Process. Syst., vol. 3, no. January, pp. 2366–2374, 2014.

11. D. Scharstein, R. Szeliski, and R. Zabih, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," Proc. - IEEE Work. Stereo Multi-Baseline Vision, SMBV 2001, no. 1, pp. 131–140, 2001, doi: 10.1109/SMBV.2001.988771.

12. Y. Furukawa and J. Ponce, "Accurate, Dense, and Robust Multiview Stereopsis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 8, pp. 1362–1376, 2010, doi: 10.1109/TPAMI.2009.161.

13. P. Favaro and S. Soatto, "A geometric approach to shape from defocus," IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 3, pp. 406–417, 2005, doi: 10.1109/TPAMI.2005.43.

14. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

15. R. Varghese and S. M., "YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness," in 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), 2024, pp. 1–6. doi: 10.1109/ADICS58448.2024.10533619.

16. J. Terven, "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," Mach. Learn. Knowl. Extr., vol. 5, no. 4, pp. 1680–1716, 2023, doi: 10.3390/make5040083.

17. X. Zhu, W. Su, L. Lu, B. Li, X. Wang, and J. Dai, "Deformable Detr: Deformable Transformers for End-To-End Object Detection," ICLR 2021 - 9th Int. Conf. Learn. Represent., pp. 1–16, 2021.

18. P. Sun et al., "Sparse R-CNN: End-to-end object detection with learnable proposals," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 14449–14458, 2021, doi: 10.1109/CVPR46437.2021.01422.

19. Ranftl and Katrin Lasinger and David Hafner and Konrad Schindler and Vladlen Koltun, "Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-Shot Cross-Dataset Transfer," IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, 2022.

Advance Automation for Transformer using Scada Data

Neeta Moharkar
Assistant Professor
Department of Electronics and Computer Science
Atharva College of Engineering
Atharva Educational Complex
Malad West, Mumbai

Gauri Vaidya
Assistant Professor
Department of Electronics and Telecommunication
Atharva College of Engineering
Atharva Educational Complex
Malad West, Mumbai

Charushila Pawar
Assistant Professor
Department of Electronics and Telecommunication
Atharva College of Engineering
Atharva Educational Complex
Malad West, Mumbai

Nilesh Shimpi
Assistant Professor
Department of Electronics and Computer Science
Atharva College of Engineering
Atharva Educational Complex
Malad West, Mumbai

ABSTRACT

Transformers are an important part of electrical transmission. Since the invention of the first voltage transformer in 1885, transformers have become indispensable equipment for the transmission, distribution and use of alternating current in power transmission lines. In this proposed system different sensors are used to detect the fault in the transformer. The SCADA and PLC combinational circuit provide required Real time operation. It is an essential step to monitor transformers before using them to avoid the risk and fault. It can reduce the cost of maintenance. This paper represents the proposed system of design and simulation of a plc based monitoring system based on SCADA to achieve computerisation to display screens to examine transformer parameters like voltage, current and voltage.

Today the functionality of Supervisory control and data acquisition (SCADA) and PLC interfacing is used to monitor the system on large application process plants where high reliability and security is important.

***KEYWORDS** : PLC, Transformer, Relay, SCADA (Supervisory Control & Data Acquisition), Sensors, Asset Management.*

INTRODUCTION

There are many special protections that depend on the value of the power distribution, including special communication to control various electrical disturbances, which often need to be adjusted according to unpredictable changes in the load. There are many disturbances in the electrical environment, and damage to transmission lines and transmission lines may occur due to natural disasters such as storms, typhoons or heavy rain. Most of the existing systems are reliable in many applications, but they are not good in terms of electricity. Real-time operation and online monitoring of transformers also require a Supervisory Control and Data Acquisition (SCADA) based system. The combination of PLC and SCADA ensures instant and reliable operation of electrical distribution equipment and accurate online monitoring of the current status of the distribution transformer of the system.

Supervisory Control and Data Acquisition (SCADA) systems collect any data from the Designed project and it can monitor the system.

SCADA interfacing with PLC can monitor the performance parameters of the transformer which are essential to evaluate the transformer safe operating condition.

This Proposed system presents automation based SCADA predictive and corrective data that can be used to monitor the performance of the transformer and help to reduce faults. Paper represent which type of data is available, and what percentage of transformers has the kind of data used for monitoring. SCADA is crucial for industrial organizations. now a days automation and AI are giving advance direction to science.[4]

These systems are used to monitor and centralize the whole working model to process variables such as

Current, Voltage and Temperature. In this proposed system We are using sensors to detect the faults as discussed in research paper ref. 2 designing and developing a Novel Distribution Automation System (DAS) in a transmission and distribution system [10].

The Introduction of modern IT and practices such as SQL (Structured Query Language) and web-based applications in SCADA software has greatly improved the efficiency, security, productivity, and reliability of SCADA systems. It allows us Ignition Integrators due to its robust and powerful nature.

With proper maintenance and operation of the power supply, the service life of the power transformer can reach 60-70 years. Programmable logic controller (PLC) automation plays an important role in reliable energy and operation. The system is an integrated SCADA and PLC system for monitoring and controlling factory control systems to ensure smooth operation and reduce the interaction between different systems. Its accuracy and efficiency play an important role in artificial intelligence and machine learning.

ASSESSMENT OF POWER TRANSFORMER IN SERVICE

Transformer condition and life are essential for transformer failure tracking. Determine if the transformer can withstand the failure. Obtain results that are used to help interpret subsequent evaluations.

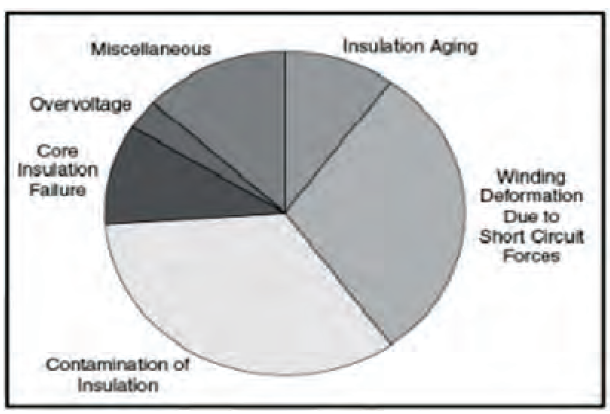


Fig. 1: Failure Data History of Transformer

When evaluating equipment, fault data on large transformers can help determine which is more important. Fault analysis of 188 power transformers in South Africa with voltages and powers of 88kV to 765kV and 20MVA to 800MVA respectively. Investigating the main components of the transformer, associated faults and diagnostic methods is important for the development of transformer inspection equipment.

A SCADA application based on the use of the most common measurement devices (temperature, gas, soil, etc.) is proposed for continuous online monitoring of transformers. It displays changes and diagnostic information online, generates alarms/alerts according to international IEC and IEEE standards[1].

CONVENTIONAL METHODS OF DIAGNOSIS

Electrical Inspections can be used to identify problems with various gas appliances including partial discharge, overheating and arcing. Insulating oil quality: A combination of physical, chemical and electrical tests are performed to assess defects, contamination and changes in electrical properties. Thermal Imager: This technology is used to measure the external temperature of the transformer. Data output from the logbook. Relays and sensors are used in practical PLC systems to identify problems of the transformer including overload, undervoltage, overvoltage, phase-to-phase and overheating problems. There is no doubt that power distribution equipment is more likely to fail, so transformer protection is important.

Use various sensors to collect real data from the environment or object. People can use the concept of Internet of Things to establish a connection between machines. Temperature, oil level and vibration are important indicators for evaluating the condition of the transformer. Transformer stability, accuracy and efficiency can be improved by using less power. With the advent of the Internet and the use of the times, people need intelligent maintenance and reliable management for all substation equipment. This can be achieved through the use of Internet of Things (IoT) technology[2].

In addition to monitoring the power transformer, the detection of winding displacement or deformation is also important. This is because these faults cause insulation changes, leading to a larger fault. Among these variables, the frequency response analysis (FRA) method is more effective in monitoring faults. In addition, the angle of the difference between the winding voltages can be used to determine the position of the power transformer windings, which makes it possible to determine the effect of radial displacement of the windings [5]. The continuous diagnostic method for large power transformers is the result of two main directions: testing and test operation. Fault testing is usually done by taking the generated signal with the defined behavior as input to the faulty switch while the switch is in use and measuring the output signal.

METHOD OF RESEARCH

A programmable logic controller (PLC) is used in the project's architecture to regulate and track the load conditions of the transformer's three phases.

Online SCADA simulation is used for monitoring.

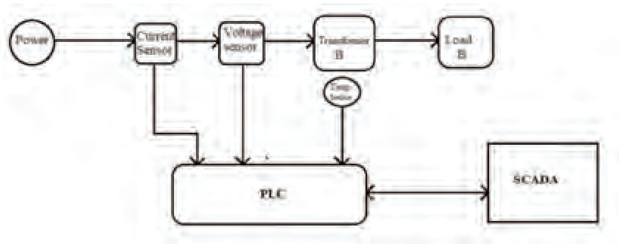


Fig. 2: Block Diagram of System

Voltage control is used to monitor the power, voltage and current of each phase when the load is connected. A transformer (PT) is used to measure the voltage of each phase to monitor the voltage amplitude, while a CT (current sensor) is used to measure the current level of the load. The system works smoothly under load and within range. When the load reaches the critical point and reaches the overload mark, the system will automatically shut down to prevent damage or destruction of the transformer. Information on overload conditions caused by short circuit overloads and overheating conditions is sent simultaneously to the sensors controlled and monitored by SCADA.

CIRCUIT DIAGRAM

Following Circuit diagram represents the interfacing of components.

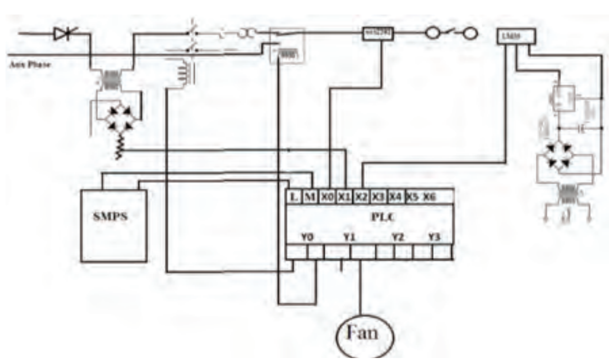


Fig. 3: Circuit Diagram of System

INTERFACING SCADA

By providing the user interface and all updates about the temperature, current and voltage, the SCADA interface will reduce the amount of work required by humans. Advanced monitoring systems based on SCADA have more sophisticated futures, and protection and control features are always changing.This article examines the optimization and duplication of substation SCADA and IEDs in the substation, including binary/digital inputs (status), binary/contact outputs (control), and analog inputs (metering).

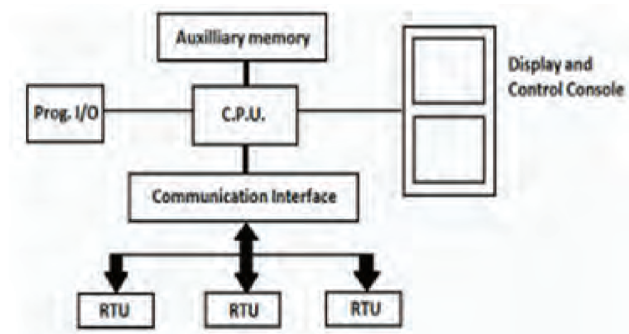


Fig. 4 Functional Diagram of SCADA

The software used by SCADA is "InTouch". Wonderware's InTouch is advanced SCADA software for monitoring and backing up large areas of production process data. Devices managed from a single workstation can be connected via OPC clients, S7 MPI, S7 PPI, Profinet (S7 1200), ModbusRTU, Modbus TCP/IP, Host System Connect (Omron), Mewtocol protocol (Panasonic). Historical data related to our SCADA system is archived.

RESULTS AND SIMULATION

Designed system use to detect overvoltage, undervoltage, overcurrent, undercurrent and over temperature faults on the primary side of the transformer (HV), as well as undervoltage, overvoltage and overcurrent faults on the secondary side of the transformer.

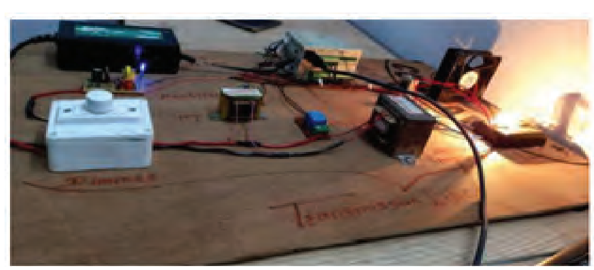


Fig. 5 Hardware of Proposed System

Current Analysis

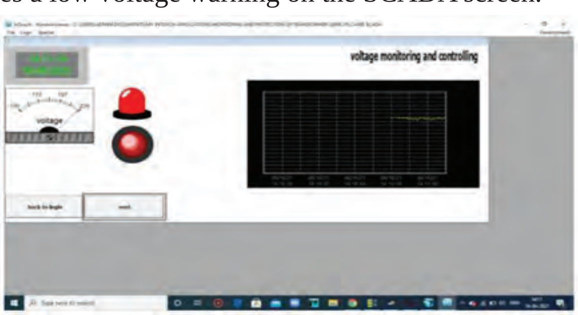
The secondary winding is connected directly to the load. The current sensor is located between the load lamp and the transformer secondary winding. We have a light load in the middle. When the electricity is applied, the light will come on. For overload, we will add an additional lamp with the help of a button and when the second lamp is turned on, the current level will exceed the value set in the PLC. Due to the current increase, the relay will trip and a short warning will appear on the SCADA screen.

a. Current Monitoring Window Without Fault

b. Current Monitoring Window When Fault Occurs

Voltage Analysis

We supply Power to the main transformer via a dimmer. We use a dimmer to change the primary voltage from 0 volts to 230 volts. When the voltage drops below 180 volts during the power change, the relay trips the circuit and gives a low voltage warning on the SCADA screen.



a. Voltage Monitoring Window Without Fault

b. Voltage Window When Fault Occurs

Temperature Analysis

For temperature sensing we have used LM35 temperature sensor. When the temperature will increase LM35 will sense the increased temperature and it will give signals to the fan so that fan starts and it turns OFF automatically when the temperature is decreased.

a. Temperature Window Without Fault

b. Temperature Window With Fault

CONCLUSION

A PLC based system has been developed for this task to monitor, control and protect the voltage, current and temperature on both sides of the transformer. The constant value of the transformer is constantly monitored by the PLC system connected with relays in the three phase system, and when the PLC detects an increase or decrease in the voltage, current or temperature level, the unit will be forced to shut down. This prevents further damage to the machine. According to the design, the power distribution equipment is more powerful against some serious problems that cause electricity, current or temperature increase. Therefore, this plan increases the stability, reliability and efficiency of distribution.

The SCADA interface will provide users with all updates related to thermal oil and fluids, reducing the amount of manual effort required. The future of SCADA-based monitoring systems is more complex than just prevention and control. This article examines the optimization and replication of substation SCADA and IEDs, including binary/digital inputs (status), binary/contact outputs (control), and analog inputs (metering).

REFRENCES

1. Arief Basuki,UK,IEEE 2018, 53rd International Universities Power Engineering Conference (UPEC)
2. Sanjay S. Tippannavar, IEEE, 2023 International Conference on Recent Trends in Electronics and Communication (ICRTEC),Mysore, India.
3. H. Jim Sim, D. L. Harris, "Current practices on power transformer monitoring and controls for smart grid applications" IEEE Power and Energy Society General Meeting, Detroit Michigan, USA, 24-28 July 2011.
4. Larsson, J. E. and S. Lee, "Managing Information Overload in Power Grid Control Centers," 9th International Workshop on Electric Power Control Centers, EPCC 9, Ullensvang, Norway, 2007.
5. J. C. Carneiro, J. A. Jardini, J. L. P. Brittes, "Substation power transformer risk management reflecting on reliability centered maintenance and monitoring", 2012 Sixth IEEE/PES Transmission and Distribution: Latin America Conference and Exposition (T&D-LA), Uruguay 03-05 September 2012
6. E. Aburaghiega, M. E. Farrag, D. M. Hepburn, B. Garcia, "Power Transformer health monitoring: A Shift from off-line to on-line detection" International universities Power Engineering Conference (UPEC), Staffordshire University, UK, 01-04 September 2015.
7. J. Q. Feng, P. Sun, W. H. Tang, D. P. Buse, Q. H. Wu, Z. Richardson, J. Fitch, "Implementation of a power transformer temperature monitoring system" in proc. International Conference on Power System Technology, PowerCon, Kunming China, 2002.
8. A. Muruganandhan D.,2020 International Conference on System, Computation, Automation and Networking (ICSCAN)Pondicherry, India.
9. Cigré WG A2.37, "Transformer reliability survey", Cigré Technical Brochure no. 642, 2015.
10. N. Baranovic, P. Andersson, I. Ivankovic, K. Zubrinic-Kostovic, D. Peharda, J.E. Larsson, "Experiences from Intelligent Alarm Processing and Decision Support Tools in Smart Grid Transmission Control Centers", 2016 Cigre Session, D2-112.

A Review of SEM & TEM Characterization Methods for Nanomaterials

Abhilasha Saini
Asst. Professor
Dept. of Applied Sciences
Atharva College of Engineering
Mumbai, Maharashtra
✉ abhilashasaini@atharvacoe.ac.in

Pinky Steffi A
Asst. Professor
Dept. of Applied Sciences
Atharva College of Engineering
Mumbai, Maharashtra
✉ pinkysteffi@atharvacoe.ac.in

P. N. Nemade
Professor
Dept. of Applied Sciences
Atharva College of Engineering
Mumbai, Maharashtra
✉ drpnnemade@yahoo.co.in

Priyanka Badani
Asst. Professor
Dept. of Applied Sciences
Atharva College of Engineering
Mumbai, Maharashtra
✉ priyankapatel@atharvacoe.ac.in

ABSTRACT

Nanotechnology brings up the manipulation, characterization, and engineering of materials at the Nano range, typically varying from one to one hundred nanometers. This interdisciplinary domain holds profound potential across various fields, inclusive of chemical science, electronics, physical science, morphology, and material science. The properties of nanomaterials often diverge significantly from their bulk counterparts, resulting in unique behaviors and applications. Contemporary research in nanoscience primarily focuses on fine-tuning the structural properties of nanoparticles and integrating them into larger material systems. This paper focuses on imparting a succinct overview about the diverse methodologies employed for characterization and synthesizing nanomaterials.

KEYWORDS *: Nanoscience, Electron microscopy, Scanning & tunneling electron microscopy, Characterizations.*

INTRODUCTION

Materials possessing nanoscale acreage have garnered significant attention in recent decades, because of their elevated surface-to-volume proportionality and diverse fundamental attributes. The rapidly advancing, interdisciplinary extent of Nano-science exerts a profound influence across a wide array of scientific and technological disciplines, including chemical & physical science, polymer science, biomedical field and even aerospace domain.

Numerous dimensional shapes, such as zero, one, two, and three-dimensional structures, are displayed by nanomaterials. One-dimensional nanotubes and nanowires are essential components of hybrid molecular and semiconductor devices, according to recent studies [1]. Metal sulphides and oxides are commonly used to create quantum dots. When analyzing characteristics including optical, electrical, mechanical, and electronic behaviors as well as real-world applications, nanoscale dimensions are crucial. Cutting tools and heat-resistant materials benefit greatly from Nano coatings because of their remarkable hardness, which surpasses that of steel and extraordinary thermal resistance [2].

CHARACTERISATION OF NANO MATERIALS BY MICROSCOPY TOOLS

To assimilate more about nanoparticles, especially those on facets, a variety of strategies are implemented. Scanning electron microscopy (SEM) and transmission electron microscopy (TEM) shed light on the dispensation, expanse, profile, and topology of particles. The expanse, profile and phase can all be established by electronic diffraction approach. Nanoparticle dimension and interior configuration can be ascertained with the use of Scanning Tunneling Microscopy (STM). The tool of gas adsorption mechanism aids in determining the particle's surface-to-volume proportion. Techniques like electron diffraction, TEM, and SEM are especially helpful for examining particles in bulk. By employing the accelerated electron beams, electron microscopy is a potent and sophisticated technology that allows for high-resolution examination of the morphology and characteristics of solid surfaces. It includes a range of probe microscopy techniques, each with unique benefits of exterior investigation.

Scanning Electron Microscope (SEM)

In a SEM, accelerated electrons possess substantial energy, which dissolute via a large extent of indications resulting from specimen's interactivities with the beam of electrons because the falling electrons get slacked up within the firm illustrative specimen. Furthermore such interactivities generate a variety of signals, including secondary generated electrons superintend for the shaping of sample representations, the backward scattered electrons, which are instrumental in determining the crystallographic conformation and orientations [3]. Moreover, these interactivities bring out the photons, including typical X-radiations for primordial exploration, also the continuous sequential X-radiations, observable region of light, and thermal energy. Each of these emitted signals provides crucial information for the comprehensive characterization of the sample. Scanning electron microscopy (SEM) frequently uses backscattered and secondary electrons for imaging. Backscattered electrons have a crucial role in emphasizing compositional differences, particularly in multiphase samples, whereas secondary electrons are especially useful for exposing the morphology and topography of samples, offering high-resolution surface details [4]. Inelastic collisions between input electrons and the electrons in the sample's distinct atomic shells produce X-rays in SEM. Excited electrons release X-rays with distinctive wavelengths as they return to lower energy states. It is therefore possible to identify the elements in the sample by analyzing the characteristic X-rays that are produced when each element is "excited" by the electron beam. These X-rays have different energies.

The X-ray radiations emitted via the electron interactivities in SEM analysis are considered "non-destructive," which means that no appreciable material volume is lost, enabling repeated investigation of the same sample [5]. The SEM can use magnifications of up to about 100,000 times to image the surface architecture of bulk objects in a variety of fields, such as biology, medicine, materials science, and earth sciences. Compared to optical micrographs, these microscopic representations are of a substantial deepness of area of activity and resolution, which makes them ideal for examining tough specimens like cracked surfaces and particle matter.

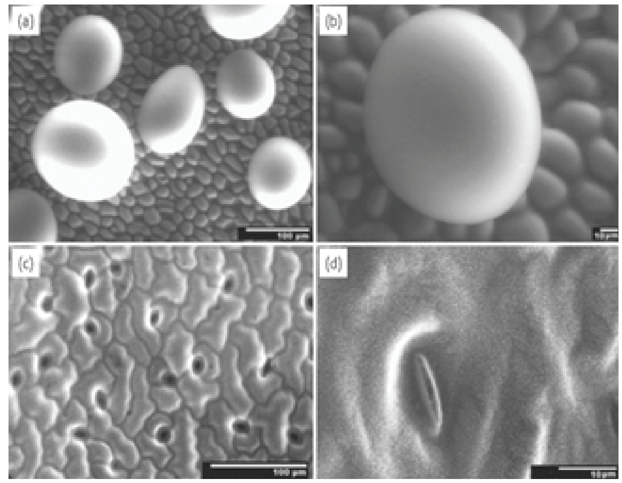


Fig. 1: Uncoated specimen leaf surfaces (a) chenopodium leaf (c) pea leaf surface with intact epidermal cells

Transmission Electron Microscopy (TEM)

This mechanism involves passing an electron beam through an extremely thin specimen, causing the sample to interact with the beam. The measured electrons that move through the sample are used in this interaction to create a picture. As the electron's wavelength is comparatively undersized that of photons, they allow for higher resolution than traditional light microscopy. An enhanced and boosted representation and a respective spectrum of diffraction phenomenon are the two different forms of information that Transmission Electron Microscopy (TEM) provides about the specimen. These two sorts of information are extremely useful in determining the composition and structure of the sample.

In this technique, the electron beam is precisely focused into an exceedingly narrow stream, which is transmitted through an ultrathin illustrative and thereupon extrapolated onto a display, thereby generating a replica representation of such chosen specimen. Thereafter the distinction of the derived depiction is intrinsically associated with the density magnitude of the chosen matter. Due to the significantly shorter wavelength of electrons, Transmission Electron Microscopy (TEM) offers markedly superior resolution compared to conventional light microscopy [6]. TEM serves as a cornerstone analytical method across a diverse array of scientific disciplines, encompassing both physical and biological sciences. Its applications extend to multidimensional areas [7].

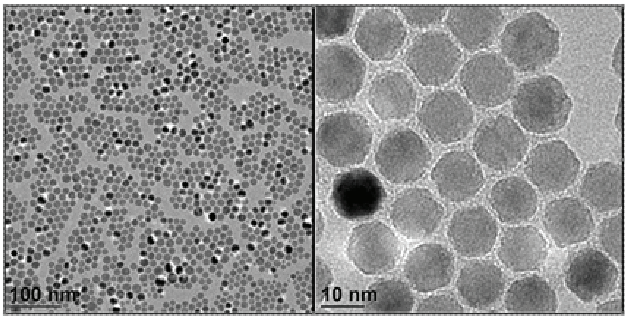


Fig. 2: TEM pictures of organic-phase-synthesized nanoparticles. There are two distinct magnifications displayed: 105000x and 810000x

The dissimilarity observed for any display of sample generated via the tunneling microscopy diverges fundamentally from that of a light microscope. In crystalline materials, the primary interactivity along with those falling electrons occurs as they get diffracted than that getting absorbed in a sample. Nonetheless, the electron's energy even after transmitting through the specimen is still influenced with its volume as well as density. As a result a very high resolution and magnified display can thus be achieved via selectively obstructing electrons that are thrown sidetrack to the microscope's axis, using an aperture that permits only un-scattered electrons to pass through. This process induces variations in electron intensity that provides critical insights into the specimen's crystalline structure, revealing intricate details of atomic arrangement.

Contrast of the nanotechnology techniques of SEM and TEM

(SEM) and (TEM) are both pivotal techniques in nanotechnology, yet they differ fundamentally in their operational principles and the type of information they provide. SEM primarily generates surface representations via the scan of the chosen illustrative using a pointed and concentrated shaft of electrons, offering superior depth of field and the ability to resolve topographical features at nanometer scales. It excels in imaging bulk surface morphology and compositional contrasts, but provides limited insight into internal structures. In contrast, TEM transmits electrons through an ultrathin specimen, allowing for the visualization of internal microstructural details at unprecedented resolution, owing to the shorter wavelength of electrons. TEM is particularly adept at revealing crystallographic information and atomic arrangements, offering both magnified imaging and diffraction patterns. While SEM is advantageous for surface analysis, TEM provides deeper, more intricate insights into the internal architecture of materials, making both indispensable for comprehensive nano-structural characterization.

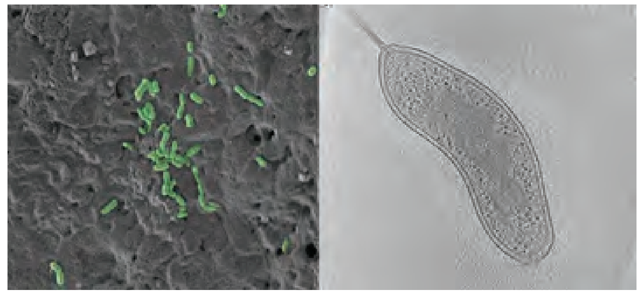


Fig. 3: TEM (right) and SEM (left) pictures of bacteria. In contrast to SEM, which displays a large number of bacteria on a surface, TEM displays the internal structure of a single bacterium

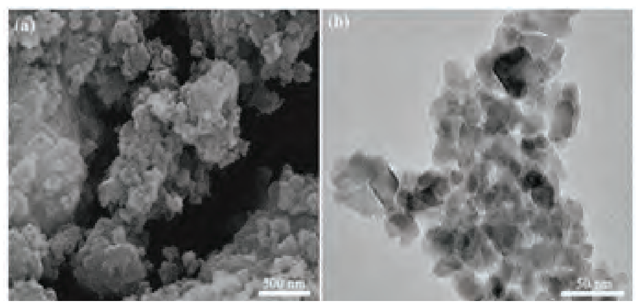


Fig. 4: Samples synthesised with varying quantities of rare earth elements but the same ratio of benzyl alcohol to 1-octylamine are displayed in SEM and TEM images (a, b)

RESULT & DISCUSSION

The comparison of (SEM) and (TEM) for the analysis of nanomaterials reveals distinct yet complementary advantages inherent to each technique. SEM provides high-resolution surface imaging, offering unparalleled depth of field and topographical details, making it ideal for examining the external morphology and compositional contrasts of nanomaterials. Its capacity for imaging bulk samples with minimal sample preparation is a notable strength, particularly for complex, irregular surfaces. Conversely, TEM offers superior resolution due to its transmission of electrons through ultrathin specimens, facilitating the visualization of internal structures at atomic scales. This enables detailed analysis of crystallographic orientations, lattice defects, and nanoparticle distributions, providing a level of structural insight that SEM cannot achieve. While SEM excels in surface characterization and is more accessible for routine analyses, TEM is indispensable for probing the fine internal architecture and nanoscale crystallography of materials. Together, these techniques provide a comprehensive toolkit for the precise characterization of nanomaterials, with SEM being more suited for surface analysis and TEM for in-depth structural evaluation.

REFERENCES

1. Huang, Y., Xiangfeng Duan, Yi Cui, Lincoln J. Lauhon, Kyoung-Ha Kim, Charles M. Lieber, Logic gates and computation from assembled nanowire building blocks, Sci., 294(5545), 1313-1317(2001). https://doi.org/10.1126/science.1066192
2. Avadhani G. S., Characterization of Nano Materials using Electron Microscopy, The "Indo-Russian workshop on nanotechnology and laser induced plasma(2009).
3. Goldstein, J., Scanning electron microscopy and x-ray microanalysis, Kluwer Academic/Plenum Publishers, 689(2003).
4. Reimer, L., Scanning electron microscopy: physics of image formation and microanalysis, Springer, 527(1998).
5. Egerton, R. F., Physical principles of electron microscopy: an introduction to TEM, SEM, and AEM, Springer, 202(2005).
6. Egerton, R., Physical principles of electron microscopy, Springer, 0387258000(2005).
7. Rose, H. H., Optics of high-performance electron microscopes, Sci. Technol. Adv. Mater., 9(1), 01- 08(2008).

Comix-Gen AI: Text to Comics Generation

Ranjana Singh
Department of Computer Engineering
Watumull Institute of Engineering & Technology
Thane, Maharashtra
✉ ranjanawatu2010@gmail.com

Hitakshi Trivedi
Department of Computer Engineering
Watumull Institute of Engineering & Technology
Thane, Maharashtra
✉ hitutrivedi9@gmail.com

Jyoti Singh
Department of Computer Engineering
Watumull Institute of Engineering & Technology
Thane, Maharashtra
✉ jyotisingh190402@gmail.com

Aditya Singh
Department of Computer Engineering
Watumull Institute of Engineering & Technology
Thane, Maharashtra
✉ jsinghadityag1006@gmail.com

ABSTRACT

Comic generation is an innovative approach to storytelling that leverages the power of machine learning and deep learning techniques to automate the creation of comic panels. This paper presents Comix-Gen, a comprehensive system designed to transform user-provided text descriptions into fully realized comic illustrations featuring characters, environments, and dialogues. The system employs advanced methodologies, including Natural Language Processing (NLP) for text classification, Generative Adversarial Networks (GANs) for image generation, and Convolutional Neural Networks (CNNs) for feature extraction and style consistency. By integrating these components, Comix-Gen addresses the challenges of visual coherence, character representation, and dialogue authenticity, paving the way for a seamless comic creation experience. Through extensive experimentation, the system demonstrates high accuracy in character and emotion detection, producing visually compelling and narratively coherent comic panels. This research not only contributes to the field of automated comic generation but also inspires new avenues for interactive storytelling in the digital age.

KEYWORDS *: Comic generation, Machine learning, GANs, CNNs,NLP, Text classification, ML, DL.*

INTRODUCTION

Comics are a unique storytelling medium that merges visual art with narrative depth, engaging audiences across different demographics. Over the years, comics have evolved significantly, adapting to cultural and technological advancements. Today, they encompass a wide range of genres, including superhero sagas, satirical works, and dramatic storytelling, captivating readers worldwide.

With the rapid expansion of digital platforms, the demand for innovative tools to streamline comic creation has increased. Traditionally, developing a comic requires extensive effort, involving collaboration between writers, artists, and editors to bring a story to life. However, advances in machine learning (ML) and deep learning (DL) offer new possibilities for automating major aspects of this creative process. By leveraging these technologies, it becomes possible to generate high-quality comic content while maintaining narrative coherence and artistic integrity.

A key technology enabling this automation is Generative Adversarial Networks (GANs) [4], which have proven highly effective in image synthesis. GANs consist of two neural networks—the generator and the discriminator—working together to produce realistic images by learning from existing datasets. While GANs have demonstrated remarkable success in generating lifelike photographs and digital art, applying them to comic generation presents distinct challenges. The unique visual language of comics—featuring bold outlines, exaggerated expressions, and stylized color schemes—requires specialized adaptations of GAN-based models. GANs have been successfully used for generating synthetic data in various fields, including tabular datasets [7][24], audio synthesis [6], and image creation [15].

Additionally, integrating Natural Language Processing (NLP) techniques is essential for converting text descriptions into visual content. NLP enables the system to analyze user input, identifying key components such as characters, settings, and dialogues—essential elements in crafting meaningful comic panels. This text-to-image generation process not only requires a strong linguistic understanding but also the ability to convey emotions, humor, and action effectively, all of which are central to comic storytelling.

In light of these advancements, this study introduces Comix-Gen, an AI-driven system designed to generate comic panels from textual descriptions. The main objectives of Comix-Gen include automating character design, creating relevant environments, and producing contextual dialogues that enhance storytelling. By tackling challenges related to visual coherence and contextual accuracy, Comix-Gen provides an intuitive platform that simplifies the comic creation process while preserving its artistic essence.

This research explores the potential of deep learning models in the field of comic generation, investigating methodologies that bridge the gap between textual input and visual representation. Through this study, we aim to demonstrate the feasibility of automated comic generation as a valuable tool for artists and storytellers in an increasingly digital world.

RELATED WORK

Artificial image synthesis focuses on generating images based on predefined constraints or inputs. This research area continues to evolve, with various approaches being explored. While some recurrent neural network (RNN)-based methods [5] have demonstrated promising results, Generative Adversarial Networks (GANs) have emerged as the most effective and efficient machine learning framework for image generation [10][19]. This section reviews key studies, divided into two primary categories: Text-to-Image GANs and comic feature extraction using Convolutional Neural Networks (CNNs).

Numerous text-to-image GAN models [15] have been designed to generate detailed images from descriptive text. The foundational framework for this process consists of a basic architecture with a single generator-discriminator pair [14]. The text input undergoes encoding, and the generated image aligns with the semantic content of the description, with a discriminator evaluating the accuracy of this correspondence.

Subsequent advancements, such as MirrorGAN [12], AttnGAN [21], and StackGAN [23], introduced additional generator networks that refine the image generation process at the word level. These models also integrate multiple discriminators to ensure that the generated visuals accurately depict both the overall text input and its specific components.

Feature extraction and encoding of the resulting images play a crucial role in improving generation quality. High-level features, along with a final vector representation, enhance the attentiveness of image synthesis and help verify alignment between the visual output and its textual input. AttnGAN [12][21], for instance, incorporates an image encoder based on the Inception v3 CNN model [18], which was pre-trained on the ImageNet dataset [2]. However, for comic-specific applications, replacing this encoder with a CNN optimized for stylized art may yield better results.

Studies on manga object detection [22][11][13] highlight the effectiveness of CNN-based feature extraction for comic illustrations. While convolutional architectures like VGG and Inception v3 [17][18] excel in feature extraction and classification, they can present challenges in implementation due to their complexity.

Generating images from text using GANs is particularly challenging when applied to comics, as conventional models often struggle to capture the stylized nature of comic illustrations. Standard text-to-image solutions may not produce visually appealing comic panels without modifications. As a result, adapting both GAN and CNN architectures with specialized enhancements is necessary to achieve higher-quality comic generation.

PROPOSED WORK

Comic Illustration Synthesis

This section describes the methodologies used to extract and generate text descriptions corresponding to comics, adapt AttnGAN for comic-based text-to-image generation, and develop a CNN-based image encoder optimized for comic art. The following subsections detail the implementation of each aspect.

AttnGAN Architecture

AttnGAN [21] is a state-of-the-art GAN framework designed for image synthesis from textual descriptions. It is trained on datasets that pair images with multiple text captions, allowing it to learn the relationship between textual and visual elements. For this study, AttnGAN was replicated and assessed using original text-image datasets

[20], serving as a baseline for experiments in comic generation. Various modifications and optimizations were applied to enhance its performance in generating high-quality comic-style illustrations.

AttnGAN's framework consists of multiple generator-discriminator pairs, each leveraging encoded text and image data. Initially, a Deep Attentional Multimodal Similarity Model (DAMSM) is trained, incorporating both a text encoder and an image encoder. The text encoder, built on a Recurrent Neural Network (RNN), converts input text into a feature vector representation, while the image encoder, which uses a Convolutional Neural Network (CNN), processes images into corresponding feature vectors. The DAMSM helps establish a strong correlation between the text input and the generated images, enabling the model to associate textual attributes (e.g., "red" in the description) with corresponding visual elements in the generated output.

Following the training of these encoders, the GAN model itself is developed. The initial generator produces a rough image based on the fully encoded text input, while subsequent generator layers refine details by integrating encoded representations of individual words from the text. Each stage outputs progressively higher-resolution images, evaluated by discriminators to ensure fidelity to the input description. The DAMSM further fine-tunes the model by comparing generated images with input text descriptions, minimizing error through backpropagation.

ComicGAN extends AttnGAN's capabilities by incorporating a text-generation pipeline for labeled comic images. It improves the comic creation process by integrating a CNN-based image encoder explicitly designed for comic-style illustrations.

CNN-Based Comic Feature Extraction

To build a specialized feature extractor for comic images, a modified Inception v3 [18] CNN was trained on the ImageNet dataset [2] for multi-class classification. This model, with adjusted final dense layers, was integrated within AttnGAN as a dedicated feature extractor. A transfer learning approach was used, training a multi-label classifier on a comic-specific dataset to improve feature extraction accuracy and enhance the GAN's ability to generate high-quality comic-style images.

The Comics Classifier was developed and tested using labeled attributes derived from comic illustrations, such as character colors and background details. Since training integrated image encoders within AttnGAN is computationally expensive and time-consuming, the effectiveness of feature extraction was evaluated through classifier accuracy. A higher classification accuracy indicated a more effective extraction of relevant visual features.

The image encoder used in AttnGAN is based on the Inception v3 architecture, designed to extract 17x17 regional features that capture localized elements such as character outlines, object placements, and color distributions. To better suit comic-style illustrations, a CNN specifically optimized for comic feature extraction was developed and integrated into the model.

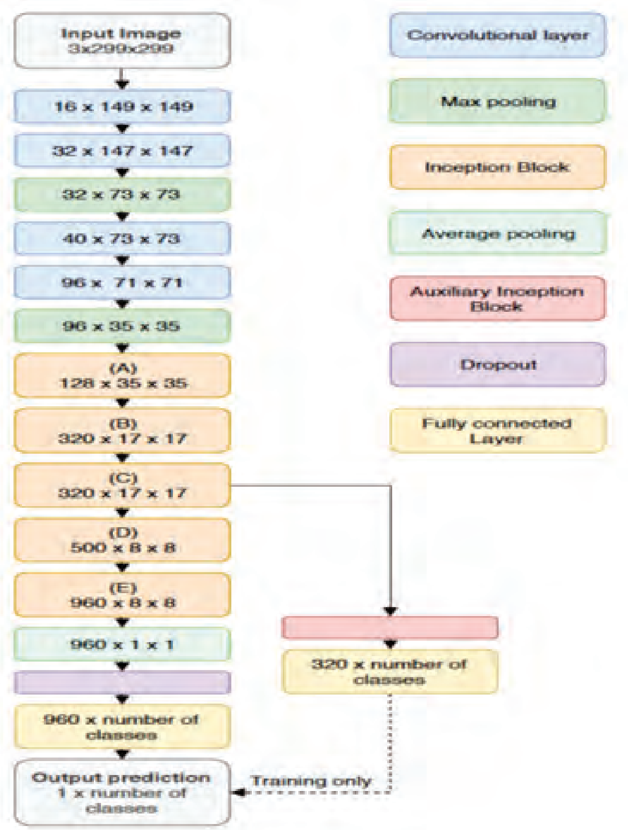


Fig. 1: Comics CNN Architecture

Input dimensions for each layer are displayed. Colored layers indicate types as per the legend on the left. The input undergoes convolution and max pooling, progressively extracting features, reducing spatial size, and increasing depth. Inception blocks enhance feature extraction by capturing multi-scale representations using varied stride lengths. The 17x17 mid-level features are flattened and passed into a fully connected layer for classification.

Comics CNN Architecture

The Comics CNN follows a structured architecture designed for efficient feature extraction. The input image, set at 299x299 pixels with three color channels (red, green, blue), passes through two convolutional blocks, each comprising two 2D convolutional layers followed by a max pooling layer. This reduces the spatial dimensions from 299 to 35 while increasing feature depth to 128.

Each inception block consists of three convolutional branches derived from the input. These branches apply different kernel sizes to process varying scales of features. Some branches use convolutional layers with different stride lengths, while others employ average or max pooling before concatenation. This technique enables the model to efficiently extract hierarchical feature representations while maintaining computational efficiency [18].

Three stacked inception blocks deepen the network and reduce spatial dimensions to 17. These outputs, representing 17x17 feature regions, undergo further processing through two additional inception layers. The final layers consist of average pooling, dropout, and a flattening operation. A fully connected layer generates the final classification results.

Following a similar approach to Inception v3, the Comics CNN leverages extracted feature regions during training to refine model performance. The 17x17 features serve as intermediate representations to guide the training process. The model's loss function is designed with binary cross-entropy (BCE) loss between predicted and actual labels, combined with fractional feature loss to ensure the network effectively captures representative characteristics of comic-style images.

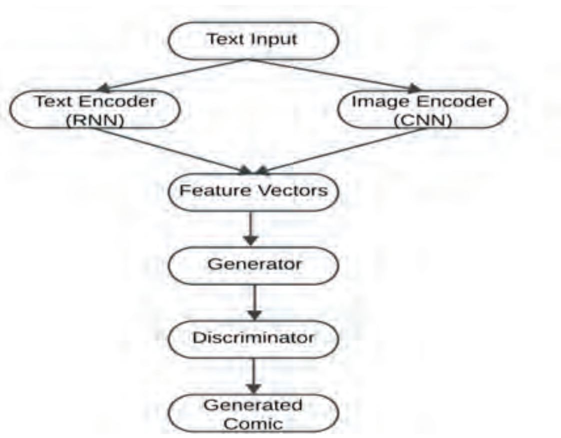


Fig. 2: Overview of COMIX-GEN Architecture

CHALLENGES IN COMIC GENERATION USING GAN TECHNOLOGY

Generating comics with GANs presents unique challenges despite their success in producing photo-realistic images. This section explores why conventional text-to-image GAN methods struggle with comic generation.

Text-Image Correlation in Comics

Using comic dialogue as input text for image generation initially seemed promising but revealed significant challenges. Text-to-image models typically require text that directly describes image elements, like objects or scenes, creating a clear semantic link. However, comic dialogue lacks such descriptive context, making text-image alignment complex.

Additionally, automatic description generation tools trained on real-world images failed to produce meaningful comic descriptions. Custom models were needed due to the stylistic differences between comics and real-life photos. Manual dataset creation became necessary, raising questions about optimal annotation types, content, and composition to ensure model learning effectiveness.

Feature Extraction Challenges

Comics differ significantly from real-world images. Their features are more stylized: bold outlines, distinct character designs, and simplified backgrounds with solid colors. Computer vision models designed for photo-realistic images struggled when applied to comics, highlighting the need for custom feature extraction techniques tailored to the unique visual traits of comics.

Evaluation Complexity

Evaluating GAN-generated comics is inherently difficult due to the lack of objective loss functions. Standard metrics like Inception Score (IS), designed for photo- realistic images, are unsuitable for comic evaluation. Since comic synthesis from text is relatively unexplored, establishing meaningful baselines for comparison remains a challenge.

Dataset Limitations

The absence of comic datasets with detailed text descriptions posed a significant barrier. Manually labeling comic panels was time-consuming and only partially automated using character and color labels. Limited time allowed labeling of just 2,500 Dilbert comics, reduced to 1,000 usable panels after filtering. Unlike standard image datasets, typical data augmentation methods like rotation and cropping are unsuitable due to comics' fixed design characteristics.

Future efforts should focus on expanding datasets with multiple comic sources and creating richer text descriptions. This would support more comprehensive evaluations and improve model performance, enabling better comic generation using GANs.

OVERVIEW OF PROPOSED EXPERIMENTS

The proposed experiments aim to evaluate the effectiveness of Comics CNN and ComicGAN in generating high-quality comic panels based on text descriptions. The experiments will leverage datasets of comic panels and associated descriptions to explore the performance of different generative models.

Dataset Preparation

1) Image Dataset:

The primary dataset will consist of cleaned comic panels, with text removed from the images to focus solely on the visual content. Non-recurring and infrequent characters will be filtered out to maintain consistency, which will also ensure that the model has enough training data for each character. For instance, side characters appearing infrequently in comics like "Garfield" were excluded due to insufficient data. Unlike standard image datasets, typical data augmentation methods like rotation and cropping are unsuitable due to comics' fixed design characteristics.

2) Text Dataset:

Several text datasets will be created to train the models, including:

I. Comic Dialogue: Automatically extracted from comic transcriptions.

II. Character Descriptions: Manually created to describe specific visual attributes of characters (e.g., "Man is wearing a white shirt and red tie").

III. General Comic Descriptions: General descriptions of the scene, including background and character placements (e.g., "Bob and Alice standing with a green background").

Evaluation Metrics

The following metrics will be used to evaluate theperformance of both Comics CNN and ComicGAN

I. Fréchet Inception Distance (FID): This metric measures the similarity between generated comic images and real comic images. Lower FID scores indicate that the generated images closely resemble real comic panels.

II. Inception Score (IS): Commonly used to evaluate photo-realistic images[16], IS may not be as effective for comics due to their stylized nature. Still, it may provide some insight into the diversity of generated images.

III. F1-Score and Accuracy: These metrics will be used to evaluate the multi-label classifiers in Comics CNN, assessing the accuracy in predicting character labels, background elements, and overall comic structure.

IV. Generator and Discriminator Losses: During training, these losses will be monitored to check for convergence and diagnose issues like vanishing gradients. They will also help evaluate the stability of the GAN models.

CONCLUSION & FUTURE WORK

The aim of this research is to design a text-to-image Generative Adversarial Network (GAN) for comic generation, capable of producing comic panels from textual descriptions. The envisioned pipeline aims to transform text inputs into comic illustrations by building upon existing methodologies and incorporating innovative enhancements tailored specifically for comic art. The proposed model, ComicGAN, integrates automatic text description generation and a custom Convolutional Neural Network (CNN) for feature extraction, specifically optimized for the stylized features found in comics.

Although the research is currently in the design and development stage, theoretical exploration suggests that this approach holds significant potential. The proposed pipeline leverages deep learning models like StabilityGAN and ComicGAN, which are expected to show notable improvements in image quality, as measured by Fréchet Inception Distance (FID) scores, compared to baseline models commonly used in image generation tasks. A core element of the approach involves generating detailed text annotations for comic panels, enabling the development of an effective training dataset for the text-to-image model.

Additionally, using a Comics CNN specialized in extracting visual features from comic art could provide a performance advantage over general-purpose models such as Inception

V3. This is due to the specialized model's ability to better capture unique comic-style features, including bold outlines, exaggerated expressions, and limited color palettes. If a sufficiently large and diverse comic dataset is curated, this custom CNN could outperform traditional architectures in terms of accuracy and feature extraction efficiency.

Looking ahead, several potential avenues for improvement have been identified. One promising direction involves developing a Multi-Label Conditional GAN, where labeled image datasets could be used instead of relying solely on text descriptions. This approach would allow label vectors to be input alongside text prompts, simplifying the training process while improving the correspondence between the input and the generated comic panels. Such a model would enable more precise control over panel attributes like character appearance, setting, and emotion.

Another potential extension involves adapting StoryGAN to support the sequential creation of entire comic strips. This adaptation would enable the generation of multi-panel comics with a coherent storyline and consistent character appearances across panels, facilitating deeper narrative development. The sequential generation approach would help bridge the gap between isolated panel generation and full comic strip creation, providing a more comprehensive solution for storytelling through AI.

Moreover, exploring alternative GAN architectures such as StyleGAN or BigGAN could further enhance the image resolution and visual consistency of generated comic panels. Incorporating advanced techniques like attention mechanisms and multimodal transformers could also improve the model's ability to align generated visuals with corresponding textual descriptions more effectively.

Overall, while the research remains in its conceptual and development stages, the proposed models and methodologies are theoretically sound and grounded in established deep learning principles. Future experimental validation, dataset preparation, and model fine-tuning are expected to provide deeper insights into the feasibility and scalability of this approach, ultimately contributing to advancements in AI-driven comic generation systems.

REFERENCES

1. A. Borji, "Pros and cons of GAN evaluation measures," Computer Vision and Image Understanding, vol. 179, pp. 41–65, 2019.
2. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2009, pp. 248–255.
3. I. Developer Staff, "Image caption generator," 2018.
4. I. J. Goodfellow et al., "Generative adversarial networks," arXiv preprint arXiv:1406.2661, 2014.
5. K. Gregor et al., "DRAW: A recurrent neural network for image generation," in Proc. Int. Conf. Mach. Learn., vol. 37, 2015, pp. 1462–1471.
6. J. Kong, J. Kim, and J. Bae, "HiFi-GAN: Generative adversarial networks for efficient and high fidelity speech synthesis," in Proc. Adv. Neural Inf. Process. Syst., vol. 33, Dec. 2020.
7. A. Kunar et al., "DTGAN: Differential private training for tabular GANs," arXiv preprint arXiv:2107.02521, 2021.
8. Y. Li et al., "StoryGAN: A sequential conditional GAN for story visualization," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2019, pp. 6329–6338.
9. L. Mescheder, S. Nowozin, and A. Geiger, "Which training methods for GANs do actually converge?," in Proc. Int. Conf. Mach. Learn., 2018.
10. S. Nasr Esfahani and S. Latifi, "Image generation with GANs-based techniques: A survey," Int. J. Comput. Sci. Inf. Technol., vol. 11, pp. 33–50, 2019.
11. T. Ogawa et al., "Object detection for comics using Manga109 annotations," arXiv preprint arXiv:1803.08670, 2018.
12. T. Qiao et al., "MirrorGAN: Learning text-to-image generation by redescription," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2019, pp. 1505–1514.
13 X. Qin et al., "A faster R-CNN based method for comic characters face detection," in Proc. 14th IAPR Int. Conf. Document Anal. Recognit., vol. 1, 2017, pp. 1074–1080.
14. A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," arXiv preprint arXiv:1511.06434, 2015.
15. S. Reed et al., "Generative adversarial text-to-image synthesis," in Proc. Int. Conf. Mach. Learn., vol. 48, 2016, pp. 1060–1069.
16. T. Salimans et al., "Improved techniques for training GANs," arXiv preprint arXiv:1606.03498, 2016.
17. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2015.

18. C. Szegedy et al., “Rethinking the inception architecture for computer vision,” arXiv preprint arXiv:1512.00567, 2015.

19. R. Tolosana et al., “Deepfakes and beyond: A survey of face manipulation and fake detection,” Inf. Fusion, vol. 64, pp. 131–148, 2020.

20. P. Welinder et al., “Caltech-UCSD Birds 200,” Tech. Rep. CNS-TR-2010-001, California Institute of Technology, 2010.

21. T. Xu et al., “AttnGAN: Fine-grained text-to-image generation with attentional generative adversarial networks,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2018, pp. 1316–1324.

22. H. Yanagisawa, T. Yamashita, and H. Watanabe, “A study on object detection method from manga images using CNN,” in Proc. Int. Workshop Adv. Image Technol., 2018.

23. H. Zhang et al., “StackGAN: Text-to-photo realistic image synthesis with stacked generative adversarial networks,” in Proc. IEEE Int. Conf. Comput. Vis., 2017, pp. 5907–5915.

24. Z. Zhao et al., “FedTGAN: Federated learning framework for synthesizing tabular data,” arXiv preprint arXiv:2108.07927, 2021.

25. Z. Zhao et al., “CTAB-GAN: Effective table data synthesizing,” arXiv preprint arXiv:2102.08369, 2021.

Integrated Machine Learning for Cricket: Predicting Match Winners

Mohammad Ahmed M. Y. Ansari
Computer Science and Engineering (Data Science)
VCET
Mumbai, Maharashtra
✉ ahmed.22396101@vcet.edu.in

Chinmay Anant Patil
Computer Science and Engineering (Data Science)
VCET
Mumbai, Maharashtra
✉ chinmay.224426105@vcet.edu.in

Aditya Manoj Pal
Computer Science and Engineering (Data Science)
VCET
Mumbai, Maharashtra
✉ aditya.224376101@vcet.edu.in

Prathamesh Liladhar Patil
Computer Science and Engineering (Data Science)
VCET
Mumbai, Maharashtra
✉ prathamesh.224446105@vcet.edu.in

ABSTRACT

Fantasy cricket has emerged as a popular pastime among cricket enthusiasts, offering an immersive experience that allows fans to engage with the sport on a deeper level. However, the process of selecting a winning fantasy cricket team presents significant challenges for users, including the need to analyze player statistics, match conditions, and navigate budget constraints and team requirements. To address these challenges the proposed system integrates optimization techniques to maximize team performance while satisfying the constraints imposed by fantasy cricket platforms. By striking a balance between performance optimization and constraint adherence, the system aims to provide users with optimal team compositions that increase their chances of success in fantasy cricket leagues. The effectiveness of the proposed system is evaluated through rigorous testing and validation procedures. Performance metrics are used to assess the predictive capabilities of the ML models and the overall performance of the system. By demonstrating the reliability and accuracy of the system, this research aims to instill confidence in users and encourage widespread adoption of the proposed approach. This research paper introduces a pioneering approach to revolutionize the fantasy cricket landscape through the integration of machine learning (ML) techniques for player selection.

***KEYWORDS** : Fantasy cricket, Machine learning (ML) Techniques, Player selection, Performance metrics, Reliability, Accuracy, Revolutionize.*

INTRODUCTION

Cricket is a team sport involving eleven players on each side, with a mix of players [24]]. Batsmen aim to score the most runs, bowlers strive to take wickets and restrict runs, while all rounders contribute in both aspects [24]. Accurately predicting a player's potential runs or wickets helps in selecting the best team for each match based on current form, past performance, and match conditions [24]. Cricket has become more popular, leading to new formats like T20 and T10 [11]. Dream11 is a popular app where you can create your dream team of 11 players [11]. It's all about skills, ideas, and a bit of luck [11]. Cricket's various formats have led to the rise of prediction tools [1]. It's hard to predict the winner until the very end due to factors like individual and team performance, and environmental conditions [1]. Planning a game strategy requires considering all these factors [1]. Artificial Intelligence, machine learning, deep learning, and data science have revolutionized various aspects of life [2]. By ML and predicting outcomes before matches, players and coaches can analyze areas for improvement [3]. It makes life easier and enhances performance analysis [3].

Fantasy sports have become a global phenomenon, captivating the hearts of sports enthusiasts all around the world. It doesn't matter if you're a die-hard football fan, a basketball aficionado, or a cricket fanatic – fantasy sports offer an exciting way. As the season progresses, you earn points based on the performance of your chosen players in actual games. It's a virtual battle of wits and strategy, where your managerial skills and knowledge of the game are put to the test. One of the key elements of success in fantasy sports is team selection [3]. It's crucial to carefully analyze the strengths and weaknesses of each player, taking into consideration factors such as their current form, past performances, and upcoming fixtures. Building a well-rounded team that covers all positions and has depth on the bench is crucial to withstand injuries, suspensions, or unexpected dips in form [5] [8].

Player analysis is another vital aspect of fantasy sports. Keeping track of player statistics, injury updates, and team dynamics can give a competitive edge [5]. By studying player trends, one can identify rising stars or potential breakout performers who might be flying under the radar.

Analyzing the strengths and weaknesses of your opponents, studying their team compositions, and predicting their moves can help you devise effective strategies to outsmart them. It's not just about picking the most popular players; it's about finding the right balance between consistent performers and those who have the potential to deliver match-winning performances [11]. In addition to the strategic aspect, fantasy sports also foster a sense of community and friendly competition. Joining leagues with friends, family, or even strangers who share your passion for the game adds an extra layer of excitement.

The proposed cutting-edge platform combines advanced data analytics, real-time updates, and personalized recommendations to simplify the decision-making process and empower you to dominate your leagues like never before [4] [6]. This means users can make timely changes to their team, adapt to match conditions, and maximize chances of success. System incorporates advanced data analytics techniques to analyze player performance trends, historical data, and match conditions. This allows them to generate personalized team recommendations tailored to user preferences, budget, and league requirements. Fantasy cricket is a social experience, so the system tends to integrate community features into our platform [11].

PROBLEM STATEMENT AND OBJECTIVES

Problem Statement

In our context, the problem at hand revolves around the challenges users face in selecting appropriate players to win. Delayed updates on injuries, match statuses, and news can be frustrating. It's crucial to have timely info for fair gameplay. Analyzing player performance without advanced data analytics is tough. It's hard to assess their history, recent form, and track records against opponents. Live scoring and player point tracking are vital for an engaging experience. Users need to monitor their teams and make real-time adjustments. Personalized team recommendations based on preferences, budget, and strategy are often missing. It's harder to optimize team selection without them. Also, some platforms focus only on cricket, but expanding to other sports can attract a wider audience. Diversification makes it more inclusive and appealing.

Objectives

The findings from the implementation of the systems bear significant relevance to the research objectives and address the identified problems.

1. Advanced Data Analytics: Significance: Simplify the decision-making process for fantasy cricket leagues by providing advanced data analytics, real-time updates, and personalized recommendations.
2. Real-Time Updates: Significance: Keep users up-to-date with the latest information on player information, match statuses, and other relevant news.
3. Personalized Recommendations: Significance: Enable users to make timely changes to their teams and adapt to match conditions for better chances of success.
4. Timely Changes: Significance: Utilize advanced data analytics techniques to analyze player performance trends, historical data, and match conditions.
5. Player Performance Analysis: Significance: Generate personalized team recommendations tailored to user preferences, budget, and league requirements.
6. User-Friendly Interface: Significance: Offer a user-friendly interface and intuitive design for easy navigation and browsing of player profiles, statistics, and team performance.
7. Addressing Challenges: Significance: Address the challenges users face in selecting appropriate players by providing timely updates, advanced data analytics, and personalized recommendations.
8. Diversification: Significance: Consider diversifying the platform to include other sports, attracting a wider audience and making it more inclusive.

LITERATURE SURVEY

Literature Review

In cricket, prediction of player performance helps team management to decide final team selection [3]. Currently cricket has the highest viewership after soccer [3]. Winning is the ultimate aim of a Cricket game [3]. Batting and bowling are the crucial factors affecting a team [10]. To stay one step ahead of the competition and win over new fans or followers, cricket must innovate [12]. IPL is a famous cricket league in India. It's loved by fans for live streaming and TV broadcasts [14]. Predicting IPL match outcomes is important for online traders and sponsors. Many factors influence the winner [19]. Batting and bowling are crucial factors for a team's success in IPL matches. Other factors like fielding, captaincy skills, and home ground advantage also play a role [10] [12]. Evaluating players' statistics helps determine their performance and suitability for the team [10] [12]. In predictive analysis, various methods have been used for making predictions. Machine learning algorithms have been employed for this task since early times [2]. Different authors have suggested different techniques, including the exploration of neural networks in this field [2]. In cricket, the team management and captain carefully choose eleven players for each match based on various attributes and players' past performances [13]. Player selection is crucial as these eleven players remain fixed for the entire match, unless there is an injury [13]. By analyzing players' past records, their performance can be predicted [13]. Machine learning is important for predicting players' performance, which helps in building a strong team [3]. Researchers use machine learning to make these predictions [2]. In cricket, many factors impact player selection, such as overall stats, performances with different teams, and the last five performances [17]. These factors are considered when choosing the playing 11 from a squad of 16 [17]. Additionally, the relative strength of the competing teams is a key feature in predicting the winner [17]. In cricket, a team has 11 players including batsmen, bowlers, and all-rounders [18]. It's crucial to have a well-balanced and diverse team for better chances of winning [18]. Factors like pitch conditions, winning the toss, and the batting or bowling order can affect the outcome too [18]. The performance of batsmen, bowlers, and fielding plays a major role in determining the match result [18]. Researchers use statistical and probabilistic methods to study these factors in more detail [18]. Machine learning is a branch of AI that tackles real-world engineering problems [1]. This allows for control, recognition, and resilience [1]. Predicting e-sports match results is tough [7]. ML models analyze player variables to calculate scores. We created a heuristic approach to predict wins [7]. We're in the data-driven era, where everything is connected and digitally recorded [9]. Machine learning (ML) is a popular technology that allows systems to learn and improve from experience without explicit programming [9]. ML offers various algorithms like classification, regression, clustering, feature engineering, association rule learning, and reinforcement learning [9]. Choosing the right algorithm for a specific application can be challenging, as their purposes and outcomes may vary based on the data characteristics [9]. Data Science and Analytics play a crucial role in gaining a deeper understanding of sports and aiding decision-making processes [16]. By adopting Data Science, we can utilize Sports Analytics to gain insights into the sports market, evaluate player performances, rank teams, predict player and team scores, make effective game day decisions, enhance team branding and income, and optimize budget and financial management [16]. It's an exciting and powerful tool for the world of sports [16].

Table I: Author, Technique, Merits and Limitations of Research Papers

Author	Technique	Merits	Limitation
Abraham Garcia-Aliaga et al [20]	Used dimensionality reduction and the RIPPER algorithm for prediction performance.	Players' positions are determined by their technical and tactical behavior.	When evaluating players, their styles are not considered.
Sait Can Yucebas [21]	The Rectifier Linear Unit (ReLU) is used for performance prediction.	The function of Gedeon is used to rank players, which helps improve the overall accuracy of performance prediction.	The players' performance was assessed solely based on their position on the field, without taking other factors into account.
Luca Pappalardo et al [22]	PlayRank algorithm is employed to identify, assess, and arrange the players.	PlayRank algorithm works well with different datasets and helps analyze individual performance efficiently.	Implementation is difficult.

Wei Gu et al [23]	Along with the nonparametric test, different ML algorithms were also used to analyze the dataset.	When compared to other algorithms, the Support Vector Machine (SVM) stands out with its accuracy rate of over 90%.	Model is more suitable for analyzing data from a single season rather than multiple seasons.
Kalpdrum Passi et al [24]	The random forest, naïve bayes, multiclass SVM, and decision tree algorithms are compared.	Random forest stands out when it comes to accuracy among other ML algorithms.	Used on small datasets only.
Mat Herold et al. [25]	Machine learning algorithms used, considered time, data size, and location for accurate predictions.	Adaptability, possession, communi-cation, and tempo are crucial for team evaluation.	Applicable to simple datasets only.

Limitations of Existing System

In the current approach, most of them have relied on basic machine learning algorithms such as Naïve Bayes and logistic regression [1]. They have considered factors like teams, toss winners, winners by runs, and winners by wickets [1]. Injuries can have a big impact on team performance and finances [5]. Some studies have used machine learning to predict football injuries, showing its potential [5]. Machine learning has been applied in various fields like healthcare and finance for tasks like cancer detection and stock market prediction [5]. However, its use in sports is still relatively new [5]. To prevent future injuries, it's crucial to understand the factors that contribute to injury risk and how they interact [8]. The data used for the machine learning model should be of high quality and can include any information that is deemed useful for predicting injuries [8]. In fantasy cricket, contestants have a fixed budget to select their team [10]. They can't choose all the top players, so they need to make smart picks based on players' past and current performance [10]. AI and machine learning can be used to model this process [10]. However, achieving the best results every day is challenging due to the unpredictable nature of the game's outcome [10]. ML models are now widely used in different sectors like healthcare, agriculture, weather forecasting, transportation, and sports [11]. Sports play a crucial role in people's lives and involve significant investments [11]. Analyzing player performance is essential in the sports industry [11]. In the past, this analysis was done manually, which took a lot of time and sometimes led to illegal activities [11]. To address these issues, ML-based performance analysis is proposed [11]. ML algorithms are used to analyze players' existing performance and create a ranking list [11].

There has been a growing focus on the complexity of tactical decisions in sports and the scrutiny coaches face [15]. However, scientific investigations into these decisions have been limited [15]. But now, there are studies using sentiment analysis on Twitter to identify important events during games and even predict game outcomes [15]. It's an exciting development that sheds light on the impact of social media in sports [15]. Real-time insights are incredibly important in today's fast-paced world [4] [6]. Traditional methods of processing data in batches and analyzing it retrospectively are no longer enough [4] [6]. With the massive amount of data being generated at unprecedented speeds, organizations need to be able to integrate and analyze data in real-time [4] [6]. This allows them to gain immediate insights into customer behavior, operational performance, market trends, and emerging opportunities [4] [6]. In today's dynamic business landscape, real-time insights are crucial for staying ahead of the curve [4] [6].

Stakeholders find older machine learning techniques like Naïve Bayes and Logistic Regression outdated [1]. However, newer machine learning algorithms, thanks to parallel processing and increased computational capacities, tend to outperform the older ones [1]. Random forest classifiers and adaboost are preferred because they can handle a large number of input features, train decision trees and stumps, and provide highly accurate predictions [1]. Evaluating players' performances is a complex task [3]. Existing studies often use limited datasets and focus on specific aspects of cricket, like batsmen's performance or a short timeframe [3]. None of them have considered a holistic dataset that includes factors like weather, venue, wickets, and cricket-related statistics [3].

PROPOSED SYSTEM

Introduction

1. Generate personalized team recommendations to users' preferences, budget constraints, and strategic objectives based on systems analysis..

2. Through data analytics techniques, we meticulously examine players' historical performances, recent form, and track records against specific opponents to identify top-performing candidates for team selection.
3. Predict the winning probability of a team.
4. Predict the runs scored by a particular team.
5. Generate personalized team recommendations to users' preferences, budget constraints, and strategic objectives based on systems analysis.
6. Offer extensive customization options for users to fine-tune recommendations according to individual strategies and preferences.
7. Expand support to include other sports beyond cricket, catering to a broader audience of fantasy sports enthusiasts.

System Architecture

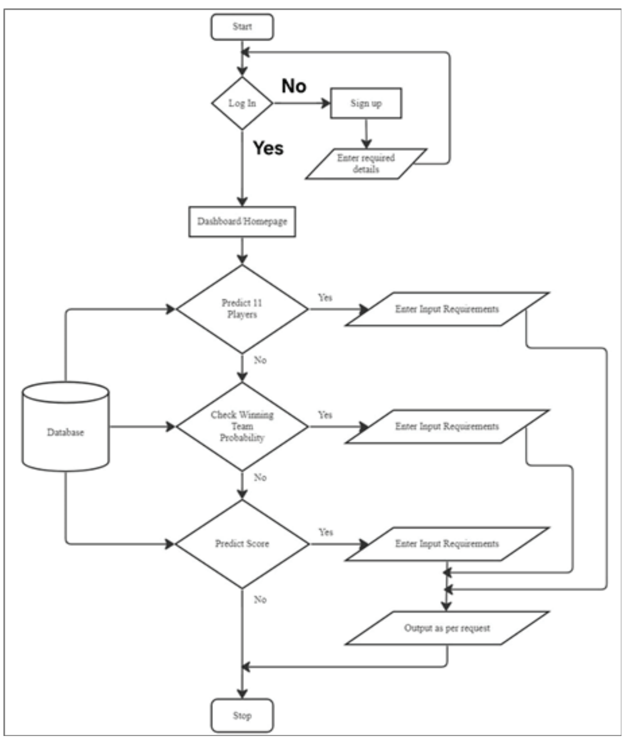


Fig. 1: Proposed System Architecture

The user starts by logging in. If they don't have an account, they can create one by signing up and entering required details. Once the signup process is successful, the user is directed back to the login page. After logging in, the user enters the home page, where they have options to choose from, such as selecting players, predicting scores, and checking probabilities. For example, if the user wants to predict the winning team probability, they would provide details like the batting team, bowling team, opposing team, and so on. Based on this information, a suitable response is generated from the database and presented to the user as output. The user can then proceed to the next option after completing one task. It's a user-friendly and interactive system.

Implementation, Results and Validity

Implementation process: The data for the proposed system was collected from Kaggle, a platform that offers a wide range of datasets for various projects. We used a machine learning algorithm to implement the model for predicting project outcomes. We took two different approaches for our project. In the first approach, we combined four different algorithms: logistic regression, random forest, decision tree, and SVM. We then calculated the combined accuracy of all four models. In the second approach, we focused specifically on using the random forest classifier to predict the names of 11 players.

Results: Implemented results are:

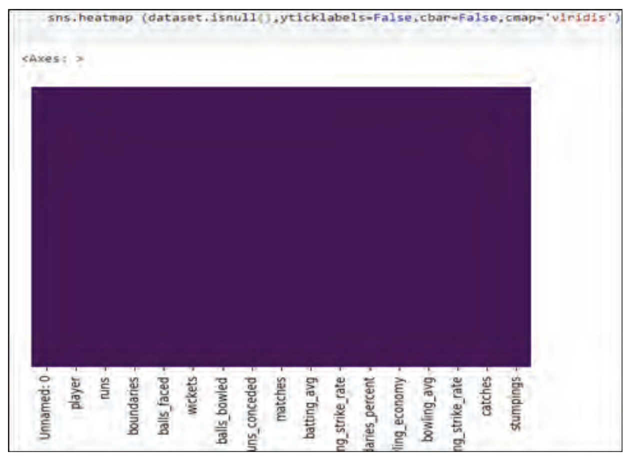
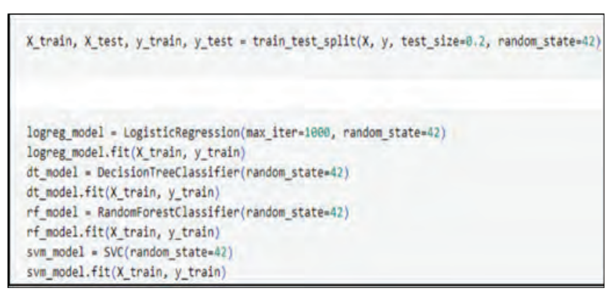


Fig. 2: Done pre-processing for handling missing values and checked using heatmap



```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

logreg_model = LogisticRegression(max_iter=1000, random_state=42)
logreg_model.fit(X_train, y_train)
dt_model = DecisionTreeClassifier(random_state=42)
dt_model.fit(X_train, y_train)
rf_model = RandomForestClassifier(random_state=42)
rf_model.fit(X_train, y_train)
svm_model = SVC(random_state=42)
svm_model.fit(X_train, y_train)
```

Fig. 3 Importing and training models

```
# Evaluate model performance
accuracy = accuracy_score(y_test, y_pred)
print(f'Accuracy: {accuracy:.4f}')

Accuracy: 0.9917
```

Fig. 4 Accuracy of Logistic regression

```
threshold = 0.5
combined_pred_binary = (combined_pred >= threshold).astype(int)

# Use the combined binary predictions for accuracy calculation
accuracy = accuracy_score(y_test, combined_pred_binary)
print(f'Combined Model Accuracy: {accuracy:.4f}')

Combined Model Accuracy: 0.9752
```

Fig. 5 Combined model accuracy

```
Classification Report:
              precision    recall  f1-score   support

           0       0.98      0.97      0.98        62
           1       0.97      0.98      0.97        59

    accuracy                           0.98       121
   macro avg       0.98      0.98      0.98       121
weighted avg       0.98      0.98      0.98       121
```

Fig. 6 Showing Classification Report

```
print(f'Precision: {precision:.4f}')
print(f'Recall: {recall:.4f}')
print(f'F1-score: {f1:.4f}')

Precision: 0.9667
Recall: 0.9831
F1-score: 0.9748

from sklearn.metrics import confusion_matrix, roc_curve, precision_recall_curve

conf_matrix = confusion_matrix(y_test, combined_pred_binary)
print('Confusion Matrix:')
print(conf_matrix)

Confusion Matrix:
[[60  2]
 [ 1 58]]
```

Fig. 7 Performance and confusion matrix

```
Accuracy: 0.9917
Precision: 0.9833
Recall: 1.0000
F1-score: 0.9916
Cross-Validation Scores: [0.96       1.         0.91666667 1.         1.        ]
Average CV Accuracy: 0.9753
```

Fig. 8 Predict using Random Forest Classifier

```
0        A Ashish Reddy
319          MEK Hussey
331              MM Ali
330           ML Hayden
329           MK Tiwary
328           MK Pandey
327           MK Lomror
324             MJ Lumb
323         MJ Guptill
321         MG Johnson
320        MF Maharoof
Name: player, dtype: object
```

Fig. 9 Names of 11 out of 604 players from the given dataset using Random Forest Classifier

Evaluation metrics: As seen in Fig. 8 the model has Accuracy: 0.9917, indicating the model is highly accurate in making correct predictions. Precision: 0.9833, suggesting that model has a high precision rate, i.e. it correctly identifies positive cases among all cases it predicts as positive. Recall: 1.0000 signifies that model has perfect recall, i.e. it correctly identifies all actual positive cases. F1-score: 0.9916 is a combination of precision and recall, providing a balance between the two metrics. A high F1-score indicates good overall performance. Cross-Validation Scores: The cross-validation scores ranging from 0.9167 to 1.0, with an average accuracy of 0.9753, indicate consistent and high performance across different subsets of the data during training and testing.

Discussion: In the first approach, we used a combination of four models, while in the second approach, we solely relied on the random forest model. Surprisingly, the random forest model demonstrated superior accuracy compared to the combined approach. This suggests that the random forest algorithm was more effective in capturing the underlying patterns and relationships within the dataset.

CONCLUSION AND FUTURE WORK

Conclusion

The. system provides a comprehensive platform that combines data analytics and makes informed decisions. By analyzing player performance, team compositions, and other factors, users can create winning fantasy teams. The integration of community features adds a social element to the experience. With its advanced features and capabilities, this fantasy cricket system empowers users to stay ahead of the game and make strategic decisions.

Future Work

Looking ahead, it could offer more in-depth player statistics, like batting averages and bowling figures Additionally, incorporating social features, such as chat rooms or forums, would allow fans to connect and discuss strategies. Lastly, a live scoring feature would provide real-time updates during matches, adding an extra level of excitement.

REFERENCES

1. A. Sahu, D. Kaushik, and A. Meena Priyadharsini, “Predictive Analysis of Cricket”, Turkish Journal of Computer and Mathematics Education, vol. 12, no. 6, pp. 5111-5124, Apr. 2021.
2. M. A. Ansari, S. Ghat, P. Patil, S. Gupta and S. Shingi, “VSAP: Technical Framework for College Academics,” 2024 11th International Conference on Computing for

Sustainable Global Development (INDIACom), New Delhi, India, 2024, pp. 178-184.

3. C. Kapadiya, A. Shah, K. Adhvaryu, and P. Barot, "Intelligent Cricket Team Selection by Predicting Individual Players' Performance using Efficient Machine Learning Technique", International Journal of Engineering and Advanced Technology (IJEAT), vol. 9, no. 3, Feb. 2020.
4. A. Ambasht, "Real-Time Data Integration and Analytics: Empowering Data-Driven Decision Making", International Journal of Computer Trends and Technology, vol. 71, no. 7, pp. 8-14, Jul. 2023.
5. A. Majumdar, R. Bakirov, D. Hodges, et al., "Machine Learning for Understanding and Predicting Injuries in Football", Sports Med - Open, vol. 8, no. 73, 2022.
6. K. Patel, Y. Sakaria, and C. Bhadane, "Real Time Data Processing Framework", International Journal of Data Mining & Knowledge Management Process (IJDKP), vol. 5, no. 5, pp. 49, Sep. 2015.
7. F. Bahrololloomi, F. Klonowski, S. Sauer, R. Horst, and R. Dörner, "E-Sports Player Performance Metrics for Predicting the Outcome of League of Legends Matches Considering Player Roles", SN Computer Science, vol. 4, no. 238, 2023.
8. H. Van Eetvelde, L. D. Mendonça, C. Ley, R. Seil, and T. Tischer, "Machine learning methods in sport injury prediction and prevention: a systematic review", Journal of Experimental Orthopaedics, vol. 8, no. 1, p. 27, Apr. 2021.
9. I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions", SN Computer Science, vol. 2, no. 160, 2021.
10. K. Karthik, G. S. Krishnan, S. Shetty, S. S. Bankapur, R. P. Kolkar, T. S. Ashwin, and M. K. Vanahalli, "Analysis and Prediction of Fantasy Cricket Contest Winners Using Machine Learning Techniques", in Evolution in Computational Intelligence, Frontiers in Intelligent Computing: Theory and Applications (FICTA), vol. 1, 2020.
11. M. Sumathi, Prabu Selvam, Rajkamal Murugesan, "Cricket Players Performance Prediction and Evaluation Using Machine Learning Algorithms", in Proceedings of the 2023 International Conference on Networking and Communications (ICNWC), 2023.
12. B. Sakthivel, A. Abinaya, M. Juraidha Rumman, T. Madhuvarshini, and B. V. Tamilsudar, "Cricket Team Prediction Using Machine Learning", International Journal of Current Engineering and Scientific Research (IJCESR), vol. 10, no. 4, 2023.
13. N. M. Patil, B. H. Sequeira, N. N. Gonsalves, and A. A. Singh, "Cricket Team Prediction Using Machine Learning Techniques", SSRN, 2020.
14. Preetham HK, Prajwal R, Prince Kumar, and Naveen Kumar, "Cricket Score Prediction Using Machine Learning", International Journal Of Innovative Research In Technology IJIRT, vol. 9, no. 8, 2023.
15. R. Rein and D. Memmert, "Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science", SpringerPlus, vol. 5, p. 1410, 2016.
16. S. Kumar, S. Prithvi, H. V. C. Nandini, "Data Science Approach to Predict the Winning Fantasy Cricket Team Dream 11 Fantasy Sports", Research Gate, 2022.
17. S. Agarwal, L. Yadav, and S. Mehta, "Cricket Team Prediction with Hadoop: Statistical Modeling Approach", Procedia Computer Science, vol. 122, pp. 525-532, 2017.
18. S. Shah, P. J. Hazarika, and J. Hazarika, "A Study on Performance of Cricket Players using Factor Analysis Approach", International Journal of Advanced Research in Computer Science, vol. 8, no. 3, 2017.
19. S. K. C., A. Khetan, B. Kumar, D. Tolani, and H. Patel, "Prediction of IPL Match Outcome Using Machine Learning Techniques", in Atlantis Highlights in Computer Sciences, vol. 4, Proceedings of the 3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC), 2021.
20. A. García-Aliaga, M. Marquina, J. Coteron, A. Rodríguez-González, and S. Luengo-Sánchez, "In-game behaviour analysis of football players using machine learning techniques based on player statistics", International Journal of Sports Science & Coaching, vol. 16, no. 1, pp. 148–157, 2021.
21. S. C. Yucebas, "A deep learning analysis for the effect of individual player performances on match results", Neural Computing and Applications, vol. 34, pp. 12967-12984, 2022.
22. L. Pappalardo, P. Cintia, E. Massucco, D. Pedreshi, and F. Giannotti, "PlayeRank: Data-driven performance Evaluation and player ranking in soccer via a machine learning approach", ACM Trans. Intell. Syst. Technol., vol. 10, no. 5, pp. 1-23, 2019.
23. W. Gu, K. Foster, J. Shang, and L. Wei, "A game-predicting expert system using big data and machine learning", Expert Systems with applications, vol. 130, pp. 293-305, 2019.

[24] K. Passi and N. Pandey, "Increased prediction accuracy in the game of cricket using machine learning", International Journal of Data Mining & Knowledge management process (IJDKP), vol. 8, no. 2, pp. 1-18, 2018.

[25] M. Herold, F. Goes, S. Nopp, P. Bauer, C. Thompson, and T. Meyer, "Machine learning in men's professional football: current applications and future directions for improving attacking play", International Journal of Sports Science & Coaching, pp. 1-20, 2019.

Comparison of ACO and ABC Optimization for Optimal Power Flow

Kshipra Pandey
Department of Electrical Engineering
Swarrnim Startup and Innovation University
Gandhinagar, Gujarat
✉ kshiprarpandey@gmail.com

Chandrakant Rathore
Department of Electrical Engineering
S. B. Jain Institute of Technology
Nagpur, Maharashtra
✉ chandrakantrathore@sbjit.edu.in

ABSTRACT

Demand for Electric Power has lofted owing to increasing consumers in the market. As there is growth of Intelligent Grids, customers can also actively participate in Energy Management Programs. These are special programs, which are also coined as Demand Response Programs that can curtail their electricity bills and can also give them dividends or perks. The prime orient of DRPs is to downsize the energy consumption during peak hours and amend the load factor. Additionally Non-Conventional Energy Sources do act as an aide source besides the grid. This further complexities the system. Specific challenges are there in pursuance of DRPs which make Demand Side Management still a fantasy. This work started with operation of optimization techniques like Whale Optimization Algorithm, Particle Swarm Optimization and Ant Colony Optimization on minor bus systems like IEEE 5 Bus System and IEEE 9 Bus System. Further correspondence of the three optimization techniques is done. Work with implementation of Gbest Artificial Bee Colony system on IEEE 57 bus was done previously. Now it's for 118 buses.

KEYWORDS *: GABC, DRP, Demand response program, Renewable energy systems.*

INTRODUCTION

One of the key elements of a smart grid that gives customers control over their electricity usage patterns is demand side management, or DSM. It's characterized as a change in consumption habits of users to make the necessary adjustments to the power systems' load curves. In order to lower the peak demand without having to install more generating capacity or reinforce the transmission and distribution line, it focuses on using power-saving techniques, dynamic electric tariffs, and incentive-based distributed generation programmes. DSM aims to minimize peak load, maximize economic gain, minimize electricity from the utility, and make use of RESs. Given the steady rise in power demand, it may eventually become necessary to install new generation units to meet demand. The DSM lowers peak load to address this problem. We can therefore postpone the new installations for a few years. But customers should behave sensibly by converting their on-peak to off-peak electricity usage in order to keep supply and demand in balance. Utility providers can lower peak loads and flatten consumer load profiles with the aid of the DSM. This reduces total operating expenses and increases the grid's sustainability. When consumption is at its highest, there is an abrupt increase in demand. As an alternative to taxing new installations and generation capacities to meet the demand, it resolves this problem by lowering power usage during peak hours . To address issues with power consumption, DSM approaches are used. The reduction of peak demand and electricity costs is the main goals of these strategies.

ANT COLONY OPTIMIZATION (ACO)

Ant Colony Optimization (ACO) is a metaheuristic optimization algorithm inspired by the foraging behavior of ants. It was introduced by Marco Dorigo in the early 1990s. The algorithm is used to solve various optimization and search problems, especially in combinatorial optimization domains. ACO is based on the behavior of real ants, who use pheromone trails to communicate with each other and find the shortest path between their nest and food sources. The algorithm simulates this behavior by creating artificial ants that traverse a solution space and deposit pheromone on the paths they explore. Over time, the paths with higher pheromone levels become more attractive to other ants, leading to the emergence of better solutions.

Explanation of how the Ant Colony Optimization algorithm works:1. Initialization: Initialize the pheromone levels on all possible paths and set up the problemparameters like the number of ants, iterations, etc.2. Ant Movement: For each ant, select a starting point and let it move through the solution space based on certain rules. The ant can choose the next step probabilistically,considering both the pheromone levels on the path and a heuristic value that represents the quality of the solution. 3. Solution Evaluation: Evaluate the solution constructed by each ant based on a predefined objective function. 4. Pheromone Update: Update the pheromone levels on the paths according to the quality of the solutions found by the ants. Paths that lead to better solutions get higher pheromone reinforcement. 5. Evaporation: Gradually decrease the pheromone levels on all paths to simulate the natural evaporation process. 6. Iteration: Repeat steps 2-5 for a certain number of iterations or until a stopping criterion are met. 7. Termination: Once the algorithm converges or reaches the stopping criterion, return the best solution found throughout the iterations. Ant Colony Optimization has been successfully applied to various optimization problems, including the Traveling Salesman Problem (TSP), Vehicle Routing Problem (VRP), Quadratic Assignment Problem (QAP), and more. It is particularly useful for solving problems where the search space is large and complex, and traditional optimization techniques might struggle. Keep in mind that ACO has various variants and parameter settings that can be adjusted based on the problem domain to achieve better performance. It's also worth noting that while ACO is effective for certain types of problems, it might not always be the best choice for every optimization problem.

ARTIFICIAL BEE COLONY OPTIMIZATION OR ABC OPTIMIZATION

Artificial Bee Colony optimization is a bio-inspired optimization algorithm that is based on honeybees foraging tendencies. Introduced by Dervis Karaboga in 2005, this algorithm simulates the foraging behavior of colonial honeybees to find peerless solutions of optimization problems. Here's a brief overview of how the ABC optimization algorithm works: 1. Initialization: The algorithm initializes a population of artificial bees, which represent implicit findings of optimization problem. 2. Employed Bees Phase: In this phase, bees which are employed to investigate the search space by conditioning their places based on the quality of the solutions they find. Each employed bee has a specific solution and evaluates its fitness. 3. Onlooker Bees Phase: Bees which are onlooker nominate solutions based on the information shared by employed bees. The probability of selecting a solution is proportional to its fitness. The better solutions have a higher chance of being chosen.4. Scout Bees Phase: Scout bees are responsible for discovering new solutions. If an employed bee exhausts its search around a solution without finding a better one, it becomes a scout bee. The scout bee then explores a new solution randomly. 5. Update: The best solution found by the colony is updated, and the process iterates till a stopping criterion is met for ex. Maximum number of iterations or a satisfactory solution. The algorithm is particularly suitable for continuous optimization problems. It's relatively simple, requires few parameters, and has shown effectiveness in finding near-optimal solutions. Parameters in ABC optimization include the quantity of the various categories of the bees such as onlooker bees, employed bees and scout bees, as well as the limit on the number of trials for each employed bee. Adjusting these parameters can impact the interpretation of the algorithm on specific problems. There are variations and extensions of the ABC algorithm tailored for specific problem domains and researchers continue to explore its applications and improvements.

CO-RELATION BETWEEN ACO AND ACB

Artificial Bee Colony (ABC) optimization and Ant Colony Optimization (ACO) are both metaheuristic optimization algorithms that are nature-inspired and draw inspiration from the collective behavior of the social insects. Despite being developed independently, there are some commonalities between ABC and ACO: 1. Nature Inspired Met heuristics: Both ABC and ACO are inspired by the foraging behavior of social insects. ABC takes inspiration from foraging behavior of honeybees, while ACO's inspiration is from foraging behavior of ants. In both cases, the algorithms mimic the way real insects communicate and cooperate to find optimal paths or solutions.2. Population-Based Approaches: Both algorithms operate with populations of solutions or paths. In ABC, the population consists of artificial bees representing potential solutions, while in ACO, it consists of artificial ants representing paths or solutions.3. Decentralized Communication: Both ABC and ACO involve decentralized communication among the agents (bees or ants). In ACO, ants communicate through pheromones to mark and follow paths, while in ABC, employed and onlooker bees share information about promising solutions through a mechanism similar to the waggle dance observed in the lifestyle of honeybees. 4. Exploration and Exploitation: Both algorithms balance exploration and exploitation. Both the techniques investigate search space for the discovery of new solutions and exploit promising regions by intensifying the search around them. This helps in searching of the optimal or at least optimal solutions to optimization problems. 5. Stigmergy: Stigmergy is a concept common to both ABC and ACO. It refers to indirect communication through modifications to the environment. In ACO, ants deposit pheromones on paths, while in ABC, employed and onlooker bees share information on solutions. The accumulated findings guide other individuals toward more promising regions in the search space. 6. Iterative Improvement: Both algorithms operate iteratively, with the solution quality improving over successive iterations. The search process continues till stopping criterion is achieved as for example the maximum number of iterations or a satisfactory solution. While there are these commonalities, it's also vital to keep a track that ABC and ACO have their unique characteristics and mechanisms tailored to their specific inspirations (bees and ants) and problem domains. Each algorithm has its strengths and weaknesses and may be more suitable for certain types of optimization problems.

Table 1: Comparison of ACO and ABC

ACO	ABC
Capable of synchronizing well with other optimization techniques	
Both algorithms belong to the group of algorithms inspired by behavior of swarms of insects and they are applied for minimizing the proper function	
Both ACO and ABC methods share the globally best solutions through the information exchange process between the ants and bees.	
Provides a solution roughly 43% faster with a limit on number of iterations	Provides roughly 30% quicker and 26% more consistent solution with a limit on maximum acceptable cost
Demerit-stagnation behavior of the ants	Demerit-time consuming global search for initial solutions by the employed bees
ACO algorithm has the disadvantage of slow convergence speed anda easy to fall into local optima, but it's robustness is strong.	ABC algorithm has the advantage of fast convergence speed and not easy to be classified into local optimal, however it has the disadvantages of early convergence and poor initial path planning effect.

The following are the main problems with the ACO algorithm:1. As the number of iterations increases, the algorithm's convergence speed rapidly changes. 2. The range of values for the constants α and β makes up the majority of the method. The ideal value of α, which is often 1.5, must be taken into account for better outcomes. Therefore, regardless of the issue, α should be taken care of for better outcomes visualization. 3. When α is negative, such as -1.5, the solution readily enters a stagnation phase in which no additional changes may be made.Due to its probabilistic nature, ACO relies heavily on its two constants, α and β. Maintaining the standard settings of this parameter is highly desired for any user. The ACO algorithm's stagnation phase places the solution in local optima from which it is impossible to recover without altering the values of α. Even if we maintain the conventional values of α and β, the convergence speed of the solution likewise changes quickly for the case with the same number of ants. Ignoring the ranges of these constants and other static parts of the method will not yield the optimal answer.

GABC (OR GBEST ABC)

In the context of Artificial Bee Colony (ABC) optimization, the term "G-best" usually refers to the best solution which is global found by the entire colony of artificial bees. This global best solution represents the most optimal solution found throughout the entire optimization process. In ABC optimization, the population of artificial bees explores the solution space, and each bee may find a candidate solution to the optimization problem. The quality of a solution is typically measured by its fitness, which is determined by the objective function of the optimization problem (i.e., the function that needs to be minimized or maximized).

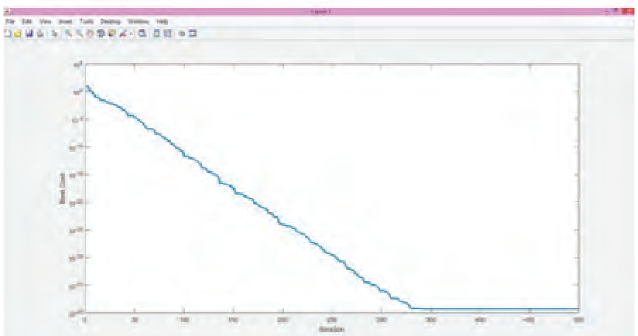
The global best solution, often denoted as G-best, is the solution with the best fitness among all the solutions discovered by the entire colony of artificial bees. This solution is crucial because it represents the overall best solution found during the optimization process.

The G-best solution is updated dynamically as the artificial bees continue to explore the solution space. If a new solution is found by an artificial bee that has a better fitness than the current G-best, the G-best is updated with the new solution

HOW G-BEST ABC IS BETTER AS COMPARED TO ABC

The search strategies between the GABC algorithm and the ABC/best algorithms are different. • The GABC algorithm searches for the new candidate solution toward the global best solution, but the ABC algorithms search for the new candidate solution around the global best solution. • The original search equations of ABC algorithms always look for better results near the global best solution. • There is still an insufficiency in ABC algorithm regarding its solution search equation, which is good at exploration but poor at exploitation. • Inspired by PSO, we propose an improved ABC algorithm called gbest-guided ABC (GABC) algorithm by incorporating the information of global best (gbest) solution into the solution search equation to improve the exploitation.

RESULTS



CONCLUSION

OPF can be implemented for higher bus system and can be enhanced plus optimized further by Gbest Artificial Bee Colony (GABC) Algorithm to create a strong, robust and more effective algorithm . Further extension to the work can be done in following ways: Including the Demand Side Management strategies to optimal load flow and Analysis of Deregulation Market.

REFERENCES

1. Ant Colony Optimization Algorithm for Electrical Power Systems Applications: A Literature Review,Ragab A. El-Sehiemy, Almoataz Y. Abdelaziz, Al-Bahrani, Applications of Nature-Inspired Computing in Renewable Energy Systems, 2022
2. Planning optimal power dispatch schedule using constrained ant colony optimization, Anand Kumar, Manoj thakur, Garima Mittal, Applied Soft Computing, 2022
3. Multi objective functions of constraint optimal power flow based on modified ant colony system optimization technique, Layth Tawfeeq Al-Bahran and Ali Qasim Abdulrasool, IOP Conference Series:Materials Science and Engineering, 2021.
4. Constraint optimal power flow based on ant colony optimization, Ali Qasim, Dr. Layth Tawfeeq Al- Bahrani, Journal of Engineering and Sustainable Development, 2020
5. Performance Comparison of Meta-Heuristic Algorithms PSO and ACO for Optimum Power Flow in Power Systems, Raheel Jawad, Ali A. Kadhum, Malak Moneer, International Conference on Communication & Information Technology, 2021
6. Comparative Analysis of Ant Colony and Particle Swarm Optimization Algorithms for Distance Optimization, Arushi Gupta, Smruti Shrivastava, Procedia Computer Science, 2020
7. Comparison of Particle Swarm and Whale Optimization Algorithms for Optimal Power Flow Solution, Ahmed S. Menesy, I. O. Habiballah , Hamdy M. Sultan, International Journal of Engineering Research & Technology (IJERT), 2022
8. Improved Whale Optimization Algorithm Based on Nonlinear Adaptive Weight and Golden Sine Operator, J. ZHANG1 AND J. S. WANG, IEEE Access, 2020
9. A cooperative PSO algorithm for Volt-VAR optimization in smart distribution grids, Myron Papadimitrakis, Aristotelis Kapnopoulas, Savvas Tsavartzidis, Alex Alexandridis, Electric Power Systems Research, 2022
10. Solving Reactive Power Optimzation Problem using weighted improved PSO Algorithm, Shaima Hamndan Shri, Md. B. Essa, Ayad Mijbas, 2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), 2021.
11. Planning optimal power dispatch schedule using constrained ant colony optimization, Anand Kumar, Manoj thakur, Garima Mittal, Applied Soft Computing, 2022.
12. Jagdish Chand Bansal , Susheel Kumar Joshi , Harish Sharma Modified global best artificial bee colony for constrained optimization problems, Computer and Electrical Engineering April 2018, Pages 365-382.
13. Marwa Hannachi , Omessaad Elbeji , Lassaad Sbita, Comparison between ABC and ACO: Tunning of On-Off MPPT for wind systems , Volume 45, Issue 6 , Sage Journals, April 28, 2021
14. Dipanshu Naware , Raviteja Badigenchala, Arghya Mitra,Debapriya Das b, Impact of demand response on battery energy storage degradation using gbest-guided artificial bee colony algorithm with forecasted solar insolation, Journal of Energy Storage, Volume 52, Part B, 15 August 2022, 10491.

15. Padmini Sahu, Bikesh Kumar Singh, Neelamshobha Nirala, An improved feature selection approach using global best guided Gaussian artificial bee colony for EMG classification Biomedical Signal Processing and Control, Volume 80, Part 2, February 2023, 104399

16. Chao Liu , Lei Wu , Wensheng Xiao , Mixing algorithm of ACO and ABC for solving path planning of mobile robot, Applied Soft Computing, Volume 148, November 2023, 110868

17. S. Shivani Reddy; B. Shravani; K. Sreeja; M. Mourya, A comparison study of MABC, ACO, ABC and PSO load balancing algorithms, S. Shivani Reddy; B. Shravani; K. Sreeja; M. Mourya, 4th Smart Cities Symposium (SCS 2021)

18. Hoang Anh Nguyen; Duc Minh Nguyen; Quoc Huy Pham; Quang Dai Pham; Duc Chinh Hoang , Energy and Time-Efficient Scheduling of Automated Guided Vehicles System: A Hybrid Artificial Bee Colony Algorithm and Improved Ant Colony Optimization Approach, 12th International Conference on Control, Automation and Information Sciences (ICCAIS) 2023.

Automated Invoice Data Extraction Using OCR, NLP, and Anchor-Based Determination for Enhanced Data Capture and Processing

Swaranjaly Sandeep Jagtap
Student
Department of Computer Engineering
Shree L.R. Tiwari College of Engineering
Mumbai, Maharashtra
✉ jagtap41swara@gmail.com

Vinayak D. Shinde
Associate Professor
Department of Computer Engineering
Shree L.R. Tiwari College of Engineering
Mumbai, Maharashtra
✉ vdshinde@gmail.com

ABSTRACT

In today's dynamic business environment, efficient processing and handling of invoice data are crucial for optimizing operations. This paper proposes an automated approach for data capture and extraction from invoice images using Optical Character Recognition (OCR) and Natural Language Processing (NLP) techniques. OCR is employed to convert invoice images into Tab-Separated Values (TSV), which are then transformed into JSON objects for systematic analysis. Relevant data fields within the invoice are identified using anchor determination with regular expressions (Regex). To validate and enrich extracted data against predefined entities such as organizations, places, and individuals, the Compromise.js NLP library is used. The processed data is further filtered into an entity set and stored in CSV format for integration with databases. The methodology was tested on two invoice layouts, yielding high accuracy rates: 99% at the character level, 98.5% at the word level, and 97% at the line level. However, limitations arise from the specificity of Regex patterns, which restrict their application to particular formats. Future work could explore more adaptable Regex patterns combined with advanced machine learning and NLP techniques to handle varied invoice formats and layouts effectively.

KEYWORDS *: Invoice processing, Optical character recognition (OCR), Natural language processing (NLP), Data extraction, Data transformation, Anchor-based data capture, Regex-based analysis.*

INTRODUCTION

An invoice is a document issued by a seller to a buyer, detailing the goods or services provided along with a request for payment. Beyond its role as a payment request, an invoice also serves as an essential record-keeping document, critical for business management, tax compliance, and accounting purposes. Invoices are indispensable for companies and organizations, as they serve as proof of purchase and provide key information required for financial operations. With the increasing volume of invoices generated daily in most organizations, efficiently managing them has become a significant challenge [3]. Manual processing of invoices is time-consuming, error-prone, and inefficient, highlighting the need for an automated system to capture and extract data accurately. The digitization of documents has given rise to advanced technologies, such as Optical Character Recognition (OCR). OCR enables the transformation of digital documents into machine-readable formats. Modern OCR systems leverage deep learning techniques, including Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Recurrent Neural Networks (RNNs), for character recognition and sequence prediction in textual data [8][10]. Natural Language Processing (NLP) further complements OCR by enabling machines to analyze and process human language data. By integrating OCR and NLP technologies, this research aims to automate the data capture and extraction process for invoices. The approach utilizes anchor determination, where predefined keywords or terms help locate and extract relevant data fields within the document. Regular expressions (Regex), a powerful tool for pattern matching and string manipulation, are employed to enhance the accuracy of data extraction. The extracted data is stored in structured formats, such as Excel sheets or databases, making it suitable for future analysis and integration into business processes. This automation not only minimizes errors but also improves operational efficiency. This paper is organized as follows: The introduction provides an overview of the technologies and concepts used in this research. Section II reviews related studies and literature. Section III describes the proposed system and algorithm, followed by the implementation and results in Section IV. Finally, Section V presents the conclusions and future scope.

RELATED WORK

Several studies have been conducted to address the challenges associated with automating invoice data extraction, each proposing different methodologies and tools. In [1], a system was presented to automate the extraction of data from invoices with varying layouts across multiple domains. The system employed OCR tools for digitizing documents and FRESCO, a language designed for modeling and extracting content using domain-specific rules. Automation exceeding 50% with an error rate of less than 1% was achieved for health insurance invoices, demonstrating high efficiency and adaptability. In [2], handwritten text recognition was addressed using deep learning techniques. Word classification was performed using Convolutional Neural Networks (CNNs), and character segmentation was handled with Long Short-Term Memory (LSTM) networks. High accuracy was achieved with the EMNIST dataset, indicating potential applications in historical document preservation and mass digitization.

In [3], invoices were scanned in batches to improve image clarity. A location information template matching algorithm was applied for data localization, while character recognition was conducted using CNNs. An accuracy of 99.8% was reported, and a manual correction mechanism was added to ensure accuracy in VAT invoices.

In [4], the Tesseract OCR engine was utilized for Gujarati character recognition. Observations were made across different font styles and sizes. While good performance was noted in general cases, challenges were identified in handling complex and ambiguous characters.

In [5], a system combining CNNs for feature extraction and RNNs for sequence-to-sequence tasks was developed. A dataset consisting of real-world invoices was used, and an accuracy of 95% was achieved in extracting relevant information. However, a decline in performance was observed for unconventional layouts and poor-quality images.

In [6], a web-based application was developed to assist retailers and wholesalers in managing business data. Pre-processing was conducted using OpenCV techniques, text data was extracted through OCR, and post-processing was performed with Regex to improve accuracy. In [7], a Python-based web application for invoice detection was created. Canny edge detection and OpenCV's findContours() function were employed, followed by a four-point perspective transformation. The processed images were then passed through OCR for data extraction. In [8], the use of Reinforcement Learning (RL) in combination with AI-OCR was proposed. The adaptability of RL to various invoice formats was emphasized, as strategies could be refined based on rewards or penalties. Scalability and efficiency for handling large datasets and diverse formats were highlighted. In [9], a system utilizing UiPath was implemented. Invoices were retrieved from emails and stored in a database after data scraping. A chatbot with speech recognition was integrated for enhanced user interaction. In [10], a dual approach for processing handwritten and printed receipts was proposed. Handwritten receipts were processed using template matching and CNNs, achieving an accuracy of 80.93%, while printed receipts were processed using YOLOv4-s, achieving an accuracy of 99.39%.

PROPOSED ALGORITHM

The proposed system has been designed to automate the process of capturing and extracting data from invoice images. The extracted information is stored in an Excel sheet to enable further integration with databases and business processes. The system utilizes Optical Character Recognition (OCR), Natural Language Processing (NLP), and anchor determination to achieve accurate data extraction. OCR is employed to convert printed or handwritten text into machine-readable data, while NLP is utilized for the classification and validation of extracted information. The system has been developed using JavaScript libraries with Node.js as the runtime environment. An Intel i3 processor (7th generation) with 8 GB RAM has been used for testing the system. The system framework, as illustrated in Fig. 1, comprises the following steps: a) Invoice image processing, b) OCRing, c) Data transformation, d) File cleaning, e) Identification of relevant data fields, f) Data validation, g) Saving extracted information, and h) Storage in an Excel sheet. The phases are detailed below:

a) Invoice Image (Input): The system takes an invoice image as input, which contains details like goods/services, quantities, prices, taxes, and the total amount due. The image is processed to extract the relevant data fields.

b) OCRing (runOcr()) : The Tesseract.js package has been used for OCR. The function runOcr() processes the invoice image and generates a TSV (Tab-Separated Values) file. This step includes noise reduction, binarization, deskewing, and segmentation of words and characters. The TSV file provides a structured representation of the text in tabular form.

c) Data Transformation (tsvtotext(), fs Module): The tsvtotext() function is responsible for handling the TSV file and converting the extracted tabular data into a readable text file. The transformation process ensures that the text retains its original structure while remaining accessible. The Node.js fs module is utilized for this conversion.

d) Data Transformation (tsvtojsonobject(), fs Module): The TSV file is further converted into JSON (JavaScript Object Notation) format using the tsvtojsonobject() function. During this process, essential details such as Invoice Number, Purchase Order Number, Billing Address, Shipping Address, and Total Amount are extracted. JSON provides a structured and machine-readable format for the extracted data.

e) Identifying Relevant Data Fields (Anchor Determination): Relevant data fields are identified using anchor determination, which involves string matching. Regular Expressions (Regex) are employed as a powerful tool for locating and matching specific patterns within the extracted text. This process ensures the precise extraction of key data points.

f) File Cleaning (Regex): The extracted text undergoes cleaning to remove unnecessary characters, standardize formats, and ensure data integrity. Regex is used to eliminate non-alphanumeric symbols, extra spaces, and other redundant elements. This step reduces errors and enhances the quality of the extracted data for further analysis.

g) Data Validation (NLP Processes): Once the data has been cleaned, it is validated using NLP techniques. The compromise.js NLP library is employed to verify specific categories such as individuals, locations, and organizations. Validation ensures that the extracted data is accurate and matches predefined standards. The validated data is consolidated into a single entity set.

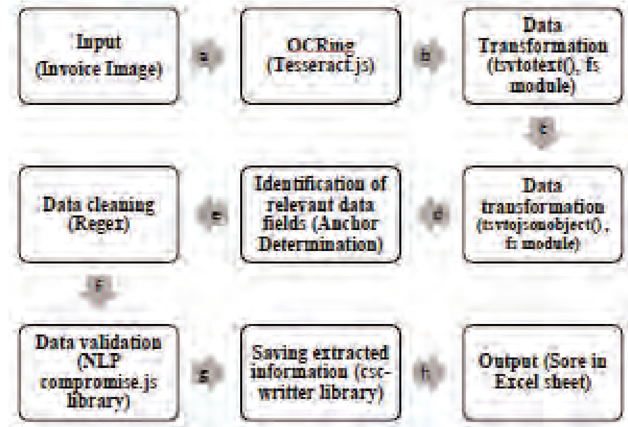


Fig. 1. Proposed System Methodology

h) Saving Extracted Information (Insertion in CSV File): The processed and validated data is saved in a CSV (Comma-Separated Values) file. This structured format ensures that the data remains easily accessible for analysis, storage, and integration with other systems.

i) Output (Save to an Excel Sheet): The validated and structured data is stored in an Excel sheet, facilitating integration with business operations. It serves as a repository for analysis, financial reporting, and auditing, ensuring accessibility for operational and compliance purposes.

Table. 1 Algorithm of proposed system

Input	:	An invoice image
Output	:	Store the invoice data in an Excel sheet
Step 1	:	Convert the image into a TSV file
Step 2	:	Convert the TSV file format into text file format
Step 3	:	Convert the text file obtained in previous step to a JSON object.
Step 4	:	Identify relevant data fields using Anchor determination.
Step 5	:	Perform data cleaning using Regex.
Step 6	:	Perform data validation using NLP process.
Step 7	:	Enter this entity set data into a CSV file.
Step 8	:	Store this data into an Excel sheet for further usage.

IMPLEMENTATION AND RESULTS

In the proposed system, invoice data is extracted and stored in an Excel sheet for further processing (Fig. 5).

The invoice image (Fig. 2) is first converted into a TSV (Tab-separated values) file using OCR, facilitated by the tesseract.js package.


East Repair Inc. INVOICE
1912 Harvest Lane
New York, NY 12210

Bill To	Ship To		
John Smith	John Smith	Invoice #	US-001
2 Court Square	3787 Pineview Drive	Invoice Date	11/02/2019
New York, NY 12210	Cambridge, MA 12210	P.O.#	2312/2019
		Due Date	26/02/2019

QTY	DESCRIPTION	UNIT PRICE	AMOUNT
1	Front and rear brake cables	100.00	100.00
2	New set of pedal arms	15.00	30.00
3	Labor 3hrs	5.00	15.00
		Subtotal	145.00
		Sales Tax 6.25%	9.06
		TOTAL	$154.06

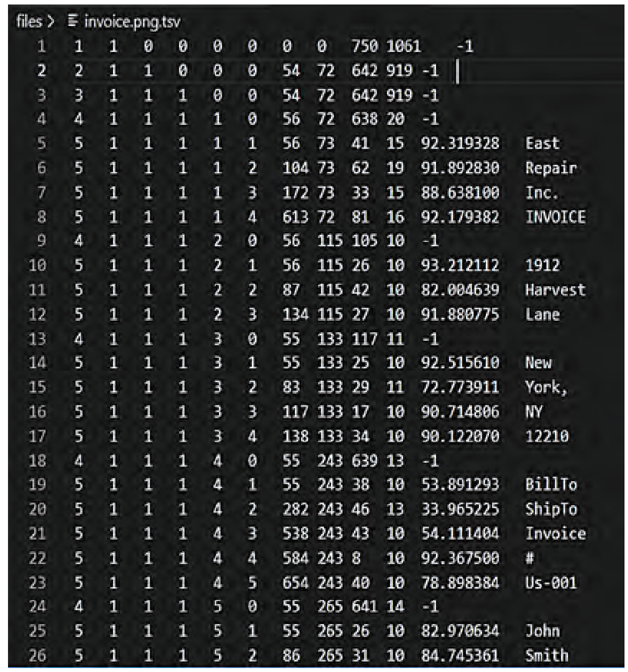
John Smith

Terms & Conditions
Payment is due within 15 days
Please make checks payable to: East Repair Inc.

Fig. 2 Invoice image(Input)

The runOcr() method processes the image, generating a TSV file (Fig. 3) that contains text positioning and hierarchical layout information. A confidence score, ranging from 0 to 100, is calculated to reflect the accuracy of the recognized text, such as a 97% confidence in "East Repair Inc." Preprocessing steps, including noise reduction, binarization, and deskewing, help ensure accurate text recognition. The TSV file is then transformed into a more readable text file using the tsvtotext() function. This step is handled by the fs library. Next, the tsvtojsonobject() function extracts essential data fields, such as Invoice No, PO No, BillTo and ShipTo addresses, and total amount, using anchor-based determination through Regex pattern matching (fig.4). Positional logic identifies additional data fields, while Regex cleaning removes unwanted characters and standardizes formats.

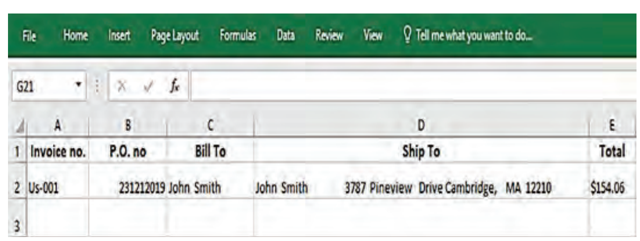
Data validation is performed using the NLP library compromise.js to ensure the accuracy of extracted entities like individuals, locations, and organizations. The cleaned data is consolidated into a set and stored in CSV format using the csv-writer library for structured storage and integration with other systems. The Regex pattern was tested on two invoice layouts, achieving the following accuracy: a) Character-level accuracy: 99%, b) Word-level accuracy: 98.5%, c) Line-level accuracy: 97%. These accuracy rates may vary with a larger dataset, as the Regex pattern is tailored to specific layouts. Variations in invoice structures and formats require distinct Regex patterns for different layouts.


files > invoice.png.tsv

1	1	1	0	0	0	0	0	0	750	1061	-1		
2	2	1	1	0	0	0	54	72	642	919	-1		
3	3	1	1	1	0	0	54	72	642	919	-1		
4	4	1	1	1	1	0	56	72	638	20	-1		
5	5	1	1	1	1	1	56	73	41	15	92.319328	East	
6	5	1	1	1	1	2	104	73	62	19	91.892830	Repair	
7	5	1	1	1	1	3	172	73	33	15	88.638100	Inc.	
8	5	1	1	1	1	4	613	72	81	16	92.179382	INVOICE	
9	4	1	1	1	2	0	56	115	105	10	-1		
10	5	1	1	1	2	1	56	115	26	10	93.212112	1912	
11	5	1	1	1	2	2	87	115	42	10	82.004639	Harvest	
12	5	1	1	1	2	3	134	115	27	10	91.880775	Lane	
13	4	1	1	1	3	0	55	133	117	11	-1		
14	5	1	1	1	3	1	55	133	25	10	92.515610	New	
15	5	1	1	1	3	2	83	133	29	11	72.773911	York,	
16	5	1	1	1	3	3	117	133	17	10	90.714806	NY	
17	5	1	1	1	3	4	138	133	34	10	90.122070	12210	
18	4	1	1	1	4	0	55	243	639	13	-1		
19	5	1	1	1	4	1	55	243	38	10	53.891293	BillTo	
20	5	1	1	1	4	2	282	243	46	13	33.965225	ShipTo	
21	5	1	1	1	4	3	538	243	43	10	54.111404	Invoice	
22	5	1	1	1	4	4	584	243	8	10	92.367500	#	
23	5	1	1	1	4	5	654	243	40	10	78.898384	Us-001	
24	4	1	1	1	5	0	55	265	641	14	-1		
25	5	1	1	1	5	1	55	265	26	10	82.970634	John	
26	5	1	1	1	5	2	86	265	31	10	84.745361	Smith	

Fig. 3 TSV (Tab-seperated values) file generated

Fig. 4 text file generated



	A	B	C	D	E
1	Invoice no.	P.O. no	Bill To	Ship To	Total
2	Us-001	231212019 John Smith	John Smith	3787 Pineview Drive Cambridge, MA 12210	$154.06
3					

Fig. 5 Output :Invoice data is stored in Excel sheet

CONCLUSION AND FUTURE SCOPE

The proposed system captures and extracts invoice data using OCR, NLP, and anchor determination techniques. The extracted data is stored in an Excel sheet for analysis and record-keeping purposes. The need for manual data entry of invoice information, which would otherwise be a time-consuming process, is eliminated. Optical Character Recognition, Natural Language Processing, and Anchor Determination are identified as optimal approaches for overall invoice processing. The string matcher was applied to two invoice layouts, achieving the following accuracy on various metrics: a) Character-level accuracy: 99%, b) Word-level accuracy: 98.5%, c) Line-level accuracy: 97%. In the future, the proposed system could be improved by implementing additional regex patterns, which would enhance its flexibility to work with a wider range of invoice layouts. Furthermore, more advanced machine learning algorithms and natural language processing techniques could be applied to further optimize the system's performance.

REFERENCES

1. T. Bayer and H. Mogg-Schneider, "A Generic System for Processing Invoices," in Proceedings of the Fourth International Conference on Document Analysis and Recognition, Ulm, Germany, 1997.
2. S.Reddy and M.Rao, "A Survey on Deep Learning For Hand Written Recognition by Digit & Character," International Journal of Creative Research Thoughts(IJCRT), vol. 8, no. 6, pp. 3310-3314, 2020.
3. S.Shi, C.Cui and Y.Xiao, "An Invoice Recognition System Using Deep Learning," in 2020 International Conference on Intelligent Computing, Automation and Systems (ICICAS), Chongqing, China, 2020.
4. M.Audichya and J.Saini, "A Study to Recognize Printed Gujarati Characters Using Tesseract OCR," International Journal for Research in Applied Science & Engineering Technology (IJRASET), vol. 5, no. IX, pp. 1505-1510, 2017.
5. T. Aung and S. P. a. H. H. K. Tin, "Automated Invoice Processing Using Image Recognition in Business Information Systems," Economics, Commerce and Trade Management: An International Journal (ECTIJ), vol. 3, pp. 73-79.
6. M. S. Satav, T. Varade, D. Kothavale, S. Thombare and P. Lokhande, "Data Extraction From Invoices Using Computer Vision," in 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), 2020.
7. H. Sidhwa, S. Kulshrestha and S. Malhotra, "Text Extraction from Bills and Invoices," in 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, 2018.
8. A. Malladhi, "Adaptive Reinforcement Learning in AI-OCR Invoice Processing for the Digital Age," International Journal of Computer Techniques , vol. 8, no. 6, pp. 1-22, 2021.
9. P. Doshi, Y. Kotak and A. Sahitya, "Automated Invoice Processing System Along with Chatbot," International Journal of Research in Engineering, Science and Management, vol. 3, no. 5, pp. 29-31, 2020.
10. C.-J. Lin and Y.-C. L. a. C.-L. Lee, "Automatic Receipt Recognition System Based on Artificial Intelligence Technology," Applied Sciences , vol. 12, no. 2, p. 853, 2022.
11. H.Ha and A.Horák, "Information Extraction from Scanned Invoice Images usingText Analysis and Layout Features," Signal Processing: Image Communication, vol. 102, 2022.

OCR Tools: A Demonstrative Study for Information Extraction

Sweta R.Sharma
Department of Computer Engineering
Ramrao Adik Institute of Technology
DY Patil Deemed to be University
Navi Mumbai, Maharashtra
✉ swetalj7011@gmail.com

Tushar H. Ghorpade
Department of Computer Engineering
Ramrao Adik Institute of Technology
DY Patil Deemed to be University
Navi Mumbai, Maharashtra
✉ tushar.ghorpade@rait.ac.in

Bharti A. Joshi
Department of Computer Engineering
Ramrao Adik Institute of Technology
DY Patil Deemed to be University
Navi Mumbai, Maharashtra
✉ bharti.joshi@rait.ac.in

ABSTRACT

A large number of organizations still use paper- based documentation. It becomes difficult and time-consuming to extract the necessary data from them. In addition to negatively affecting cost and time efficiency, this excessive paper consumption also hurts the environment due to the deforestation required to create these papers. To cut down on paper use, Usually, these paper-based documents are scanned and saved as images during their transition. This creates a requirement for technology that can identify and retrieve data from scanned images of paper documents. OCR is the process of turning handwritten or picture text into a format that computers can read. The process of electronically translating handwritten, typewritten, or printed text into automatically translated images is known as optical character recognition, or OCR. It is frequently used as a kind of data entry from original paper sources, such as bank or consultation documents, passport documents, bills, statements, receipts, cards, correspondence, or any other written records. Computer vision, artificial intelligence, and pattern recognition are all combined in the research of OCR. OCR is also dependent on the quality of the image. The accuracy of text recognition might be impacted by low-quality scans or photos. Handwriting, illegible or unclear letters, and various fonts all reduce OCR accuracy; also, each language has its own peculiarities and fonts, making it challenging to identify text written in multiple languages. This paper focuses on the study of various OCR tools and extraction of information from unstructured data.

KEYWORDS *: OCR, Image extraction, Natural language Processing, Text recognition, Information retrieval.*

INTRODUCTION

Optical character recognition is a technique that locates and recognizes text in scanned documents, pictures, or photos such as .jpg, .jpeg, .png etc. [2]. OCR software employs a technique which extracts data from PDFs, scanned documents, and other non editable forms by converting the data into machine-readable text data that may be changed and saved more effectively for further processing. In [1], the primary functions of an OCR system are feature extraction and feature discrimination/classification (based on patterns). As a branch of OCR, handwritten OCR is gaining popularity. Based on the input data, handwritten text is further divided into offline system [4] and online system. In [1], feature extraction andIdentify applicable funding agency here. If none, delete this.

feature discrimination/classification (based on patterns) are the main functions of an OCR system. Handwritten OCR is becoming more and more popular as a subfield of OCR. Handwritten text is further divided into offline system [4] and online system based on the input data. While the online system employs a more dynamic technique that relies on the movement of a pen tip with a specific velocity, projection angle, position, and locus point, the offline system in [1] uses scanned photographs as its static input. Because it eliminates the offline system's input data overlapping issue, an online system is therefore seen as more sophisticated and advanced. The objective is to study various OCR tools, extract text from images, and study their advantages and limitations. We need to examine and compare the performance of various OCR tools on the single page input text image and develop an end-to-end OCR model to retrieve the appropriate value using pytesseract and executable application setup using a self- assembled handwritten text image.To promote reproducibility, the implementation used in this study are publicly available at https://github.com/swetalj/Folder. The basic goal is to have a comprehensive evaluation of various OCR models on a single input handwritten text image dataset (a custom model is created).

This study evaluates the performance of various OCR tools, demonstrating their application on custom datasets and com- paring them based on predefined metrics. Key contributions include the creation of a custom dataset and the implementa- tion of a reproducible PyTesseract-based OCR model, with the code accessible in a linked repository. The paper is explored as follows, the survey of OCR tools is described in Section 2 with its sample of demonstrations, its origin, advantages, limitations, and applications. Section 3 elaborates on the comparative analysis of many OCR tools. Section 4 concludes with a study of various OCR tools.

OCR PROCESS WORKFLOW

The generalized workflow of OCR is depicted in Figure 1. It consists of the following phases:

1) Preprocessing: In [3],the foundational stage of character recognition, pre-processing plays a crucial role in achieving a high recognition rate. The main goals of pre-processing are to standardize strokes and eliminate irregularities that could hinder accurate recognition and reduce the recognition rate. These irregularities may include uneven text dimensions, missing points from pen movements, jitter, bends in handwriting, and inconsistent spacing between points. Common pre-processing tech- niques include interpolating missing points, smoothing, slant correction, size normalization, centering, and point resampling.

2) Feature Extraction: It is a crucial phase that entails extracting data in some way from the image to solve character recognition or other image recognition issues. The following are some popular feature extraction techniques [4]Use of methods like HOG, LBP and GF which are described below respectively:

    1) Histogram of Oriented Gradients (HOG): This application splits an image into pixels and determines the gradient's magnitude and orientation, which are then utilized by classification models.

    2) Local Binary Pattern: In this case, a binary pattern that is subsequently utilized for character recognition is created by comparing each pixel with its neighbor.

    3) The Gabor Filter is a bypass filter that can identify texture and edges in an input image.

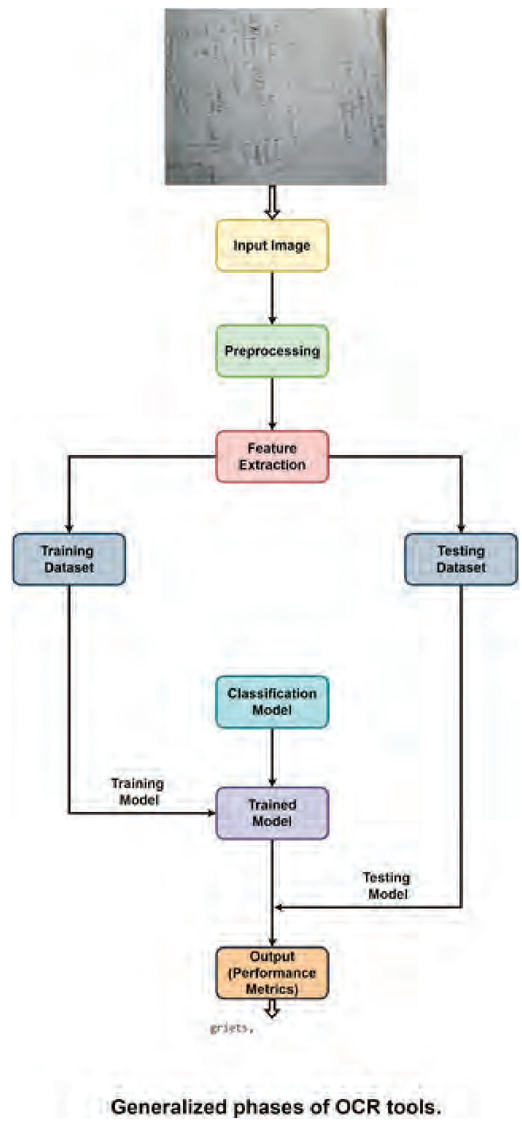


Fig. 1. OCR Lifecycle

3) Classification: The above-discussed machine learning techniques in previous steps are employed to determine the new data class based on historical datasets[4].

OCR processes use a variety of classification models; some straightforward models include: Employing machine learning models such as k-NN, SVM, or CNN which are described below respectively:

The k-Nearest Neighbors (k-NN) algorithm predicts the class of new data by comparing it to previously collected data. Out of the top k linked classes, the most prevalent one is selected as the new data class. Distance, typically Euclidean distance, serves as the basis for the correlation.

A hyperplane based on kernel function divides the classes in Support Vector Machines (SVM). When two classes are separated by a line, the subsequent income dataset class is indicated.

Convolutional neural networks (CNNs) are deep learning algorithms that use numerous layers of convolution filters to directly learn high-level features from input images.

4) Post-processing: Post-processing is the process of using linguistic expertise to correct results that were incorrectly classified [3]. The processing of the shape recognition result is called post-processing. Pure shape recognition accuracy can be improved by using language information.

SURVEY OF OCR TOOLS

This section provides a detailed analysis of various OCR tools.

Nanonet

Nanonets is an AI-based OCR application that serves as a cognitive capture automation tool for intelligent document processing. It is typically used for processing ID cards, receipts, invoices, and other documentation, extracting pertinent information from unstructured data through cutting- edge OCR, machine learning, and deep learning algorithms [8]. The platform utilizes advanced machine learning, image processing, and deep learning techniques to extract relevant facts from unstructured data. It provides several Zapier integrations, is quick, precise, and easy to use, and allows users to create bespoke OCR models from scratch. Despite its strong engine and ease of use, Nanonets requires users to create a processing model [14]. If a new or different invoice type is introduced, the current model may not work, making it less reliable and introducing longer workflows, which could be a challenge for service providers across multiple industries. Nanonets automates data collection for intelligent document processing, including ID cards, invoices, and receipts, using its AI-based OCR program. It offers several advantages, such as a modern user interface, the ability to handle large volumes of documents, affordability, ease of use, and cognitive data capture that minimizes the need for human intervention. It is particularly effective for automating document data extraction workflows using OCR and machine learning. Despite these strengths, one of the limitations of Nanonets is its table capture user interface, which is not user-friendly. The capabilities of Nanonets in recognizing words and sentences are illustrated in Figure 2, which displays text within square boxes, demonstrating its ability to process and extract data.

Nanonets is an AI-based OCR application that serves as a cognitive capture automation tool for intelligent document processing. It is typically used for processing ID cards, receipts, invoices, and other documentation, extracting pertinent information from unstructured data through cutting- edge OCR, machine learning, and deep learning algorithms [8]. The platform utilizes advanced machine learning, image processing, and deep learning techniques to extract relevant facts from unstructured data. It supports text extraction in over 100+ languages. It provides several Zapier integrations, is quick, precise, and easy to use, and allows users to create bespoke OCR models from scratch. Despite its strong engine and ease of use, Nanonets requires users to create a processing model [14]. If a new or different invoice type is introduced, the current model may not work, making it less reliable and introducing longer workflows, which could be a challenge for service providers across multiple industries. [16] Nanonets automates data collection for intelligent document processing, including ID cards, invoices, and receipts, using its AI-based OCR program. It offers several advantages, such as a modern user interface, the ability to handle large volumes of documents, affordability, ease of use, and cognitive data capture that minimizes the need for human intervention. It is particularly effective for automating document data extraction workflows using OCR and machine learning. Despite these strengths, one of the limitations of Nanonets is its table capture user interface, which is not user-friendly.

LightPDF

LightPDF offers free cloud-based services powered by AI, providing a simple and efficient way to read, edit, convert, sign, annotate, merge, manage, and distribute PDF files.

This private and secure PDF editor allows users to edit PDF documents just like Word files, making digital work easier to manage and more productive. LightPDF's applications include cloud-based PDF storage, ID photo creation, a PDF editor and reader, as well as a mobile scanner for on-the-go use. The platform is particularly useful for converting scanned images into editable documents, offering a free solution for converting images to PDF, Excel, or Word files. LightPDF's advantages include being a solid OCR program, with its basic version meeting the needs of most users, and being best for converting images into editable formats at no cost. However, one limitation of the tool is that its editor and tools can be slow and not always accurate.

Fig. 2. Sample output generated using Nanonet

The platform's AI-powered capabilities are demonstrated in Figure 3, which highlights the words and sentences that LightPDF can recognize in grey square boxes.

LightPDF offers free cloud-based services powered by AI, providing a simple and efficient way to read, edit, convert, sign, annotate, merge, manage, and distribute PDF files. This private and secure PDF editor allows users to edit PDF documents just like Word files, making digital work easier to manage and more productive. Even if the text is blurred or slanted, LightPDF can help you convert it into an editable format. LightPDF's applications include cloud-based PDF storage, ID photo creation, a PDF editor and reader, as well as a mobile scanner for on-the-go use. The platform is particularly useful for converting scanned images into editable documents, offering a free solution for converting images to PDF, Excel, or Word files. LightPDF's advantages include being a solid OCR program, with its basic version meeting the needs of most users, and being best for converting images into editable formats at no cost. However, one limitation of the tool is that its editor and tools can be slow and not always accurate.

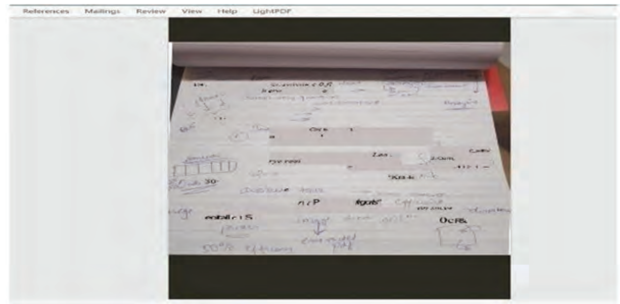


Fig. 3. Document processed using LightPDF OCR, showing extracted text

OCRSpace

OCRSpace is an web-based OCR service that converts scanned photos of text documents into editable files. This tool has a limitation in that the size of the PDF or image cannot exceed 5MB. The platform recognizes text and presents it in an editable format, with Figure 4 illustrating the words and sentences that OCRSpace can identify in its different output editors. OCRSpace can be accessed from any compatible browser on both PC and mobile devices, making it versatile and easy to use. It is best suited for the free online conversion of scanned photos or PDF files into editable Word documents. However, the tool has some limitations, thatis, being a bit slower and it is not suitable for processing larger images due to its file size constraint.

SimpleOCR

Simple OCR is a basic optical character recognition applica- tion that allows users to process all files in a given folder or se- lect individual images to extract text from them. Additionally, developers can utilize Simple OCR as a royalty-free OCR SDK in their own applications. Simple OCR is available for free to all users, including corporate users, educational institutions, and home users [6]. This tool enables fast and cost-effective scanning of documents with a scanner, eliminating the need for manual retyping. Simple OCR is an excellent solution for quickly converting paper documents into editable text. Figure 5 illustrates the words and sentences that Simple OCR can recognize in its output editor, along with suggestions for unpredictable words. The application provides several useful features, such as a despeckle option that improves accuracy for unclear documents, such as faxes or copies of copies. It also retains specific formatting elements such as font sizes, bold, italic, and underlining, preserving up to 99% of the original document format in some cases. Simple OCR is particularly effective at capturing and preserving images in addition to text, reducing the need for alternative methods to import photos from documents. The software can also extract plain text, which can be formatted later in word processors or HTML/web editors.The advantages of Simple OCR is that,it boasts high accuracy in converting scanned images to editable documents and can reduce speckles or dots in scanned images. It supports documents with non-standard fonts, multi-column layouts, and tables, making it ideal for batch OCR conversion of scanned images on Windows. However, the input and output formats supported by the application are limited, which may not meet the needs of all users.

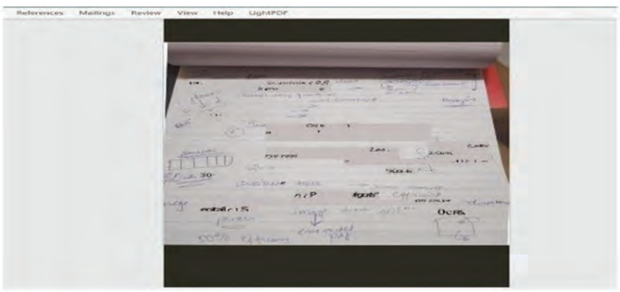


Fig. 4. Sample of text extracted using OCRSpace

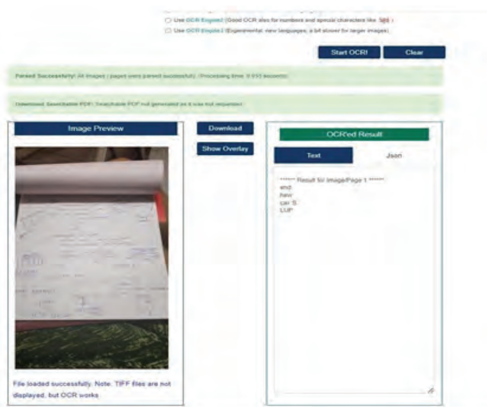


Fig. 5. Sample output generated from SimpleOCR

Fig. 6. Extracted text using NewOCR

NewOCR

NewOCR.com offers a free online optical character recognition (OCR) tool capable of analyzing text in uploaded image files and converting it into editable text on a computer. Supported input formats include JPG, TIFF, PDF, ZIP, JFIF, PNG, GIF, BMP, PBM, PGM, PPM, and PCX, as well as

compressed files like gzip, bzip2, bzip, and UNIX compress. The tool can process documents with poor scan quality, low resolution, and text in any orientation, as well as recognize mathematical formulas. It also supports documents with numerous pages, such as TIFF, PDF, DOCX, ODT files with images, and ZIP archives containing multiple images. Output formats include Microsoft Word (DOC), Adobe Acrobat (PDF), and plain text (TXT), with support for 75 to 106 languages depending on file types. The OCR service is highly effective for extracting text from PDFs, images, and other document types, making it particularly useful for processing newspaper and magazine articles with its integrated multi- column support. This service simplifies document conversion without the need to install software or create an account and imposes no restrictions on the number of uploads, enabling users to quickly and easily convert scanned documents and photos into editable text. However, limitations include the automatic deletion of files from the server after use and occasional reliance on APIs, which may become unavailable during server outages. Figure 6 highlights the tool's ability to recognize words and sentences in its output editors [7][10].

OnlineOCR

OnlineOCR.net is a free web-based optical character recognition (OCR) tool that enables users to convert scanned documents, PDF files (including multipage files), faxes, photos, and digital camera images into editable and searchable electronic documents without any file upload restrictions. The tool supports various input formats, such as JPEG/JPG, BMP, PCX, PNG, GIF, and ZIP, and can convert text images into output formats like RTF, Microsoft Word, Plain Text, Microsoft Excel, and Adobe PDF. It accommodates files up to 100 MB in size and supports 46 different languages. Users can effortlessly transform scanned documents and photos into editable text without prior registration, with the added option of registering for free to unlock advanced features, such as multi-page PDF conversion on any device. Figure 7 illustrates the words and sentences that OnlineOCR can recognize in different output editors, and it also provides downloadable output files. This tool is particularly suited for converting scanned images and PDF files online for free, making it a valuable resource for handling document conversions. Advantages include the ability to convert up to 15 pages per hour using the free version and the availability of advanced features upon registration. However, limitations exist, as the tool may struggle to recognize complex handwriting and non-standard fonts [9][7].

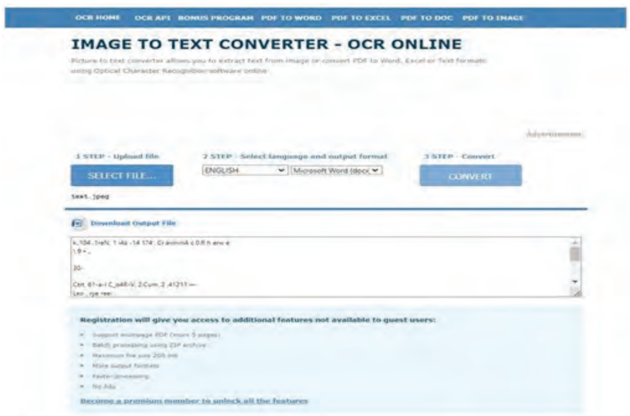


Fig. 7. Text extracted from a scanned image using OnlineOCR

IMPLEMENTATION OF A CUSTOM OCR MODEL

The Tesseract OCR engine, through its Python wrapper Pytesseract, is widely used in Python-based OCR applications, particularly for recognizing and reading text in images [5]. Pytesseract, also known as Python-tesseract, supports vari- ous image formats such as jpeg, png, gif, bmp, and tiff, thanks to compatibility with Pillow and Leptonica image libraries. As a standalone script for Tesseract, Pytesseract is particularly useful in applications requiring text extraction from images. Preprocessing steps like morphological thinning, detailing, normalization, denoising, binarization, and scaling are critical for ensuring the accuracy of Tesseract results [5, 15]. Images are preprocessed before being fed into the Pytesseract engine, which extracts unstructured text that can then be further processed into the desired format. This makes it highly effective in OCR use cases where images need to be converted into text, as illustrated in Figure 2, which demonstrates OCR implementation via API with Pytesseract. The tool is advantageous due to Tesseract's broad language support, recognizing text in over 100 languages, making it ideal for international applications and projects with diverse linguistic requirements. However, Tesseract has limitations in accuracy when compared to advanced AI-based methods and is prone to errors when the foreground and background of an image have minimal contrast. Additionally, creating a custom Tesseract OCR system demands significant effort and resources, which may pose challenges for developers. To illustrate the OCR process, a custom model using PyTesseract was implemented.

Dataset Description: A handwritten text dataset was cre- ated with diverse handwriting styles.

Preprocessing Techniques: The image undergoes grayscale conversion to simplify processing by reducing color com- plexity, focusing on brightness variations essential for text detection. Following this, thresholding is applied to convert the grayscale image into black and white, enhancing contrast between the text and background, which improves the OCR engine's ability to distinguish the text from noise.

Code Snippet

import matplotlib.pyplot as plt import pytesseract

import PIL import re

%matplotlib inline

img = PIL.Image.open(r'C:\Users\sweta\ Downloads\sample_inputs\test.jpeg') plt.imshow(img) pytesseract.pytesseract.tesseract_cmd='C:/ Program Files/Tesseract-OCR/tesseract' TESSDATA_PREFIX = 'C:/Program Files/ Tesseract-OCR' text_data=pytesseract.image_to_string(img. convert('RGB'),lang='eng') print(text_data)

Application Setup: The OCR task is conducted in a Python environment utilizing libraries such as matplotlib for display- ing images, PIL for image manipulation, and pytesseract to interface with the Tesseract OCR engine for text extraction. The Tesseract OCR engine is configured using its installation path and language data files. The input image (test.jpeg) is loaded with PIL, preprocessed, and converted to RGB before being passed to Tesseract for text extraction. The extracted text is then printed, and the original image is visualized using matplotlib for validation. The current implementation of the project is carried out in Jupyter Notebook, where data is pro- cessed, models are experimented with, and results are collected and analyzed. This environment facilitates testing various algorithms, preprocessing data, and visualizing intermediate steps, making it an efficient platform for experimentation and evaluation.

Fig. 8. Pytesseract output on English Text

COMPARATIVE ANALYSIS

Table 1 presents a comparison of OCR tools based on key metrics.

APPLICATIONS AND USE CASES

OCR technology has wide applications, including: label=•

- Information Documentation. This area includes ad- vances for accessing a large amount of sparse data. The frameworks are explained by looking at a very limited selection of printed characters, mostly numbers and a few special pictures. Record numbers, client distinguishing proof, article numbers, currency measures, and other information are among the things they are meant to examine. Each record in the paper designs must have a certain number of fixed lines for perusal[18].
- Banking. The applications of optical character recognition vary across different domains. When it comes to banking, people understand its application, in which automation technologies are used for optical character recognition; ATM is the greatest illustration of this. Another example from banking that comes to mind is the handling of checks, when a handwritten check is scanned and validated. It is possible to embed a register with a machine, filter its contents instantly, and transfer the appropriate amount of money. This has been nearly perfected for printed cheques and was initially accurate for transcribed cheques as well, but it occasionally needs manual confirmation. This generally reduces hold-up times in many banks[18].

Applications for optical character recognition vary across different domains. One may see its use in banking, where automation technologies are used for optical character recognition; ATMs are the finest illustration of this. Additionally, we can consider the handling of checks in banking, where a handwritten check is scanned and validated. It is possible to integrate a register with a machine, filter its contents instantly, and transfer the appropriate amount of money. This invention is nearly perfect for printed checks and was initially accurate for transcribed checks as well, although it occasionally needs manual confirmation. In many banks, this generally results in shorter hold-up periods.

Table 1 Comparison Table of OCR Models

Models Vs Measures	Nanonet	LightPDF	OCRSpace	Simple OCR	New OCR	Tesseract	Online OCR
OpenSource/ Online/API	Online	OpenSource	Online	OpenSource	Online	API	Online
Number of Input Files	1–10	1	1	1	1	1 (varies on dataset)	1
WordCount	No	No	No	Yes	No	No	No
Downloadable to Word/Text	Yes	No	Yes	No	Yes	No	No
Suggest Similar Words	No	No	No	Yes	No	No	Yes
Text (Type/ Heuristic)	Heuristic	Type	Type	Heuristic	Heuristic	Heuristic	Heuristic
Image Support	Any format	Any format	specifically png, jpg but is flexible with JPEG	JPG only	Any format	Any format	PNG, JPG

- Aid for individuals with visual impairments. Before the advent of sophisticated PCs and the need for partic- ipation, a lot of information was generated; Support for individuals with visual impairments was the intended area of use for comprehension machines. Accompanied by an such a reader would enable the ignorant to understand printed archives through a discourse union framework. Today's technology allows it to read, scan, and identify text. Another benefit for the blind is that we can turn the scanned text into voiced speech. This was the intended area of use for comprehending machines back in the day, before the development of sophisticated PCs and the need for participation, when a lot of information arose. Such a reader, when combined with a discourse union framework, would enable the uninformed to understand written archives. It can read, scan, and recognize text thanks to modern technologies. For the benefit of the blind, we may also turn the scanned text into voiced speech[18].
- HealthCare. Optical character recognition modification has also being used more frequently in healthcare to accomplish administrative tasks. Every victim needs to have a large number of structures taken care of by a medical care specialist, including both general wellbeing structures and defense frames. Entering relevant data into a device information base that is accessible as essential is useful for keeping track of all of this information. Devices under OCR control are able to separate data from structures and store it in databases, thus each victim's information is instantly recorded. Medical care providers can therefore focus on providing each patient with the best possible support[18].
- Number-plate Recognition. OCR is a technique employed in this technology that scans the numbers on the license plate from pictures and can be used to locate a car. To store the images captured by the cameras, ANPR has also been created which include the number that was taken from the permit plate. The district has explicitly innovated automatic number-plate recognition, which has led to plate assortment everywhere. These are employed by many law enforcement agencies as a solution for voltaic cost assortment on pay-per-use streets and to document the evolution of traffic or pedestrian This technology scans the numerals on the license plate from pictures using a process known as optical charac- ter recognition (OCR), which can be used to locate a car. Additionally, ANPR has been designed to store the images captured by the cameras that include the permit plate number. Since automatic number-plate identification is a district-specific innovation, it has contributed to plate assortment everywhere. These are employed by many law enforcement agencies as a method of charging for services on pay-per-use streets and to track the movement of traffic or people[18].

CONCLUSION AND FUTURE WORK

OCR is a technology that helps in detecting and recognizing words. Here we have studied the need for OCR, and various OCR tools and observed their measures of performance on different parameters analytically and experimentally respectively by taking sample inputs of many languages like Hindi, German, Spanish, Japanese, and Arabic apart from English and outcomes the same amount of output. To ensure reproducibility and facilitate future advancements, the complete implementation are accessible at https://github.com/swetalj/Folder. However, initially, optical characters that are recognized by the products were mostly developed by large technology companies. Individual researchers are enabled by the development of machine learning and deep learning to develop algorithms and techniques, which can recognize manuscripts with greater accuracy.

This study evaluated multiple OCR tools, demonstrating their strengths and limitations. PyTesseract was implemented as a custom OCR solution, showcasing its effectiveness. Future work will focus on enhancing handwritten OCR and support- ing additional languages.

FUTURE WORK

For extracting and analyzing the information from limited text and images associated with it we used deep learning to get semantic information and improve the richness of text data. Future works can be extended by taking input as, a large text and evaluating the grammar instead of image extraction as discussed here, and applying deep learning methods on that by developing the architecture to determine the font, text duplication, and making applications inferenced from GPT.

REFERENCES

1. J. Menon, M. Sami, R. A. Khan, and M. Uddin, "Handwritten Optical Character Recognition (OCR): A Comprehensive

Systematic Literature Review (SLR)," IEEE Access, 2020. DOI: https://doi.org/10.1109/ACCESS.2020.3012542 10.1109/AC- CESS.2020.3012542.

2. R. Ayyodri, S. Lukas, and H. Tjahyadi, "OCR for Text-Recognition and Its Post Processing Method: A Literature Review," IEEE, 2022. DOI: https://doi.org/10.1109/ICTIIA54654.2022.9935961 10.1109/IC-TIIA54654.2022.9935961.

3. A. Purohit and S. S. Chauhan, "A Literature Survey on Handwritten Character Recognition," IEEE, 2023. DOI: https://doi.org/10.1109/ICIDCA56705.2023.10100213 10.1109/ICIDCA56705.2023.10100213.

4. T. Kumar, C. Khosla, and K. Vashistha, "Character Recognition Techniques Using Machine Learning," IEEE, 2023. DOI: https://doi.org/10.1109/ICAISC58445.2023.10199173 10.1109/ICAISC58445.2023.10199173.

5. A. Kumar, P. Singh, and K. Lata, "Comparative Study of Different Optical Character Recognition Models on Handwritten and Printed Medical Reports," IEEE, 2023. DOI: https://doi.org/10.1109/ICIDCA56705.2023.1010021 10.1109/ICIDCA56705.2023.1010021.

6. B. Sharma and A. K. Rao, "OCR Related Technology Methods," International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE), 2020. DOI: https://doi.org/10.30534/IJATCSE/2020/47932020 10.30534/IJATCSE/2020/47932020.

7. S. Vijayarani and A. Sakila, "Performance Comparison of OCR Tools," International Journal of UbiComp, vol. 6, pp. 19-30, 2015. DOI: https://doi.org/10.5121/iju.2015.630310.5121/iju.2015.6303.

8. S. Ekande, Y. Chavan, V. Gurav, R. Joshi, and V. Rupnar, "A Compar- ative Study on OCR Engines for Invoice Analysis," 2023.

9. Online OCR Tool. Available: http://www.onlineocr.net/, accessed Dec. 15, 2023.

10. New OCR Tool. Available: http://www.newocr.com/, accessed Dec. 20, 2023.

11. Y. Fogel, N. Josman, and S. Rosenblum, "Functional abilities as reflected through temporal handwriting measures among adolescents with neuro-developmental disabilities," Pattern Recognition Letters, vol. 121, pp. 13-18, 2019. DOI: https://doi.org/10.1016/j.patrec.2018.07.00610.1016/j.patrec.2018.07.006.

12. W. Y. Lee, "Development of character recognition system based on the image processing techniques," The Society of Convergence Knowledge Transactions, vol. 5, pp. 99-103, 2017.

13. Simple OCR Tool. Available: https://www.simpleocr.com/, accessed Jan. 20, 2024.

14. S. Louis, P. Sonar, and P. Kaul, "An Image-based Intelli- gent System for Data Extraction," Asian Journal for Conver- gence in Technology (AJCT), vol. 8, no. 2, pp. 1-4, 2022. DOI: https://doi.org/10.33130/AJCT.v08i02.00110.33130/AJCT.v08i02.001.

15. C. Patel, A. Patel, and D. Patel, "Optical Character Recognition by Open Source OCR Tool Tesseract: A Case Study," International Journal of Computer Applications, vol. 55, no. 10, pp. 50–56, 2012.

16. Nanonets. Available: https://nanonets.com/online-ocr, Accessed: 15- Dec-2024.

17. LightPDF. Available: https://lightpdf.com/ocr-api, Accessed: 15-Dec- 2024.

18. Proddutur Shruthi and Dr. Devaraj Verma C, A Detailed Study and Recent Research on OCR, International Journal of Computer Science and Information Security (IJCSIS), Vol. 19, No. 2, February 2021.

JARVIS: Voice Assistant with Smart Home Automation

Rishabh Shah
Department of Artificial Intelligence and Data Science
D. J. Sanghvi College of Engineering
Mumbai, Maharashtra
✉ rishabhshah203@gmail.com

Yukta Saraf
Department of Artificial Intelligence and Data Science
D. J. Sanghvi College of Engineering
Mumbai, Maharashtra
✉ yukta2710@gmail.com

Alan George
Department of Artificial Intelligence and Data Science
D. J. Sanghvi College of Engineering
Mumbai, Maharashtra
✉ alangeorge3603@gmail.com

Rujuta Jariwala
Department of Artificial Intelligence and Data Science
D. J. Sanghvi College of Engineering
Mumbai, Maharashtra
✉ rujutajariwala@gmail.com

ABSTRACT

This paper explores the integration of virtual assistant technology with home automation systems. The voice assistant JARVIS leverages natural language processing to understand user requests and perform tasks such as information retrieval from Wikipedia, open YouTube, and perform mathematical calculations and manage alarms. It answers user queries using LLM models and simplifies daily routines, thereby revolutionizing our interaction with the technology in our homes. We utilize various artificial intelligence (AI) and natural language processing (NLP) libraries for improving speech recognition of our system. Our setup involves using Raspberry Pi 3B+ to link with IoT devices and sending signals over the Internet for home automation. We prioritize the security of the user by ensuring that the data does not get stored in the cache memory. This research takes a thorough look at how we can blend virtual assistant technology into home automation systems to make living spaces more accessible to users.

KEYWORDS *: Virtual assistant, Smart home automation, Artificial intelligence, Natural language processing, Raspberry Pi 3B+, Speech recognition, Large language models, Internet of things.*

INTRODUCTION

Home automation allows users to automate tasks and control the devices in their homes. This includes anything from remotely turning on lights, or adjusting the thermostat. The goal of this project is to create a convenient and efficient living environment for the user.

Internet of Things (IoT) are the interconnected devices that collect and share data, allowing researchers to make daily life easier through automation that streamlines daily tasks. The concepts of Home Automation and IoT are combined to form a Smart Home Automation system, which assists users to control various appliances at once.

Existing virtual home assistant (VHA) technology offers convenience for basic voice-based tasks like switching ON/OFF lights, operating smart plugs to control appliances, setting alarms, creating shopping lists etc.

Current virtual assistants excel at handling straightforward commands like adjusting lights or playing music. However, their limitations in natural language processing limit their ability to understand complex user requests. This research is focused on enhancing the potential of virtual assistants and transforming them into intelligent tools that make our life more efficient and convenient. Specific requests require a deeper understanding of user context and intent, which is where this research comes in.

We propose an approach that uses Large Language Models (LLMs), which are powerful AI models pre-trained on vast amounts of text data, which gives them an exceptional grasp of the English language. By incorporating LLMs into our virtual assistants, we aim to enhance their NLP capabilities. This will allow them to understand complex commands and enable our virtual assistants to understand the true meaning behind our words.

The virtual assistant will be able to tailor its responses based on the situation where real-time monitoring is necessary, by offering responses which are aware of the user's context, by analyzing the tone of voice and language patterns.

Security and user privacy are also extremely valued in our increasingly connected world. This research prioritizes user data protection by ensuring voice commands are not stored on the device itself.

To validate our proposed system, we will construct a real-world smart home system using Raspberry Pi 3B+, which allows us to develop and test a personalized system that seamlessly integrates into daily routines.

By empowering our assistants with advanced NLP capabilities and prioritizing user data privacy, we pave the way for a future where these become intelligent companions, working in the background to make our lives easier and more comfortable.

LITERATURE SURVEY

Recently there have been a lot of advancements in the fields of Artificial Intelligence and Home Automation, leading to development in studies, helping researchers in carrying out tasks, making the ideal benchmark for real-life use case scenarios of home automation, which in the last few years were not comprehensible. Some of the recent studies, as well as the related works of various researchers have set the benchmark/base for individual tasks, functions occurring in the realm of home automations.

Recent studies [2],[7] explored a secure and intelligent home automation system using Internet of Things (IoT) devices. One area of focus is secure communication. This might involve using Bluetooth for short- range communication within the home, combined with an internet- based verification process to ensure only authorized devices can connect [4]. Another key concept is a decentralized approach. This means devices can communicate and manage some functions directly with each other, without relying solely on a central hub. This can improve both reliability and security [3],[7]. Additionally, some studies explore using machine learning algorithms to categorize device states. For instance, the system might learn to identify when a light is on or off based on sensor data and then automate actions accordingly [8],[9].

In another recent study [1],[6] , a prototype for a home automation system built around a Raspberry Pi minicomputer was presented. The proposed system used various sensors to collect environmental data and offers remote appliance control and automation functionalities [5]. The key benefits of the automation system described in the paper were the remote monitoring and control which helped avoid situations where there would be a wastage of energy on forgotten appliances. The system also allowed scheduling the use of appliances where the system could be programmed to automatically switch ON/OFF appliances at particular times such as turning the lights off as soon as you left home etc. By using sensor data, the proposed system adjusted the appliance's usage of energy which optimized energy consumption [5],[6].

There were also some challenges regarding the proposed system which included the security of the web application used for controlling and monitoring the system [4],[8], integration with real world appliances outside the LEDs used for testing purposes to check the real world efficiency of the proposed system and the compatibility with various brands and appliances and a user friendly UI/UX interface while also keeping in mind the overall cost of the system including the hardware and software needed to be weighed against the benefits of the system [9].

SYSTEM ARCHITECTURE

The UI interface comprises various web pages which help in navigating to different modules. The paper examines four modules in detail, each designed for the distinct functionalities.

Module 1 (Deterministic Task Module)

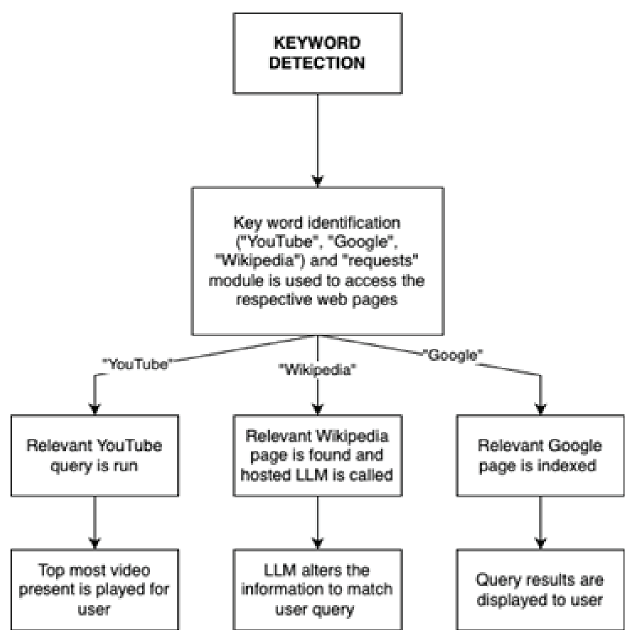


Fig. 1: Keyword Detection Module Architecture

This module is responsible for identifying predefined keywords or commands within user-generated queries. Upon detection, the system leverages these keywords to execute specific actions. The captured audio is then pre-processed to remove background noise, silence, and other unwanted sounds that could interfere with speech recognition. This can involve techniques like spectral filtering and noise gating. This step involves detecting the presence of the wake word in the audio input.

Module 2 (Mathematical and Reminder Module):

This module handles specific user commands such as performing mathematical calculations or setting reminders. The system accesses the appropriate services or datetime functionalities and displays results or confirmations to the user based on their request. This was achieved with the help of multi-threading, which allows for multiple tasks to run simultaneously.

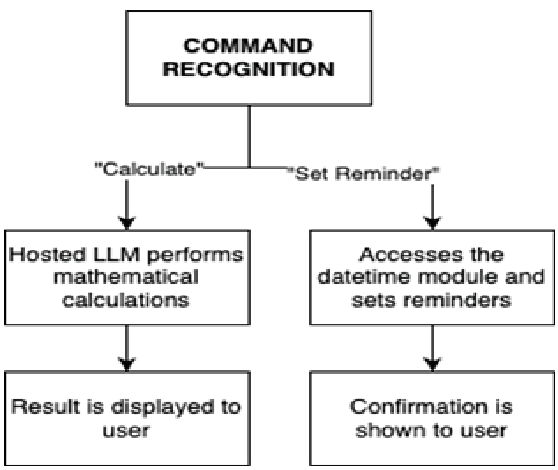


Fig. 2: Command Recognition Module Architecture

Module 3 (LLM Query Processing)

In this module, user queries are processed by detecting keywords. The system then forwards the query, with additional instructions, to the hosted LLM. The model interprets the query and returns a relevant response, which is shown to the user. We hosted the LLM (Llama 3.1 8b) via Groq which is then used for inference via an API. Prompt Engineering was put in place to filter out private data, harmful context, and prevent granting unauthorised data.

Module 4 (Home Automation)

This module interacts with smart home devices. It sends a signal to the Azure IoT Hub, updates the device status in a database, and communicates with the Raspberry Pi to control home appliances (turning them on or off) based on the user's voice command.

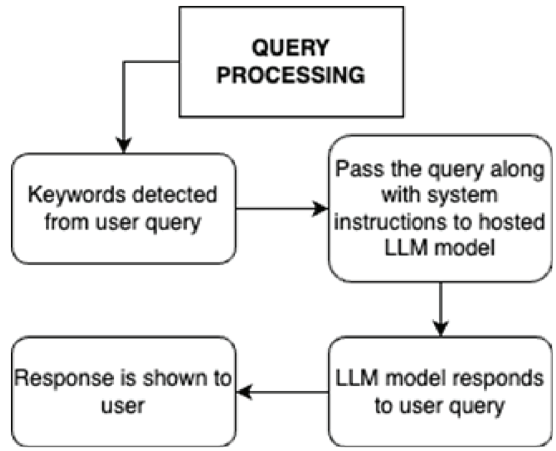


Fig.3: Query Processing Module Architecture

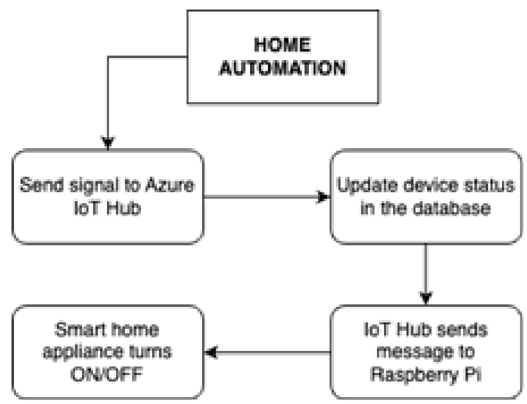


Fig. 4: Home Automation Module Architecture

PROPOSED METHODOLOGY

For every module, the user's audio is taken through a microphone and is processed to ensure seamless communication between the user and the smart home environment. This is a crucial part of the system architecture as it simulates the experience of talking to a personalized assistant.

The accuracy of speech recognition is improved by filtering out unwanted noise components present in the background while the user is conversing with the assistant. In the phase of Automatic Speech Recognition, the captured audio signals are converted into text format.The voice assistant is programmed to activate upon detecting the predefined wake word 'JARVIS' in the incoming audio stream. This triggers the interaction between the user and the assistant.

The text extracted from the Automatic Speech Recognition (ASR) component undergoes further analysis to extract meaning and context. Our system aims to decode the semantic meaning of the text output, discerning the user's intents and preferences.

The system's smart responses and personalized assistance is mainly due to the integration of a Large Language Model (LLM). The LLM is trained using multiple prompts which direct the system on how to respond based on different use cases ensuring relevant and accurate replies. Prompts are implemented such that the system filters out harmful queries and prevents exposure of confidential data. Based on the feedback from the user interactions, the model continuously updates its parameters and caters to the user's preferences over time.Our system then generates and synthesizes a response based on the context of the conversation history. This is crucial to ensure a coherent and contextually appropriate response.

To launch the various applications in the JARVIS system, we use PyAutoGUI to simulate keyboard and mouse actions. When the user provides a voice command to open an application, JARVIS uses PyAutoGUI to ensure smooth interaction with the OS. This guarantees that tasks such as opening applications or web browsers directly on the user's device.

To provide additional functionalities such as third party integration, the system makes use of external services and APIs for smart home automation.

We have utilized the following libraries and methodologies for the execution of the working system:

pysstx3: Pytt3x is a cross-platform text to speech library which is platform independent. The major advantage of using this library for text-to-speech conversion is that it works offline.

Speech Recognition: This allows us to convert audio into text for further processing.

Django: A high-level Python web framework that encourages rapid development and clean, pragmatic design. It includes a robust ORM, built-in admin interface, and tools for building secure web applications.

Djang-allauth: Django-Allauth is an authentication library for Django that provides user registration, login, and social authentication features, streamlining user management and enhancing security.

Cryptography: Cryptography secures information by transforming it into unreadable formats, ensuring data integrity and confidentiality through techniques like encryption and digital signatures.

Groq: Groq is a high-performance computing platform optimized for machine learning and data-intensive tasks, designed to accelerate workloads through its specialized architecture.

Azure Iot Hub: Azure IoT Hub is a managed service that enables secure communication between IoT devices and cloud applications, facilitating device management and data exchange at scale.

Requests: This is an elegant and simple HTTP library for Python that allows you to send HTTP/1.1 requests extremely easily.

RPi.GPIO: This is a Python library for controlling GPIO pins on Raspberry Pi, allowing users to set pin modes and manage inputs and outputs.

RESULTS AND CONCLUSION

For the implementation of the voice assistant with smart home automation technology integrated, JARVIS is first trained to respond to various natural language commands such that it can recognise any keywords spoken by an individual. This speech recognition technology is then leveraged to interact with the human queries and provide an efficient output.

While acting as a standard voice assistant, JARVIS is integrated with web based resources such as search engines and video sharing platforms. This enables users to access content which has been posted online without manually entering the query. Some of the most popular search engines and video sharing platforms accessed by JARVIS are Google and YouTube. The assistant can fetch information from Wikipedia and using LLMs, after retrieving the appropriate Wikipedia page, the content is paraphrased and altered to accurately answer the user's query.

Beyond this, JARVIS is also equipped to handle other tasks such as handling basic mathematical calculations,

setting alarms and creating reminders based on the user's schedule. Moreover, it can interact with the user in a way that feels engaging and personalized, enhancing user satisfaction by providing a more natural interface.

Furthermore, the voice assistant is also capable of home automation and can control various appliances via voice commands. Users can effortlessly switch smart lights and fans on and off, thereby enhancing convenience.

Apart from this, security is one of the key aspects of the system, with access being granted to trusted users only. This ensures that the interactions between the system and the user are kept private and that any information shared with the system is secured. The system also provides personalized recommendations based on user preferences and adapts its suggestions and responses based on the individual's tastes and needs, such as making suggestions for recipes based on the availability of ingredients in one's pantry.

To control home appliances through JARVIS, the user must first sign up for an account and access the "Devices" section of the interface. Here, the user is able to add new devices to their account or access previously added devices. After adding their preferred devices, users need to set up an IoT Azure code which is specific to a device. This allows for a secure connection and communication channel. After setting up the IoT Hub, a custom name is given to the device for easier identification and control.

Once this device specific configuration is saved, the users can control their devices via an ON/OFF button provided in the application interface. When the user presses a button, the system sends a signal to the IoT Azure Hub, which relays it to the connected Raspberry Pi 3B+. This device acts as an intermediate between the appliance and the system, controlling the state of the appliance by activating or deactivating the GPIO pins on it. This in turn switches on or off the connected appliances.

The features performed by JARVIS are:

- The voice assistant recognizes and responds to natural language commands
- It performs standard voice assistant functions such as using search engines to access sites like Google, YouTube, Wikipedia and others
- Smart lights/fans/other appliances can be turned ON/OFF via voice activation
- System is equipped to answer questions accurately and proficiently
- The system has security features to ensure user privacy and data protection
- Mathematical calculations can be performed seamlessly
- Setting up alarms and reminders
- Provides personalized recommendations based on user preferences.

Overall, JARVIS is a versatile system that combines the facilities of voice commanded control of household appliances with retrieval of information, basic computational tasks and provides personalized experiences to its users. The features of tailored recommendations as well as the security of the system, make it a robust solution for smart home environments.

FUTURE SCOPE

So far we have explored the functionalities of JARVIS as a smart voice based assistant integrated with a home automation segment. However, the capabilities of this system are not limited to the realm of smart homes and can be expanded to include various domains.

To build a smarter future, robust security protocols, anonymized user data, and user control over data collection are crucial for trust. Standardized protocols and open-source platforms could improve compatibility between smart home devices and voice assistants, while advancements in speech recognition algorithms will ensure understanding of diverse languages and accents. Voice-controlled smart homes of the future could utilize AR/VR activated by voice commands, allowing for virtual furniture placement and real-time appliance repair instructions. Voice assistants can become powerful tools for increased productivity and accessibility for users with physical limitations, allowing them to manage calendars, schedule appointments, or dictate documents using just their voice.

One critical challenge for JARVIS is a latency issue i.e. time lag between the user asking the question and the response generated. This directly affects real time interactions. This could be improved by optimising algorithms for faster response time and by incorporating edge/cloud based computing models. Furthermore, preloading commonly used tasks or commands can significantly reduce the latency for frequently accessed functionalities.

Another area of improvement is offline inference where the voice assistant can process basic tasks and respond to user queries even without a stable internet connection. This could offer offline responses but it comes with a trade-off, the pre-trained model's knowledge and capabilities will be limited to the data it was trained on, potentially hindering its ability to respond to new information. Future research can explore techniques for enabling local inference and improve the latency of the model.

Speech recognition technology has been advancing significantly, but ensuring high accuracy in noisy environments or interacting with people having diverse accents is a challenge. Future research can focus on areas such as expanding the training datasets to reflect diverse speech patterns, personalising models for better adaptation to individual speech patterns.

By addressing current limitations and embracing these future trends, voice assistants have the potential to transform our living spaces and enhance our lives in many ways.

REFERENCES

1. Valov, N., & Valova, I. (2020). Home Automation System with Raspberry Pi. 2020 7th International Conference on Energy Efficiency and Agricultural Engineering (EE&AE), Ruse, Bulgaria, pp. 1–5. DOI: 10.1109/EEAE49144.2020.9278998.
2. Tonoy, R. B. A., Mahmudunnabi, H., Zilany, H. M., & Rahman, R. M. (2021). A Smart and Intelligent Home Automation System. 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, pp. 446–452. DOI: 10.1109/UEMCON53757.2021.9666614.
3. Wang, Y., Chen, C., & Zhang, W. (2023). Design of Smart Home Control System Based on Internet of Things. 2023 IEEE 6th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chongqing, China, pp. 1680–1683. DOI: 10.1109/ITNEC56291.2023.10082440.
4. Reddy, P. S. N., Reddy, K. T. K., Reddy, P. A. K., Ramaiah, G. N. K., & Kishor, S. N. (2016). An IoT-Based Home Automation Using Android Application. 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES), Paralakhemundi, India, pp. 285–290. DOI: 10.1109/SCOPES.2016.7955836.
5. Iqbal, S., et al. (2021). Internet-of-Things-Based Smart Home Automation System Using Android Phone. Sir Syed University Research Journal of Engineering & Technology (SSURJET), 11(2).
6. Dubey, P., Chourasia, H., & Ghode, S. (2023). A Novel Approach to Design Home Automation Using IoT Applications. 2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT), Gharuan, India, pp. 801–806. DOI: 10.1109/InCACCT57535.2023.10141754.
7. Wang, Y., Rawat, S., & Metze, F. (2014). Semi-Automatic Audio Semantic Concept Discovery for Multimedia Retrieval. 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, pp. 1375–1379. DOI: 10.1109/ICASSP.2014.6853822.
8. Feki, I., Ammar, A. B., & Alimi, A. M. (2014). Query Sound-by-Example Video Retrieval Framework. 2014 14th International Conference on Hybrid Intelligent Systems, Kuwait, Kuwait, pp. 297–302. DOI: 10.1109/HIS.2014.7086165.
9. Chengalvarayan, R., & Deng, L. (1997). HMM-Based Speech Recognition Using State-Dependent, Discriminatively Derived Transforms on Mel-Warped DFT Features. IEEE Transactions on Speech and Audio Processing, 5(3), pp. 243–256. DOI: 10.1109/89.568731.

A Scalable Solution for CKD Diagnosis and Decision Support Tool for Early Intervention

Ragini Sharma
CSE- Data Science Department
Saraswati College of Engineering
Kharghar, Maharashtra
✉ ragini.sharma@it.sce.edu.in

Suyash Ghorpade
CSE- Data Science Department
Saraswati College of Engineering
Kharghar, Maharashtra
✉ suyash.ghorpade22@comp.sce.edu.in

Rajdeep Girkar
CSE- Data Science Department
Saraswati College of Engineering
Kharghar, Maharashtra
✉ rajdeep.girkar22@comp.sce.edu.in

Sahil Madge
CSE- Data Science Department
Saraswati College of Engineering
Kharghar, Maharashtra
✉ sahil.madge22@comp.sce.edu.in

ABSTRACT

The progressive decline of kidney function marks Chronic Kidney Disease (CKD) leading to renal failure when treatment fails to occur or disease diagnosis is late. A predictive model for CKD required machine learning implementation into an easy-to-use web application for the proposed system development. We ran supervised learning algorithms against clinical and laboratory parameters (such as age, blood pressure and serum creatinine) in order to detect indicator patterns for CKD. The dataset required cleaning through preprocessing followed by feature selection procedures to build an effective predictive model. Several tests including accuracy along with precision and recall and F1-score were applied during our evaluation to assess our machine learning models' performance. The selected model demonstrated solid stage prediction capabilities which establish it as a practical decision tool for healthcare professionals. Our team created a web application powered by a predictive model which integrates with a database through flask to deliver an easy-to-work-with platform for enhanced accessibility and usability. Users get 'predictions' instantly together with timely follow-up actions through our application. Research findings demonstrate how combining machine learning with web technologies solves critical healthcare challenges through scalable cost-effective diagnostic tools.

KEYWORDS *: CKD, Predictive model, Machine learning model, Web technologies, Diagnostic tools.*

INTRODUCTION

Chronic kidney disease destroys feathers to such an extent that blood filtration functions decline. The bloodstream releases waste and poison through the organs which primarily perform this function namely feathers [1]. This is how urine is made. Chronic kidney disease uses this term when kidneys accumulate waste products in the body. The onset of this condition continues for an extended duration which classifies it as habitual. The medical condition affects people around the world at present [2]. Many health problems stem from this condition which involves chronic kidney disease. Multiple factors beyond diabetes, high blood pressure or heart complaints contribute to CKD development. Age and gender affect the predisposition rates for developing CKD [3]. Those who face dyspepsia or diarrhoea along with defective feather function must consider persisting back pain as a symptom. Your feathers may sustain permanent damage from high blood pressure and diabetes according to medical literature [4]. The prevention of CKD requires successful management of these health conditions. The scarcity of symptoms in kidney disease makes people unaware they have it until advanced stages when treatment becomes challenging.

In this proposed system, we have evaluated key challenges surrounding predictive capabilities for CKD and propose solutions within our proposed system. These challenges include:

Complexity of CKD progression: Patients receive late-stage CKD diagnosis after development completes because CKD lacks recognizable warning signs. Our algorithms reveal early CKD detection mechanisms which identify hidden signals in patient statistics to facilitate prompt medical intervention decisions.

Integration of diverse data sources: Laboratory results and co-existing medical conditions merge with demographic and clinical symptom information to identify CKD in patients. The integration of multiple data types enables our approach to develop an extensive machine learning structure that improves the accuracy of CKD prediction.

Optimal feature selection and data preprocessing: Joint data preprocessing with feature selection methods enabled by correlation analysis as well as chi square evaluation and variance threshold boundaries enhances our model's diagnostic capability.

Model selection and validation: The selection process for an ML model to predict CKD demands finding the optimal fit for prediction. Our study evaluates multiple predictive models through performance assessment metrics to identify which one most accurately forecasts recurring cases of CKD.

LITERATURE SURVEY

M.A. Islam et al. (2023) [5] performed an analytical work which focused on CKD early detection with machine learning methods. Researchers worked with 400 patient cases using 24 attributes, including 13 categorical and 11 numerical variables. Principal Component Analysis was used to analyze key features following preprocessing to establish CKD predictions. The XGBoost classifier delivered outstanding results with original data accuracy at 98.33% and achieved 99.16% accuracy when subjected to PCA processing.

R. Sawhney (2023) [6], utilized 400 CKD cases with 24 core features including numeric and categorical elements for AI model development to detect and assess condition risks. Backpropagation enabled their use of a Multilayer Perceptron (MLP) neural network component. The Artificial Neural Network (ANN) achieved superior performance when used as a classifier because testing accuracy reached 100% surpassing Support Vector Machine (SVM) at 82% and Logistic Regression (LR) at 96%.

Poonia RC et al. (2022) [7] created a kidney disease detection methodology leveraging 400 patient cases with 24 characteristic markers. The proposed system used KNN, ANN, Naive Bayes and SVM algorithms while employing RFE and chi square to determine relevant features. A Chi square selected feature within an enhanced Logistic Regression model delivered 98.75% accuracy.

Pal. S. (2022) [8] applied a model prediction of CKD through analysis of a dataset containing 400 cases based on 24 UCI machine learning repository features. The research implemented the classifiers Logistic Regression (LR) and Decision Tree (DT) along with Support Vector Machine (SVM) before optimizing their performance through bagging ensemble methods. When the Decision Tree (DT) classifier was combined with the bagging method, accuracy improved from 95.92% to 97.23%.

PROBLEM STATEMENT

Chronic Kidney Disease exists as a global public health issue which develops silently within many patients before intensifying. People find it difficult to get early diagnosis during the early stages of the disease because preventing permanent kidney injury remains challenging while healthcare costs increase. The present assessment methods rely on costly laboratory tests and time-consuming clinical evaluations that make it difficult for individuals to access them. The self-generated analysis of medical data faces delays and imprecisions when attempting early detection.

The research addresses limitations of previous diagnostic in CKD detection through the development of an affordable accessible method to assist healthcare professionals in their diagnostic work.

METHODOLOGY

In order to analyse the CKD dataset, a number of experiments were carried out using different types of machine-learning algorithm, including Support Vector Machine, Decision Tree and Logistic Regression.

Table 1: In-depth descriptions of each attribute present in the main CKD dataset

CKD dataset	Attributes meaning	Category	Scale
age	Age	Numerical	Years
bp	Blood pressure	Numerical	mm/hg
sg	Specific gravity	Nominal	1.005-1.025
al	Albumin	Nominal	0 to 5
su	Sugar	Nominal	0 to 5
rbc	Red blood cells	Nominal	Abnormal, normal

pc	Pus cells	Nominal	Abnormal, normal
pcc	Pus cells clump	Nominal	Not present, present
ba	Bacteria	Nominal	Not present, present
bgr	Blood glucose random	Numerical	mgs/dl
bu	Blood urea	Numerical	Mgs/dl
sc	Serum creatinine	Numerical	Mgs/dl
sod	Sodium	Numerical	mEq/L
pot	Potassium	Numerical	mEq/L
hemo	Hemoglobin	Numerical	gms
pcv	Packed cell volume	Numerical	Pcv
wc	White blood cell count	Numerical	Cells/cumm
rc	Red blood cell count	Nominal	Millions/ cm m
htn	Hypertension	Nominal	No, yes
dm	Diabetes mellitus	Nominal	No, yes
cad	Coronary artery disease	Nominal	No, yes
appet	Appetite	Nominal	Poor, good
pe	Peda edema	Nominal	No, yes
nae	Anaemia	Nominal	No, yes
Classification	Class	Nominal	Not ckd, ckd,

The research employed CKD dataset [9] including 400 records spread across 14 columns. The "class" output column contains data points maintaining either "yes" or "no" values. Responses indicating "yes" obtained numeric value "1" whereas "no" yielded numerical value "0". Each instance of CKD received the value "1" while instances of non-CKD received the value "0". A different set of features was used by each machine learning model across all data points.

Fig. (1) displays a confusion matrix for Chronic Kidney Disease (CKD) classification which evaluates actual cases alongside predictions. This classification shows True Positive (TP) results for correct CKD predictions and False Positive (FP) cases for misdiagnosed non-CKD conditions yet it shows False Negative (FN) and True Negative (TN) outcomes when detecting CKD patients correctly and non-CKD subjects accurately respectively. The matrix helps in measuring the model accuracy and precision together with its recall and F1-score performance for detecting Chronic Kidney Disease cases [10].

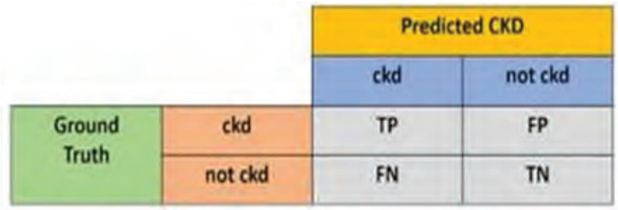


Fig. 1: Confusion Matrix Format

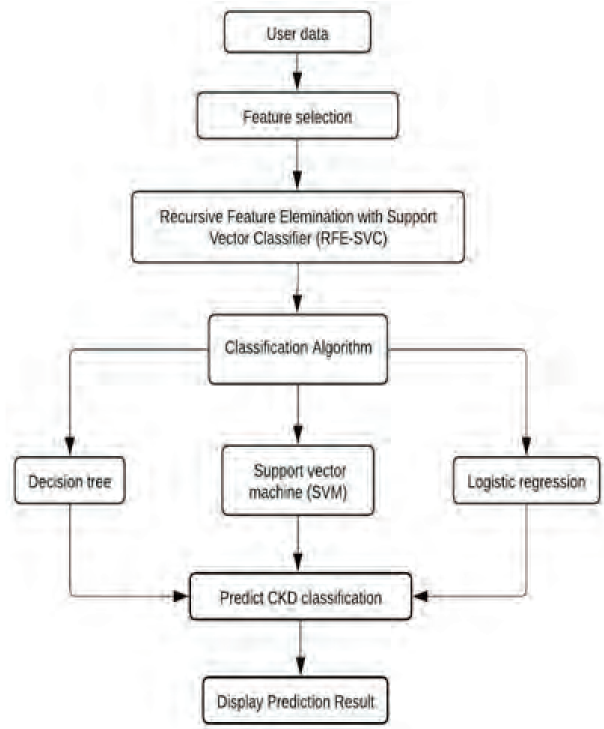


Fig. 2: Workflow of CKD Prediction Using

Classification Models

The Fig. (2) illustration demonstrates how recursive element elimination with a probative vector classifier (RFE-SVC) removes nonrelevant features while significant features serve to refine classification. Different algorithms including Decision Tree and Support Vector Machine (SVM) and Logistics Regression use named features for implementation. Each algorithm produces predictions about CKD presence through analysis of specified properties.

The high-dimensional data processing capability of SVM exists alongside decision tree algorithms which demonstrate interpretability and flexibility. The statistical model of logistic regression shows robust performance for binary classification problems including CKD prediction. The processing of estimated results leads to user classification as CKD or non-CKD. The system presents an intuitive action forecast outcome to users after processing the results.

This approach utilizes data attribute selection methods together with multiple distinct modelling approaches to build an accurate decision framework that supports CKD diagnosis.

RESULTS

A data-driven approach to predict chronic kidney disease generates significant value that improves both clinical diagnostic tools and healthcare operational systems. RAM Challenge supports low-cost patient access to diagnostic protocols for CKD that otherwise would be inaccessible because of rising healthcare expenditures. Simple health input into the system generates immediate predictions of CKD status for users since the recorded data helps improve disease diagnosis and accessibility features. This innovative solution enhances early detection, enabling timely interventions and personalized treatment plans.

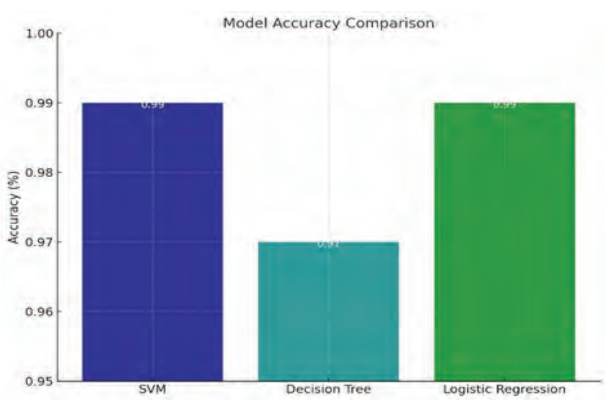


Fig. 3: Model Accuracy Comparison

In Fig. (3), the bar chart illustrates a comparative analysis of the accuracy of three machine learning models: Support Vector Machine (SVM), Decision Tree and Logistic Regression. The graph displays accuracy scores related to y predictions together with model performance results for x. All three machine learning models demonstrate similar performance based on accuracy scores close to 1.0 (100%). All projected results display identical accuracy values because each model demonstrates exceptional prediction capabilities for the dataset.

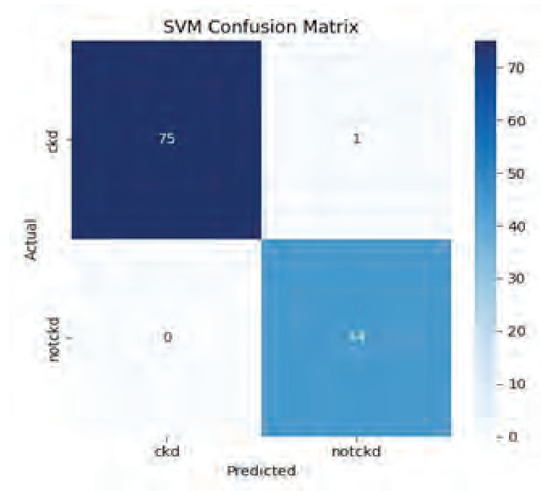


Fig. 4: Confusion Matrix of SVM

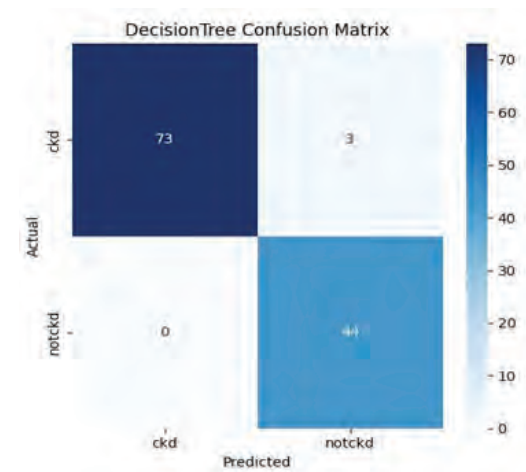


Fig. 5: Confusion matrix of Decision Tree

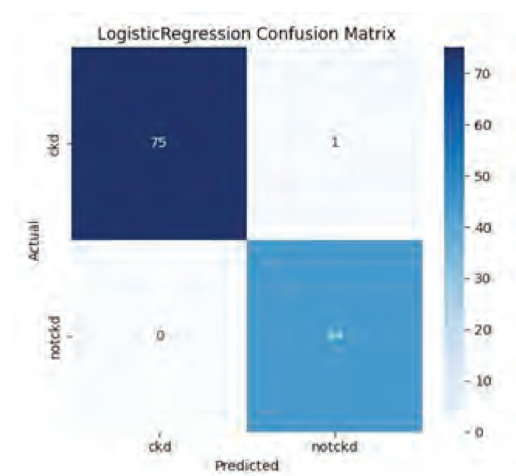


Fig. 6: Confusion Matrix of Logistic Regression

The assessment of chronic kidney disease (CKD) through three models appears in confusion matrices Fig. (4), Fig. (5) and Fig. (6). These models consist of Support Vector Machine (SVM) and Decision Tree and Logistic Regression. The confusion matrices demonstrate actual labels in rows and predicted labels in columns by tracking true-positive, true-negative and false-positive and false-negative cases. The performance of SVM matches that of Logistic Regression as they both produce 75 true positives, 44 true negatives as well as no false positives and 1 false negative. The Decision Tree model achieved 73 true positives but revealed three false negatives alongside its results.

Results from the trained SVM, Decision Tree, and Logistic Regression models were evaluated in the below table which displays the accuracy, precision and recall rates along with F1 scores to determine the chronic kidney disease (CKD) classification.

Table 2: Classification report of models

Model	Accuracy	Precession	Recall	F1 score
SVM	99.1666%	0.99%	0.99%	0.99%
Decision Tree	97.5%	0.97%	0.98%	0.97%
Logistic Regression	99.1666%	0.99%	0.99%	0.99%

The combined precision, recall, and F1 scores for SVM and Logistic Regression models reached 0.99 during the execution of all metrics and maintained an accuracy rate of 99.166667%. The accuracy of the Decision Tree model reached 97.5% while presenting precision at 0.97%, recall at 0.98% and an F1 score of 0.97% but exhibited minimal trade-offs between these metrics.

Both SVM and Logistic Regression algorithms showed consistent excellent performance and precise-recall balance throughout various experimental conditions in CKD classification.

CONCLUSION

Our study aimed to create a machine-learning diagnostic system for early chronic kidney disease (CKD) detection. Several research studies identified various interconnected predictors that scientists turned into simplified combinations while discarding overlapping details. Our filtering process selected haemoglobin along with albumin and specific gravity as vital predictors that exposed renal disease diagnostic patterns in medical practice.

We created a web application focused on user needs to facilitate the entry of kidney health parameters alongside related risk factors. The resulting dataset along with our trained classifiers enables our tool to deliver quick diagnostic and risk assessment solutions for kidney diseases thus demonstrating the real-life application of our research for early health prevention.

REFERENCES

1. Chen Z, Zhang X, Zhang Z, “Clinical risk assessment of patients with chronic kidney disease by using clinical data and multivariate models.” Int Urol Nephrol, vol. 48, pp. 2069–2075, June 2016.

2. Charleonnan A, Fufaung T, Niyomwong T, Chokchueypattanakit W, Suwannawach S, Ninchawee N, “Predictive analytics for chronic kidney disease using machine learning techniques.” in 2016 Management and Innovation Technology International Conference (MIT icon), Sep. 2017.

3. Subasi A, Alickovic E, Kevric J, “Diagnosis of chronic kidney disease by using random forest.” CMBEBIH 2017 vol. 62, pp 589-594, Mar. 2017.

4. Zhang L, Wang F, Wang L, et al, “Prevalence of chronic kidney disease in China: a cross-sectional survey.” in The Lancet vol. 379, pp. 815-822, Mar. 2012

5. Islam MdA, Majumder MdZ, Hussein MdA, “Chronic kidney disease prediction based on machine learning algorithms” Journal of Pathology Informatics vol. 14, 2023.

6. Sawhney R, A Malik, S Shilpa, V Narayan, “A comparative assessment of artificial intelligence models used for early prediction and evaluation of chronic kidney disease” Decision Analyst Journal, vol. 6, Mar. 2023.

7. Poonia RC, Gupta MK, Abunadi I, “Intelligent diagnostic prediction and classification models for detection of kidney disease” Healthcare vol. 10, Feb. 2022.

8. Pal S, “Chronic Kidney Disease prediction using machine learning techniques” Biomedical Materials and Devices, vol. 1, pp. 534-540, Aug. 2022.

9. L. Rubini, P. Soundarapandian, P. Eswaran, “Chronic Kidney Disease”, UCI Machine Learning Repository, Feb. 2017.

10. Rácz A, Bajusz D, Héberger K, “Multi-level comparison of machine learning classifiers and their performance metrics”, Molecules, vol. 24, Aug. 2019.

Electric Vehicle Adoption and Its Impact on Smart Cities

Mayuri Lohar
Assistant Professor
Information Technology
Atharva College of Engineering
Malad, Maharashtra

Snehal Mahajan
Assistant Professor
Information Technology
Atharva College of Engineering
Malad, Maharashtra

Kalpana Mane
Assistant Professor
Information Technology
Atharva College of Engineering
Malad, Maharashtra

Snehal Gholap
Assistant Professor
Information Technology
Atharva College of Engineering
Malad, Maharashtra

ABSTRACT

Electric vehicle (EV) adoption is transforming urban mobility and is a key factor in the creation of smart cities. EVs provide a sustainable substitute for conventional fossil fuel-powered automobiles, helping to lower greenhouse gas emissions while also improving air quality and reducing noise pollution. The extensive usage of EVs fits in perfectly with smart cities' priorities of sustainability, energy efficiency, and technology integration.

This paper explores the multifaceted impact of EV adoption on smart cities, focusing on key areas such as infrastructure, energy management, and urban planning. It emphasizes how EV charging networks can be integrated into intelligent energy grids to facilitate effective energy distribution and encourage the use of renewable energy sources. Additionally, it examines the role of data-driven solutions, such as IoT-enabled sensors and real-time analytics, in optimizing EV usage and traffic management.

The paper also addresses challenges, including the need for scalable charging infrastructure, equitable access to EV technology, and the mitigation of increased energy demand. The switch to EVs can hasten the development of smart cities into resilient, connected, and sustainable urban settings by encouraging cooperation between communities, businesses, and governments. According to this study, the adoption of EVs is a revolutionary force that is influencing urban living in the future rather than only being a technological advancement.

***KEYWORDS** : Electric vehicles, Smart cities, Sustainable transportation, Charging infrastructure, Smart grids, IoT, Urban mobility, Renewable energy, Environmental sustainability, Autonomous vehicles.*

INTRODUCTION

The 21st century's fast urbanization has created serious problems with energy use, transportation, and environmental damage. Due to their heavy reliance on fossil fuels, traditional transportation systems significantly increase carbon emissions and air pollution. There has never been a more pressing need for cleaner, greener, and more effective mobility solutions as the globe struggles with the negative effects of climate change. As a sustainable option that lessens environmental effect and supports international efforts to reach net-zero emissions, electric vehicles (EVs) have emerged as a top contender in tackling these issues.

At the same time, the idea of smart cities has become a major focus of urban development plans. In order to improve urban living standards, smart cities seek to leverage cutting-edge technologies like big data, artificial intelligence (AI), and the Internet of Things (IoT). These cities foster an environment where technology and urban design meet by placing a high priority on sustainability, connectivity, and effective resource management. Since EVs are essential to environmentally friendly transportation, the smart city concept is a perfect fit for them. Their incorporation into urban environments represents a rethinking of how cities operate and prosper in the face of 21st-century difficulties, not just a technological advancement.

This paper delves into the symbiotic relationship between EV adoption and smart city development, exploring how the two domains can synergize to create sustainable, connected, and resilient urban environments. It investigates the environmental, economic, and technological dimensions of this transformation, providing a comprehensive framework for policymakers, urban planners, and industry stakeholders. By examining both opportunities and challenges, this study seeks to illuminate the path toward a future where EVs and smart cities coexist harmoniously, driving innovation and raising the standard of living for those living in cities across the globe.

THE ROLE OF EVS IN SMART CITIES

Environmental Benefits

Greenhouse gas emissions are greatly decreased by EVs, especially when they are fuelled by renewable energy. This reduction contributes to improved air quality, which is vital for public health and urban livability. Smart cities prioritize green technologies, and EVs seamlessly integrate into this framework by lowering the carbon footprint of transportation.

Enhanced Urban Mobility

The adoption of EVs supports multimodal transportation systems, including shared mobility solutions such as electric buses, e-scooters, and ride-hailing services. This facilitates efficient, sustainable, and convenient urban mobility, a key feature of smart cities.

Noise Pollution Reduction

Compared to internal combustion engine vehicles, electric vehicles are quieter, which greatly lowers noise levels in cities. This improves the general quality of life by making urban areas more livable and pleasant.

INFRASTRUCTURE REQUIREMENTS

Charging Infrastructure

A vast network of charging stations is necessary for the broad adoption of EVs. Smart cities must integrate charging infrastructure into urban planning, ensuring accessibility and convenience for EV users. Innovative solutions such as wireless charging and fast- charging technologies are critical to addressing this demand.

Smart Grids

Efficient energy management is made possible by the integration of EVs with smart energy infrastructures. EVs can function as mobile energy storage devices thanks to vehicle-to-grid (V2G) technology, which promotes grid stability and the use of renewable energy sources. This bidirectional energy flow is a cornerstone of sustainable urban energy systems.

Urban Design and Planning

Smart cities must adapt their urban landscapes to accommodate EVs, including dedicated parking spaces with charging facilities, designated EV lanes, and optimized traffic flow systems. Urban planning must align with the evolving transportation needs of EV users.

TECHNOLOGICAL INTEGRATION

Internet of Things (IoT) and Big

Real-time data analytics and IoT-enabled sensors are essential for tracking and controlling EV usage. By analyzing EV patterns, smart traffic management systems can improve overall transportation efficiency, ease congestion, and optimize routes.

Artificial Intelligence (AI)

AI-powered applications can predict energy demand, optimize charging schedules, and improve battery efficiency. These advancements ensure that EVs are seamlessly integrated into the urban ecosystem.

Connectivity and Automation

The convergence of EVs with autonomous vehicle technology has the potential to transform urban mobility further. Connected EVs can communicate with city infrastructure, traffic systems, and other vehicles, creating safer and more efficient transportation networks.

FUTURE SCOPE

One major factor influencing how smart cities will develop in the future is the uptake of electric vehicles, or EVs. It is anticipated that the incorporation of EVs into infrastructure and urban design would have a major influence on a number of industries. The impact of electric car adoption on smart cities and their potential future scope are broken down below:

Environmental Impact

- Reduction in Emissions: Since EVs have no exhaust emissions; air pollution is decreased, especially in crowded cities. This is in line with the environmental objectives of smart cities, which include lowering carbon footprints and attaining sustainable growth.
- Lower Noise Pollution: EVs are generally quieter than internal combustion engine (ICE) vehicles, contributing to quieter urban environments, especially in residential areas.
- Energy Transition: The transition to renewable energy sources may be aided by the growing use of EVs. EVs may be charged using solar, wind, and other renewable energy sources, which lessens the need for fossil fuels.

Energy Infrastructure and Grid Management

- Demand on the Power Grid: EVs increase electricity demand, especially if charging stations are not well-distributed or if many people charge vehicles at peak times. Smart grids can manage this demand by balancing energy use through real-time data and optimizing when and how EVs are charged.
- Vehicle-to-Grid (V2G) Technology: V2G essentially transforms automobiles into mobile power storage by enabling EVs to transmit electricity back to the grid. This can facilitate a more decentralized energy distribution system and aid in grid stabilization during periods of high demand.

Urban Mobility and Traffic Management

- Integrated Mobility Systems: EVs, along with autonomous vehicles (AVs), could revolutionize urban transportation. Smart cities can integrate EVs with other modes of transportation like shared bicycles, buses, and trains to provide seamless, efficient mobility.
- Reduced Traffic Congestion: By lowering the number of private vehicles on the road, EVs are frequently utilized in ride-hailing and car-sharing services, which help to ease traffic congestion. As a result, there is less demand for substantial parking infrastructure and transportation networks operate more efficiently.

Infrastructure Development

- Charging Stations: Smart cities will need to establish an extensive network of EV charging stations. These stations will be integrated into urban infrastructure, with some smart charging systems offering fast charging and optimized energy use, powered by renewable sources.
- Urban Planning: Smart cities may develop more EV-friendly urban layouts, including dedicated lanes for EVs, charging facilities in public spaces, and green areas designed for reduced vehicular traffic.

Economic Impact

- Job Creation: The EV industry, including manufacturing, infrastructure development, and maintenance services, is expected to create numerous jobs. This contributes to local economic growth and offers new opportunities in tech-driven sectors.
- Energy Cost Savings: Because EVs require less gasoline and maintenance than conventional cars, they typically have cheaper running expenses. Over time, this may lower the cost of urban travel.

Data and Connectivity

- Smart Transportation Systems: Real-time information on traffic, weather, and road conditions can be obtained by equipping EVs with sensors and connecting them to the Internet of Things (IoT). By maximizing traffic flow and lowering accident rates, this can improve the effectiveness of urban transportation networks.
- Autonomous EVs: The development of self-driving EVs can further streamline urban mobility by reducing the need for human drivers, resulting in more effective and safe transportation.

Challenges and Future Scope

- Battery Recycling and Sustainability: As EV adoption grows, so does the need for sustainable battery recycling practices to avoid environmental harm. To guarantee that the environmental advantages of EVs are fully realized, it will be essential to develop technology for efficient EV battery recycling.
- Cyber security: With the increased connectivity of EVs and their integration into smart cities, cyber security becomes a significant concern. Safeguarding EVs, charging stations, and grid systems against

cyber-attacks will be essential for the future of smart cities.

- EV Accessibility: Ensuring equitable access to EV technology and charging infrastructure is a challenge. Smart cities will need to address issues related to the affordability of EVs and access to charging stations for people in lower-income or underserved areas.

Smart City Synergy

The adoption of EVs in smart cities is not just about the vehicles themselves, but about how they interact with other technologies and systems:

- Blockchain for EV Tracking: Blockchain could be used to securely track EV ownership, charging, and maintenance history, contributing to greater transparency in the EV ecosystem.
- Artificial Intelligence (AI): AI can optimize traffic management and predict charging demand, contributing to smoother traffic flow and better energy distribution.
- Urban Sustainability: The objectives of lowering greenhouse gas emissions, enhancing air quality, and establishing a healthier urban environment are more likely to be accomplished by including EVs into a smart city's sustainability framework.

CHALLENGES AND OPPORTUNITIES

Challenges

Infrastructure Investment:

Establishing a comprehensive charging network requires significant financial investment and careful planning.

Energy Demand:

Increased EV adoption can strain energy grids, necessitating advancements in grid capacity and renewable energy integration.

Equity and Accessibility:

Ensuring that EV technology is accessible to all socioeconomic groups is a critical challenge for policymakers.

Opportunities

- Economic Growth: The EV industry fosters innovation, creates jobs, and stimulates economic growth in urban areas.

Public-Private Partnerships

Collaboration between governments, private companies, and communities can drive the development of EV infrastructure and policies.

Sustainability Goals

EV adoption accelerates the achievement of sustainability and climate goals, enhancing a city's reputation and livability.

CASE STUDIES

This section examines successful examples of EV integration in smart cities worldwide, including:

- Oslo, Norway: A leader in EV adoption with robust charging infrastructure and policy incentives.
- Singapore: Leveraging technology and data to promote EV usage and optimize urban mobility.
- San Francisco, USA: Integrating EVs into shared mobility networks and renewable energy grids.

CONCLUSION

The adoption of electric vehicles is a transformative force that aligns seamlessly with the goals of smart cities. By addressing environmental concerns, enhancing urban mobility, and fostering technological innovation, EVs contribute to sustainable, connected, and resilient urban environments. However, achieving this vision requires overcoming challenges related to infrastructure, energy management, and accessibility. Through collaborative efforts and innovative solutions, EVs have the potential to redefine urban living and drive the evolution of smart cities.

REFERENCES

1. Mullan, D., et al. (2022). "The Role of EVs in Smart Cities: Insights from the European Union." Journal of Urban Technology, 29(1), 73-98.

2. International Energy Agency (IEA) (2021). "Global EV Outlook 2021." IEA.

4. Sperling, D., & Gordon, D. (2020). Two Billion Cars: Driving Toward Sustainability. Oxford University Press.

5. Zhao, Y., & Gao, Y. (2020). "Electric Vehicles in Smart Cities: Key Technologies and Future Trends." IEEE Access, 8, 190271–190285.

6. Tushar, W., et al. (2020). "Smart Grid and Electric Vehicles: A Comprehensive Review." Renewable and Sustainable Energy Reviews, 119, 109600.

7. Alonso, M., & Figueroa, A. (2019). "Electric Vehicle Integration in Smart Cities: A Literature Review." Renewable and Sustainable Energy Reviews, 113, 109276.
8. Jochem, P., et al. (2016). "The Impact of Electric Vehicles on the Environment and Economy." Energy Policy, 93, 144-154.
9. He, H., et al. (2019). "Electric Vehicle Charging Infrastructure Planning for Smart Cities." Energy, 177, 255-267.
10. Cuenot, F., et al. (2017). "Electric Vehicles and Smart Grids: A New Era of Energy Systems." Energies, 10(7), 972.
11. Gkatzoflias, D., et al. (2020). "Electric Vehicles and Smart Cities: The Role of Digital Technologies and Connectivity." Environmental Science & Technology, 54(9), 5317-5328.
12. Liu, Y., et al. (2020). "Smart Charging for Electric Vehicles: A Survey." IEEE Transactions on Industrial Informatics, 16(2), 1343-1354.
13. Chien, S., et al. (2021). "A Review of the Impacts of Electric Vehicle Adoption on Urban Mobility." Transportation Research Part D: Transport and Environment, 88, 102529.
14. IEA (2022). "The Role of Electric Vehicles in Clean Energy Transitions." International Energy Agency.
15. Rajasekaran, A., et al. (2021). Cybersecurity in Electric Vehicle Infrastructure: Challenges and Solutions for Smart Cities." IEEE Transactions on Smart Grid, 12(3), 2622-2632.

Blockchain based Decentralised Social Media Platform

Vaishali V. Bodade
Associate Professor
Dept. of Information Technology
Fr. C. Rodrigues Institute of Technology
Navi Mumbai, Maharashtra
✉ vaishali.bodade@fcrit.ac.in

Jyoti S. More
Dept. of Computer Engineering
Fr. C. Rodrigues Institute of Technology
Navi Mumbai, Maharashtra
✉ jyoti.more@fcrit.ac.in

ABSTRACT

A social networking platform called Decentragram uses blockchain technology to support the design of applications and smart contracts. It is produced via blockchain platforms and protocols. The social media creation process is facilitated by blockchain technologies. Decentralised applications like Ethereum, Steem, and Stellar are just a few examples. Decentralisation refers to the absence of a single proprietary authority managing all the data in blockchain social media networks. Instead, the information is evenly disseminated and kept on servers at each network node. Users that create profiles on integrated social platforms must link their real-world identities, such as their electronic mail addresses or contact details. Moreover, these networks usually use encrypted public keys for authentication of accounts rather than trusting on just one company to secure user data. While this offers benefits in terms of data security, there are shortcomings as well. Blockchain-based social media systems provide end-to-end encryptions for every contact while still supporting social networking, content sharing, and even blogging. This decentralised social media offers an additional solution for data security and privacy.

KEYWORDS *: Blockchain, Security, Decentralised, Social media.*

INTRODUCTION

One's life now includes a significant amount of social media. We use it to create communities, strengthen bonds, and stay in contact with pals who live far away. We use it to communicate our aspirations, anxieties, travel images, and pet videos. We also employ it to keep up with local, national, and worldwide news as well as politics. But centralised social media keeps violating our confidence. Appearingly innocent games and quizzes convey information about ourselves and our friends to massive organisations with electoral votes and financial gain in mind. Media streams don't provide precise and comprehensive viewpoints on present-day affairs. You're followed by marketing firms while you visit other websites.

Indeed, a new generation of creative software developers is creating novel platforms that, via decentralisation, avoid the risks associated with the present social network system. Instead of using a centralised server controlled by a company, decentralised social media platforms function on servers that are independently managed. This social network offers more power and liberty. A person may design their own social network, choose how it functions, and limit what people can post. The creator of an interconnected social network can set the rules of appropriate behaviour for the website rather than having material regulated by a business.

For the first time since the development of popular social media platforms, fewer Americans in 2018 have at least one social media profile. Social media networks are being discontinued by users for a variety of reasons, ranging from their adverse impact on mental health to users' overall suspicion of platform owners.

Users are growing more and more conscious of the control that a few individuals have over their data. A shift towards decentralised social media, however, may return control to the user. Email addresses, phone numbers, and other forms of personal identification are frequently linked to social media platforms. However, social media marketers that rely on leveraging consumer data to create personalised content must have access to user information. Making content for bogus bots and false views on social media platforms is thought to be costing social media marketers up to $16.4 billion annually. Blockchain is a database that is verified by many different people as opposed to a single central authority. The technology can give people back control over their personal data.

LITERATURE REVIEW

Social media is a platform that focuses on collaboration, sharing of material, engagement, and community-based feedback. The loss of personal privacy, data protection, and ownership of information are the three fundamental negatives of social media. Decentralised social media, which uses a Blockchain-based social media platform to facilitate social networking, content sharing, and even blogging with end-to-end encryption, is a solution to this challenge.

By using a decentralised method to online social networking Ching-man Au Yeung, Ilaria Liccardi, Kanghao Lu, Oshani Seneviratne, Tim Berners-Lee [1] demonstrated how the aforementioned issues might be resolved. Users are not restricted by a certain social networking service using this method. When compared to many of the prominent social networking sites available today, this can offer an equal or even greater amount of user involvement. Additionally, it gives individuals more authority over their own data. On open technologies like Linked Data, Semantic Web ontologies, open single-sign on identity systems, and access control, a decentralised social networking infrastructure is outlined. The decentralised framework may be distributed and expandable because to the widespread usage of Uniform Resource Identifiers (URIs) as identifiers since people, apps, and data can all be connected to by using their URIs.

Online social networks (OSNs) are growing more and more common in people's lives, yet because of their centralised data management system, they have a privacy leakage issue. The introduction of Distributed OSNs (DOSNs) can address this privacy concern, however they have shortcomings in terms of delivering key features like access control and data accessibility. Thomas Paul, Antonino Famulari , Thorsten Strufe [2] provided a categorization of prior work under following two major dimensions. The DOSN's three key technological concerns, decentralising content storage, authentication, and interaction/signaling, are addressed in (i) different kinds of methods with regard to resource provisioning devices and (ii) selected techniques. A novel DOSN framework that combines the benefits of both conventional centralised OSNs and DOSNs is designed using the developing blockchain approach in light of the aforementioned issues faced by OSNs and DOSNs. By combining smart contracts, the blockchain is used as a trusted server to offer central authority tasks. Due to the isolation of the storage services, users have complete authority over their data. Users have total control over their data since the storage services are isolated. Real-world data sets used in the experiment attest to the usefulness of the suggested approach.

Jiang, L., & Zhang, X. [3] presented an innovative blockchain-based frame-work to tackle the above-mentioned challenges of OSNs and DOSNs. Each registered account on the blockchain has a distinct address, which is utilised in this framework to identify each user. Additionally, user profiles are encrypted on the blockchain using a number of complex encryption techniques. The blockchain's immutability makes it possible to guarantee the originality of the data friend's view. By combining the two previously described functionalities, the framework can stop fraudulent users from forging or stealing identities In contrast to conventional servers, Jiang, L., & Zhang, X. [3] explained the separation of the regulating service from the storage service. A distributed storage service implements the storage service. The blockchain is used as a centralised server that offers practical functions for the regulating service when combined with smart contracts. One of the benefits is that a user may utilise the smart contract's event mechanism to get alerts (such post updates) on time rather of having to periodically check in with each buddy.

Due to the centralization of the majority of well-known OSNs, there are several kinds of management, security, and confidentiality problems. A decentralised architecture based on blockchain technology has the potential to address the issues described mentioned earlier. A blockchain-based OSN service is developed and its decentralised functioning is illustrated by Ningyuan Chen, David Siu-Yeung Cho [4]. The Interplanetary Filesystem (IPFS) allows for the storage of an enormous amount of data with generally low security standards in order to decentralise data. Due to a decentralised autonomous organisation that was created for user independence, users may autonomously self-manage the OSN.

Access control is typically necessary in DOSNs to protect data privacy. R. Nasim and S. Buchegger [5] created a standard access management framework based on XACML and included an appropriate authentication technique [in Security Assertion Markup Language (SAML)] in order to protect user privacy. Additionally, De Salve, P. Mori, and L. Ricci [6] suggested using XACML to define fine-grained privacy restrictions. An innovative user-to-user relationship-based control (UURAC) architecture that enables users to specify more complex and fine-grained access control policies was presented by Y. Cheng, J. Park, and R. Sandhu [7]. More works, however, O. Bodriagov, G. Kreitz, and S. Buchegger [8] were created with encryption-based access restriction. Predicate encryption (PE) was first presented by O. Bodriagov, G. Kreitz, and S. Buchegger [8] and modified for the DOSN environment.

Loopholes of existing methods

Due to privacy and security issues, traditional social media platforms have always been in the spotlight. Social media has undoubtedly transformed how people communicate. Nevertheless, they have total access over user data, and as a result, they may manage the servers that are held by any company that runs centralised social networking platforms that track user behaviour. In addition, these platforms sell the combined information to the highest bidder in order to give them better targeting for advertising and marketing. All platforms, including Facebook, Whatsapp, Twitter, and others, are centralised, so they may easily abuse the content of users. Few loopholes of traditional methods are as given below:

- Users—not clients—are the product.

Centrally managed social media is a poor strategy for fostering unity across races. While Facebook may link individuals more effectively than any other site, it comes at a price: it collects vast amounts of participant data and utilizes that data to market to specific audiences. Advertisers include businesses who want us to buy their products and services as well as individuals who try to get our decisions by influencing our perspectives on an opponent in politics or social status issue. Users' data is an outcome of centralised social media platforms like Facebook, Instagram, LinkedIn, Twitter, and YouTube, where marketers are the clients.

- Feeds that are selected by an algorithm may be more biased and polarizing

78% of social media users who are engaged have little trouble believing what they read online. Depending on the client's prior choices and behaviours, the data given might occasionally be biased or polarized. Users may take very definite actions when they frequently encounter stuff that gives them comfort. Furthermore, if a well-known individual or company spreads this knowledge, it may have a cascading influence on society.

- Centralized systems are simple to breach

Anyone who uses centralized social networks like Facebook, Twitter, Instagram, and Snapchat has significant concerns about user privacy. Their rules let them to access, retrieve, and store user data on an individual basis on a centralized server, increasing the potential hazards and implications of a data breach.

PROPOSED ARCHITECTURE

The proposed architecture allows users to post their creative stuff just like other social sites. Fig. 1 depicts the block diagram of proposed architecture. The portal is accessible via the website and will enable login following credential authentication.

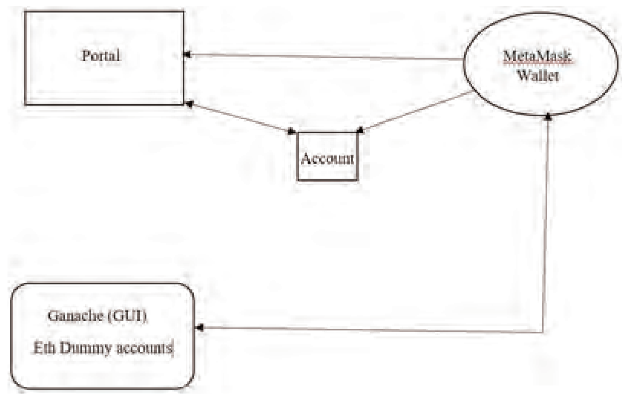


Fig. 1 Block diagram of Proposed Architecture

A well-known and well-respected browser plugin called MetaMask serves as an Ethereum blockchain-connected coin wallet. The user account is connected to MetaMask. The Ganache application provides dummy accounts to the decentralized social networking application to carry out Ethereum transactions.

A well-known and well-respected browser plugin called MetaMask serves as an Ethereum blockchain-connected coin wallet. The MetaMask user account is linked to it. The decentralised social networking application uses the Ganache programme's dummy accounts to conduct Ethereum transactions.

Blockchain social media offers the following three features:

A Check on user commodification

Data is one of the most lucrative assets nowadays. Data controllers have the power to influence a data producer's thoughts, actions, and other qualities. Furthermore, centralised social media platforms are treasure troves for user information. The fact that companies like Facebook and Instagram, among others, exploit user data to enable customised advertising for its customers, is an open secret.

These risks of user commodification are eliminated by distributed social media platforms' peer-to-peer network, end-to-end encryption, and lack of "third parties". Additionally, by extension, it gives consumers more control over their data. The usage of Private and Public keys is the primary method for ensuring this.

The user can start a transaction using the private key, and the specifics of that transaction will be recorded in a data block on the blockchain. Only when this block has been approved by the majority of the other network participants it is included to the chain. The intended receiver is the only one who can access the data as a result of the transaction. These blocks include data in the form of a cryptographic "hash." These are one-way techniques, therefore a hash can't (theoretically) be decoded to reveal the original data.

Privacy of freedom and expression

Those who are in charge of managing centralised social media have access to every contact, even private communications. This poses an enormous risk to the users' privacy, among other things. Additionally, social media posts are increasingly being utilised to stifle opposition and advance political goals. Understandably, this poses a serious threat to the democratic values of expression and free speech. Blockchain social media platforms guarantee better confidentiality and support the users' right to free expression thanks to its decentralised consensus methods. However, a lack of constraints like that can bring about fresher dangers. Such problems must be resolved over time if new decentralised platforms are to equivalent the appeal of the current social media juggernauts.

Scope for crowd funding

Blockchain social media also provides tangible incentives in addition to non-financial ones. The majority of decentralised social networks compensate their users (nodes) for uploading content there. Such benefits occur in the form of the network's native cryptocurrency, which is a feature shared by the majority of blockchain ecosystems. Additionally, these social media platforms give users an opportunity to run successful crowd funding projects using token sale-like procedures. The native cryptocurrency is also used by them to enable user-to-user financial transactions.

Fig. 2 shows the workflow of proposed system. It comprises five major components. These are React, Ganache, Truffle, Decentralized Application (DAPP) center and InterPlanetary File System (IPFS).

All content producers have access to a front end React js web client to upload their work. For the effective production, tipping, and updating of the material, smart contracts are designed using Solidity, and this smart contract is assembled using Truffle.

The object-oriented high-level language Solidity may be used to build smart contracts. The behaviour of accounts in the Ethereum state is managed by programmes referred to as smart contracts. Curly brackets are the language of solidity. It is designed for the Ethereum virtual machine (EVM) and has JavaScript, C++, and Python features. It facilitates inheritance, libraries, and advanced user-defined types, among other things, and is statically typed. For tasks like voting, crowdsourcing, blind bidding, and multi-signature wallets, Solidity allows you to create contracts. The injected wallet (MetaMask) is connected to the front-end source using Web3.js. Accounts are imported from Ganache into the MetaMask wallet.

Ganache provides dummy Ethereum accounts. A private blockchain called Ganache makes it simple to build Corda and Ethereum distributed apps. You may create, deploy, and test your decentralised apps (dApps) with Ganache during the whole development cycle in an environment that is secure and monitored. A User Interface (UI) and command-line interface (CLI) are the two flavours of ganache. CLI is a command-line interface that provides users with more system control. A desktop programme called Ganache UI supports both Corda and Ethereum. For Ethereum development, ganache, originally known as the TestRPC, is a command-line tool. For Windows, Mac, and Linux, there are all available versions of Ganache.

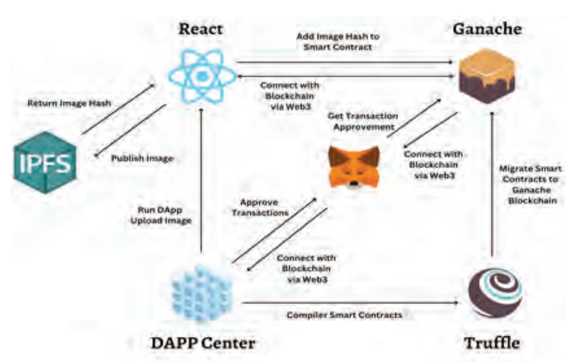


Fig. 2 Workflow of Proposed System

Ethereum TestRPC is a quick and configurable blockchain emulator. It enables calling the blockchain without the costs associated with operating a real Ethereum node. With a fixed quantity of Ether, accounts may be recycled, reset, and instantiated (mining or faucets are not required).

A top-notch programming environment, testing framework, and asset pipeline for blockchains built on the Ethereum Virtual Machine (EVM) that promises to render work easier for developers. Benefits of truffle are as follows:

- built-in binary management, linking deployment, and smart contract compilation.
- automated contract testing for quick development.
- a framework for deployment and migration that is scriptable and extensible.
- network administration for several private and public networks.
- management of packages using the erc190 standard with the ethernet package manager (EthPM) and NPM.
- Direct contract communication through an interactive interface.
- A build process that is configurable and supports tight integration.
- External script runner for Truffle setups that runs scripts.

All of the content's specifics have been kept in a distributed storage system utilizing InterPlanetary File System (IPFS) .Using SHA 256 encryption, IPFS transforms the material into hash codes that are then added to the block design. After publishing and tipping material, the user's account will be updated. According to tipped Ethereum, content would be arranged. Files, websites, programmes, and data are stored and accessed via a distributed system known as IPFS. That sweet, sweet aardvark data may be located using IPFS based on its contents rather than its physical location. Instead of accessing one of Wikipedia's computers for the page, your computer utilises IPFS to request the page from a number of machines across the world. The string of numbers in the middle of the URL (QmXo...) denotes the IPFSified version of the aardvark information. It is not limited to using Wikipedia to find facts on aardvarks. IPFS enables this for all types of computer files, including documents, emails, and even database entries, in addition to web pages.

Fig. 3 Screenshot 1 of Uploading Image

Fig. 4 Screenshot 2 of Uploading Image

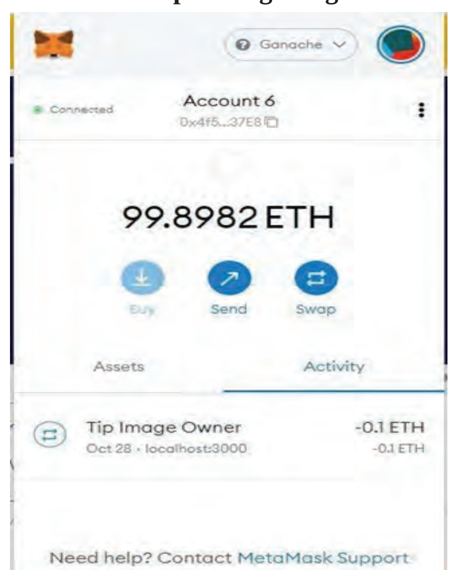


Fig. 5 MetaMask wallet

An Ethereum blockchain-connected coins wallet called MetaMask is a well-known and respected browser plugin. MetaMask lets users to interact with the Ethereum ecosystem, which offers a wide variety of decentralised applications (Dapps), without having to download the full blockchain onto their device. It is therefore one of the best Ethereum wallet choices for rapid access to decentralised exchanges (DEX), gaming software, gambling websites, and many other applications. The wallet works with all of the most popular browsers, including Chrome, Firefox, Brave, and Microsoft Edge. The native currency of

Ethereum, ETH, as well as tokens made in compliance with the ERC-20 and ERC721 standards are both kept by MetaMask. Fig. 3 and Fig. 4 are the screenshots of uploading image. Fig. 5 and Fig. 6 show MetaMask wallet and Ganache respectively.

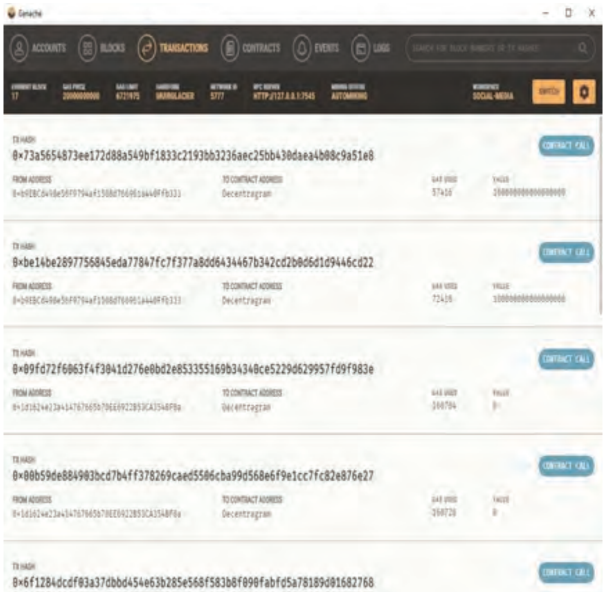


Fig. 6 Ganache

CONCLUSION

A new era of social networks is being launched in by blockchain technology. With the help of this new technology, current social media models may now have the highest level of privacy. Data sales, algorithm modifications, and useless data constraints could all become obsolete. It's because these social networks using blockchain technology are run and controlled by the users themselves. Because there is no centralised authority, users on these networks enjoy more consequences for their opinions on social media. On the majority of decentralised social media sites, individuals may privacy, which supports the right to free speech and expression and spares people the suffering of facing legal earn cryptocurrency by publishing and chatting. Messaging that is directly connected to the crypto address (domain name, for example, unstoppable. crypto) is made possible using a decentralised messenger known as the Dchat protocol. The features like as team conversations, auto-destruct features, stickers applications, games, etc. may be separately developed and released on top of any wallet or blockchain application. On a P2P storage network, it will be constructed.

ACKNOWLEDGMENT

Special thanks to the Principal Dr. S. M. Khot, FCRIT, Vashi, Navi Mumbai and R & D incharge Dr. Sushil Thale, FCRIT, Vashi, Navi Mumbai for their constant support, motivation and timely guidance.

REFERENCES

1. Ching-man Au Yeung , Ilaria Liccardi, Kanghao Lu ,Oshani Seneviratne , Tim Berners-Lee , " Decentralization: The future of online social networking," 2008. https://www.w3.org/2008/09/msnws/papers/decentralization.pdf
2. Thomas Paul, Antonino Famulari , Thorsten Strufe, " A survey on decentralized Online Social Networks Elsevier Computer Networks," Volume 75, Part A, 24 December 2014. https://www.sciencedirect.com/science/article/abs/pii/S1389128614003600
3. Jiang, L., & Zhang, X., "BCOSN: A blockchain-based decentralized online social network," IEEE Transactions on Computational Social Systems, 6(6), 1454–1466 2019. https://doi.org/10.1109/tcss.2019.2941650
4. Ningyuan Chen; David Siu-Yeung Cho,"A Blockchain based autonomous decentralized online social network," 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE) . https://ieeexplore.ieee.org/document/9342564
5. R. Nasim and S. Buchegger, "XACML-based access control for decen- tralized online social networks," in Proc. IEEE Int. Conf. Utility Cloud Comput., Dec. 2014, pp. 671–676. https://www.academia.edu/56656173/XACML_Based_Access_Control_for_Decentralized_Online_Social_Networks
6. A. De Salve, P. Mori, and L. Ricci, "A privacy-aware framework for decentralized online social networks," in Proc. Int. Conf. Database Expert Syst. Appl. Cham, Switzerland: Springer, 2015, pp. 479–490. https://core.ac.uk/download/pdf/37831573.pdf
7. Y. Cheng, J. Park, and R. Sandhu, "An access control model for online social networks using user-to-user relationships," IEEE Trans. Dependable Secure Comput., vol. 13, no. 4, pp. 424–436, Jul./Aug. 2016. https://dl.acm.org/doi/10.1109/TDSC.2015.2406705
8. O. Bodriagov, G. Kreitz, and S. Buchegger, "Access control in decen-tralized online social networks: Applying a policy-hiding cryptographic scheme and evaluating its performance," in Proc. IEEE Int. Conf. Pervas. Comput. Commun. Workshops, 2014, pp. 622–628. https://www.researchgate.net/publication/271419767_Access_control_in_decentralized_online_social_networks_Applying_a_policy-hiding_cryptographic_scheme_and_evaluating_its_performance

Enhancing Image Classification using Transfer Learning with AlexNet and Ensemble Model

Juhi Janjua
Computer Engineering Department
Thadomal Shahani Engineering College
Mumbai University
Mumbai, Maharashtra
✉ ragini.sharma@it.sce.edu.in

Nabanita Mandal
Computer Engineering Department
Thadomal Shahani Engineering College
Mumbai University
Mumbai, Maharashtra
✉ nabanita.mandal@thadomal.org

Archana Patankar
Computer Engineering Department
Thadomal Shahani Engineering College
Mumbai University
Mumbai, Maharashtra
✉ archana.patankar@thadomal.org

Darkhshan Khan
Computer Engineering Department
Thadomal Shahani Engineering College
Mumbai University
Mumbai, Maharashtra
✉ darakshan.khan@thadomal.org

ABSTRACT

Image classification is a fundamental task in computer vision with broad applications. Accurate image classification is essential to the development of many industries, including manufacturing, agriculture, autonomous systems, healthcare, and more. In the realm of machine learning, ensemble methods have garnered attention for their capability to boost the efficacy of model predictions. This is a research paper exploring the head-to-head performance of the stacking ensemble technique and the standalone AlexNet architecture in image classification using the Fashion- MNIST dataset. Preliminary results are astonishing and reveal that the ensemble method can outperform even when trained on a fraction of the epochs set for AlexNet. The results highlight the resilience of ensemble approaches, especially stacking, which opens a great deal of scope for exploiting the combined strength of different architectures. This work, apart from illuminating the subtle benefits of ensemble methods, paves the way for further research aimed at optimizing and exploiting these methods for complex classification problems.

KEYWORDS *: Image classification, AlexNet, Ensemble model, Fashion-MNIST.*

INTRODUCTION

With the advent of big data, digital imagery has revolutionized a new wave of possibilities across healthcare, autonomous vehicles, entertainment, and security. But this vast abundance of visual data brings along with it the difficult task of categorizing and extracting meaningful information from these enormous collections of images. This is the main problem of the image classification domain in the computer vision field.

Image classification is training computers in the recognition and further categorization of images through their content. It allows algorithms to learn which objects, patterns, and features inside an image are related to which given categories. This technology is important in a wide array of applications including facial recognition, object detection, and medical diagnosis. It equips machines with the capacity to read visual data to help in analytics and streamline multiple tasks across any industry. So far, challenges in this sector are developing an algorithm that, on its own, can make sense of these complex patterns as well as characteristic features in any image, which goes on to lead to consistent and accurate identification results across a varied set of images.

In recent times, researchers have used many strong techniques that include transfer learning and ensemble methods to simplify and later resolve the complexity of image classification. Transfer learning makes it possible for images and patterns learned from extensive data sets to head into pre-trained neural network models. Instead of building a model entirely, transfer learning fine-tunes the pre-built models for different tasks in an image classification model by using this pre-existing knowledge. This improves model development acceleration and reduces requirements on data storage; more so, it will have higher accuracy while classifying these images, so this is actually an effective way across all disciplines [1]. The ensemble is designed to maximize model performance by pooling together the individual predictions from other models. They integrate multiple individual models, hence combining their predictive insights. Often more accurate and robust than reliance on a single model, aggregation of the output of diverse models through ensemble methods reduces overfitting, improves generalization, and increases the overall classification accuracy [2]. The paper seeks to explore ensemble techniques applied using pre-trained models with potential in the improvement of performance in image classification.

The primary objective of this research paper is to perform an in-depth comparative analysis of the methods for image classification by making use of transfer learning with AlexNet and ensemble methods by employing ResNet, VGG, and MobileNet architectures. This will be beneficial in determining the strengths and weaknesses of the AlexNet transfer learning model as well as the ensemble techniques and can guide practitioners to select the best possible approach to tackle specific challenges of image classification.

The manuscript is ordered as follows. Section 2 illustrates the exhaustive literature review followed by methodology in section 3. Section 4 depicts the experimental work which is compared with machine learning models in section 5. At the end section 6 concludes the paper.

LITERATURE REVIEW

Image classification is one area of significant advancements and especially by means of combining deep learning with the concept of transfer learning. For example, this has been applied using the pre-trained VGG16 model for daisy and dandelion classification from images. The network described had an input shape of (7,7,512), ReLU activation, and 0.5 dropout and showed a respectable accuracy of 90% after 100 epochs [3]. It highlights the capability of transfer learning and pre-trained models in the image classification task.

One study explored the performance of eleven pre-trained models on five different datasets. Surprisingly, the ResNet-50 model was particularly impressive, averaging 92.5% accuracy across the datasets [4]. Such results have lent credibility to transfer learning, especially when the training data are scarce. With such architectures as VGG-16, ResNet-50, and DenseNet-121, it identified the former as a better model using a mAP score of 0.9629 [5]. This shows potential in environmental safety while highlighting a large role in better classification of multilabels between transfer learning and data augmentation.

Medical image classification has not been an exception to these advancements. A recent paper on how ensemble learning can optimize deep CNNs for medical images demonstrated that using techniques such as stacking augmented performance resulted in as high as a 13% F1-score lift. In addition, augmenting and cross-validation based bagging helped immensely in improvements to performance [6]. In a similar direction, a weighted ensemble was proposed for the task of image classification. The same model, which was trained using an Adam optimizer, was promisingly performing on X-ray image datasets, even surpassing other established models from previous studies [7]. The method of assigning weights based on individual model accuracies and deriving predictions from a weighted average of all models' predictions is proving to be a potent approach.

One outstanding ensemble learning-based technique was contributed to the analysis for multi-spectral image classification. Ensemble-based learning was done for the image data classification which belonged to multiple spectra. On using this strategy, it results in the phenomenal outperformance and shows a truly excellent classification of 99.96% compared to the overall strong base classification while conducting any kind of multi-spectral analysis of land use and land cover. This high level of accuracy underscores the potential of ensemble learning methods in handling complex classification tasks within multi-spectral image datasets and solidifies its importance in advancing land use and environmental monitoring applications [8].

AlexNet was one of the pioneering architectures for deep learning. Since it came in 2012 and bagged first place at ILSVRC ImageNet Large Scale Visual Recognition Challenge, various other architectures were proposed based on the performance on deep neural network. It proved a tremendous capability by delivering extremely accurate image classification capabilities as well as proving deep learning capabilities are indeed capable of superseding the old traditional computer vision methods.

Moreover, its architecture is characterized by using multiple layers that enable hierarchical feature learning. Its design further makes it exceptional because it uses different layers that allow for hierarchical feature learning. From raw pixel values, AlexNet could analyze images and get complex patterns and representations [9]. However, it is not exempted from flaws.

The depth and complexity in AlexNet are computationally intensive, requiring high resource and memory. It also follows the multi-layer architecture that enables it to learn hierarchical features. AlexNet analyzes raw pixel values, capable of finding complicated patterns and representations in images [9]. However, these come with disadvantages. The complexity and depth of AlexNet, though, are computationally demanding, taking enormous powers and memory resources. It is therefore often not suitable for devices under considerable constraint or in real-time applications. Moreover, it has a high parameter count that leads to overfitting especially if trained on smaller sets of data; thus, some regularization techniques need to be employed [10].

Ensemble methods are touted for their higher robustness and generalization capabilities. The beauty of ensembling is that predictions from a plethora of models can be combined to reduce overfitting and facilitate more reliable outcomes. One peculiar characteristic of an ensemble is the ability to represent a wide range of base models with different architectures, thereby making the predictions richer and the overall performance of the ensemble stronger [11]. Perhaps one of the shining methods in the ensemble techniques' arsenal is stacking, which makes use of meta-learning. It provides models with a way to take insights from what other models have predicted, packaging complex relationships in outputs from the base models into something that works better than conventional ensemble methods many times. This will harmonize what the base model does well, while diminishing their limitations, therefore generalizing to be better than the rest [8]. However, ensemble methods also have their share of problems. Their design and maintenance tend to add a heightened complexity over the process of training a single model. They also become computationally intensive and time- consuming, which might delay their usage in real-time applications. Besides, stacking introduces additional layers of complexity, particularly at the architecture of the models and the tuning of hyperparameters. The identification of the appropriate ensemble of base models and the meta-learner might be challenging as well [12]. Finally, data requirements are rather high when it comes to stacking. While both the ensemble models and the meta-learner require a pretty large dataset in order to successfully train, scarce data may let the meta-learner overfit in sparse cases.

METHODOLOGY

This paper proposes two distinct yet powerful approaches to tackle the challenge of classifying Fashion MNIST images: one based on transfer learning using AlexNet and the other harnessing the combined strength of ensemble learning, integrating ResNet50, VGG16, and MobileNet.

Both methodologies are rooted in the philosophy of utilizing the strengths of pre-trained models, either by transferring their learned features or by aggregating their capabilities. In the subsequent sections, we delve deeper into the intricacies of both approaches, elucidating the architecture details & training procedures.

AlexNet is a notable convolutional neural network architecture consisting of five convolutional layers followed by three fully connected layers.

The architecture begins with an input layer for a 224 x 224 x 3 image, followed by multiple convolutional layers: the first with 96 filters of 11 x 11 kernel size, the second with 256 filters of 5 x 5 kernel size, and the subsequent layers with smaller filters and kernel sizes [10] as shown in Fig. 1.

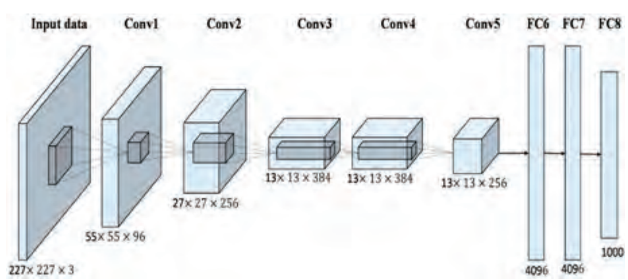


Fig. 1. AlexNet Architecture

To decrease the number of features max-pooling layers follow the first, second, and fifth convolutional layers. For training AlexNet on the Fashion MNIST dataset, we used a total of 25 epochs. The convolutional layers are followed by three fully connected layers, the last of which connects to the output layer, employing a softmax function for multi-class classification. The use of ReLU activations, dropout before the first two fully connected layers, overlapping max- pooling, and data augmentation and normalization are additional prominent features of AlexNet.

The Fashion MNIST dataset, another critical element in the machine learning field, comprises 70,000 28x28 pixel grayscale images of fashion items, split into 60,000 training and 10,000 testing images. Each image is labeled with one of 10 class labels, denoting different types of clothing such as T- shirts, trousers, sandals, etc. Developed as a more sophisticated alternative to the classic MNIST dataset of handwritten digits, Fashion MNIST challenges machine learning algorithms with more intricate patterns associated with clothing items [10].

To harness the pre-trained AlexNet for the Fashion MNIST dataset from its original architecture, we implement a series of modifications to the network, embodying the essence of transfer learning. Transfer learning involves using a pre- trained model, like AlexNet in this case, and fine-tuning it for a new task, such as classifying images in the Fashion MNIST dataset. The following adjustments are done on the AlexNet model for transfer learning.

Input Layer Adjustment

The Fashion MNIST dataset is characterized by grayscale images with a dimension of (28, 28, 1).

AlexNet, originally configured for (227, 227, 3) images, is adapted to work with (28, 28, 1) images, averting the loss of fine details and eliminating the need to resize the dataset images to 227*227 [10]. This modification allows us to keep the input layer intact, demonstrating the flexibility and adaptability of transfer learning.

Output Layer Adjustment

In creating AlexNet compatible with the 10 distinct categories of the Fashion MNIST dataset, the output layer of AlexNet, originally tuned for 1000 categories of the ImageNet dataset, is replaced with a new fully connected layer with 10 nodes [10]. With a softmax activation function, this conversion is crucial in producing class probabilities to accurately classify fashion articles and tailor fit AlexNet into the specific needs and categorizations of the Fashion MNIST dataset.

The fully connected layers of AlexNet are replaced, while the convolutional layers are retained and adjusted to match the input shape and number of classes of the Fashion MNIST dataset. This is one of the fundamental aspects of training in transfer learning, where a few layers are frozen to retain their pre-trained weights, thereby helping in good feature extraction. We intentionally freeze the early layers that capture lower-semantic features to retain their pre-learned capabilities. Concurrently, we fine-tune some later layers of the architecture, allowing certain pre-trained layers' weights to be updated during the training. This balanced strategy enables the model to hold onto the generic lower-level features while adeptly modifying the higher-level features for the specific task of Fashion MNIST classification.

The second approach adopts an ensemble learning approach. In the realm of ensemble learning, delving into the architectures of ResNet50, VGG16, and MobileNet provides a comprehensive view. The ensemble approach for classifying Fashion MNIST images was trained over 3 epochs. ResNet, or Residual Networks, uses skip or shortcut connections to ensure efficient training of deeper networks by allowing gradients to bypass certain layers. This architecture aids in mitigating the vanishing gradient problem, enhancing the learning capability of the network [13]. VGG or Visual Geometry Group, another deep architecture, employs small receptive fields and multiple layers to extract intricate features from the input data. The depth of the network allows it to capture a diverse range of features, enhancing its classification ability [14]. MobileNet, tailored for mobile and edge devices, emphasizes computational efficiency. It employs depthwise separable convolutions to minimize the computational load while maintaining robust performance [15].

For Fashion MNIST dataset, the preprocessing steps include normalization of pixel values and reshaping images before feeding it to the networks. Preprocessing makes sure that the dataset fits the model architecture and improves the performance of the models by giving standardized input data.

One core step in developing a more efficient machine learning approach lies in combining the use of various models into addressing one specific problem: ensemble methods. The possibility of deriving predictions that can outperform even those made with one single, highly potent model. Stacking or stacked generalization diverges by employing a more layered approach. The construction of a stacking ensemble begins with the selection and training of multiple diverse base models on the complete dataset. These models, each with its own predictive power, output predictions on a validation or test dataset.

Finally, the performance of the stacking ensemble is carefully tested and measured over a test dataset with multiple measures to test its performance in prediction [17].

Such a detailed, painstaking, multi-layered approach in stacking ensemble constitutes a considerable stride towards attaining improved performance of a model as depicted in Fig. 2.

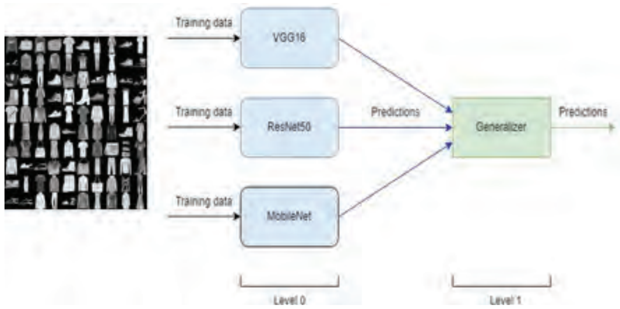


Fig. 2. Architecture of Ensemble Model

For unbiased effectiveness estimation, the model is tested against unseen data. The assessment is usually performed in terms of accuracy, precision, recall, and F1-score along with other tools, such as a confusion matrix [18]. A more extensive analysis through such methods provides refinement and thus, suitability to actual real-world application.

EXPERIMENTAL RESULTS

Model evaluation metrics highlight the performance of image classification models. In this research four performance metrics are used i.e., accuracy, precision, recall and F1-score. Accuracy measures how often model, like AlexNet or the ensemble of ResNet, VGG, and MobileNet, correctly classify fashion items in the Fashion MNIST dataset [1] as shown in equation 1.

$$Accuracy = \frac{Count\ of\ true\ estimate}{Count\ of\ total\ estimates}, \quad (1)$$

Equation 2 indicates precision (P), which evaluates the model's reliability, while equation 3 indicates recall, R which assesses their completeness in class identification [18].

$$P = \frac{Count\ of\ true\ positive\ estimates}{Count\ of\ total\ positive\ estimates}, \quad (2)$$

$$R = \frac{Count\ of\ true\ positive\ estimates}{Count\ of\ (true\ positive\ estimates + false\ estimates)}, \quad (3)$$

The F1-score balances precision and recall, highlighting strengths and weaknesses [19] as shown in equation 4.

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}, \quad (4)$$

The research provides a nuanced evaluation of AlexNet and ensemble models, highlighting their strengths and vulnerabilities in fashion image classification, using diverse metrics for a rigorous comparative analysis.

The comparative results shown in Table. I between the ensemble approach and AlexNet provide intriguing insights into their performance dynamics. Despite being trained for only 3 epochs, the ensemble method achieves an accuracy of 0.905, outpacing the AlexNet's performance of 0.877, which was trained for a considerably longer 25 epochs. This discrepancy is even more pronounced when considering the precision, recall, and F1-score metrics, where the ensemble consistently surpasses AlexNet across the board.

One could surmise that the inherent strength of the ensemble approach, which combines the predictive power of multiple architectures, allows it to rapidly converge to a more accurate solution, even with a relatively limited number of iterations.

Table 1. Evaluation of Ensemble & AlexNet model

Parameter	Ensemble (epoch = 3)	AlexNet (epoch = 25)
Accuracy	0.905	0.877
Precision	0.909	0.897
Recall	0.905	0.877
F1-score	0.904	0.886

It's particularly noteworthy that, with such a constrained number of epochs, the ensemble method not only matches but exceeds the performance of AlexNet as shown in Fig. 3.

Projecting these results into future experiments, it's reasonable to hypothesize that increasing the number of epochs for the ensemble method could further amplify its accuracy and other metric scores. If the ensemble's performance can be this robust with a mere three epochs, it's tantalizing to consider the potential advancements in accuracy, precision, recall, and F1- score that could be achieved with more extensive training. This observation underscores the potential efficacy of ensemble methods in image classification tasks, especially when juxtaposed against single, albeit powerful, architectures like AlexNet.

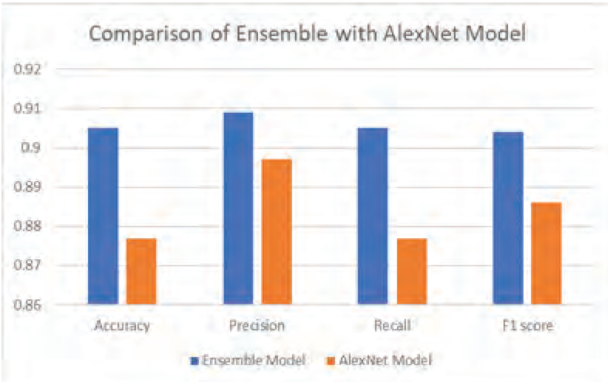


Fig. 3. Comparison of Ensemble & AlexNet model

V. COMPARISON WITH TRADITIONAL MACHINE LEARNING MODEL

In our study using the Fashion-MNIST dataset, we evaluated the performance of various machine learning models such as Decision Tree, K-Nearest Neighbors (KNN), Random Forest, Feedforward Neural Network (FNN). These results are as shown in Table. II.

All models delivered strong results, with Random Forest and FNN nearly matching the Ensemble Model's performance. Surprisingly, the basic FNN achieved remarkable results, challenging the notion that complex models are always necessary. Intriguingly, FNN's precision, recall, and F1 score were quite like those of the Ensemble Model as shown in Fig. 4.

This suggests that the FNN, with a straightforward architecture, can also offer competitive results. Furthermore, we explored the potential for improving FNN performance by deepening its architecture.

Table 2. Traditional Machine Learning Models comparison

Parameter	Decision Tree	KNN	Random Forest	FNN	Ensemble Model
Accuracy	0.788	0.854	0.878	0.860	0.905
Precision	0.790	0.857	0.877	0.898	0.909
Recall	0.788	0.854	0.878	0.860	0.905
F1-score	0.789	0.853	0.877	0.878	0.904

Deep learning models have shown the capacity to capture intricate patterns, and enhancing the FNN with more layers could potentially surpass the Ensemble Model. Increasing the number of layers of a network can potentially be computationally less expensive than an ensemble model with different individual models.

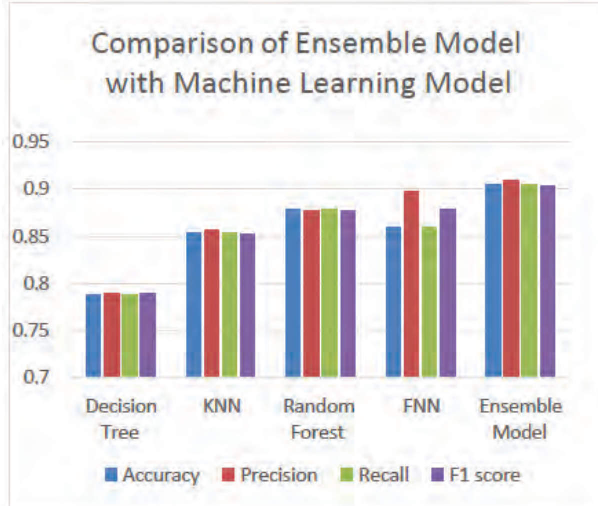


Fig. 4. Traditional Machine Learning Models Comparison

CONCLUSION

In the realm of computer vision, image classification is a fundamental activity that is essential to the resolution of many real-world issues. The research discussed here sheds light on the power of image classification, exemplified through a comprehensive study comparing ensemble methods with individual pre-trained model architectures like AlexNet.

In the exploration of machine learning ensemble methods, the stacking technique emerged as a potent strategy for enhancing model performance. Through a detailed study comparing an ensemble method and AlexNet on the Fashion MNIST dataset, the ensemble, even with a mere 3 epochs of training, outpaced AlexNet, which underwent a much lengthier

25 epochs of training. This superiority underscores the potential of leveraging the combined strengths of various architectures in ensemble methods. The observed efficiency of the ensemble suggests further research avenues, such as extended training or the integration of a broader range of base models, could yield even more refined performance in image classification tasks.

Beyond the comparison of transfer learning methods, our exploration extended to traditional machine learning models on the Fashion-MNIST dataset. Notably, Random Forest and Feed forward Neural Network (FNN) displayed performance comparable to the ensemble model. Particularly, the basic FNN architecture showed remarkable potential in image classification. It was intriguing to see that adding more layers in FNN led to a deep architecture that appeared to outperform the ensemble model and at the same time, reduce computational cost.

In conclusion, the research conducted on ensemble methods opens avenues to further deepen the traditional machine learning models and to open possible future advancements of image classification.

REFERENCES

1. J. Janjua and A. Patankar, "Framework of CNN Architecture for Fashion Image Classification," Applied Computational Technologies, pp. 96–103, 2022. doi:10.1007/978-981-19-2719-5_9.
2. A. Mohammed and R. Kora, "A comprehensive review on Ensemble Deep Learning: Opportunities and challenges," Journal of King Saud University - Computer and Information Sciences, vol. 35, no. 2, pp. 757–774, 2023. doi:10.1016/j.jksuci.2023.01.014.
3. C. Desai, "Image classification using transfer learning and Deep Learning," International Journal of Engineering and Computer Science, vol. 10, no. 9, pp. 25394–25398, 2021. doi:10.18535/ijecs/v10i9.4622
4. N. Abou Baker, N. Zengeler, and U. Handmann, "A transfer learning evaluation of deep neural networks for Image Classification," Machine Learning and Knowledge Extraction, vol. 4, no. 1, pp. 22–41, 2022. doi:10.3390/make4010002.
5. M. Park, D. Q. Tran, S. Lee, and S. Park, "Multilabel image classification with Deep Transfer Learning for decision support on wildfire response," Remote Sensing, vol. 13, no. 19, p. 3985, 2021. doi:10.3390/rs13193985.
6. D. Muller, I. Soto-Rey, and F. Kramer, "An analysis on ensemble learning optimized medical image classification with deep convolutional Neural Networks," IEEE Access, vol. 10, pp. 66467– 66480, Oct. 2021. doi:10.1109/access.2022.3182399.
7. T. Iqball and M. A. Wani, "Weighted Ensemble model for Image Classification," International Journal of Information Technology, vol. 15, no. 2, pp. 557–564, 2023. doi:10.1007/s41870-022-01149-8
8. T. Aboneh, A. Rorissa, and R. Srinivasagan, "Stacking-based ensemble learning method for multi-spectral image classification," Technologies, vol. 10, no. 1, p. 17, 2022. doi:10.3390/technologies10010017.
9. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional Neural Networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017. doi:10.1145/3065386.
10. Y. Xue, "An Overview of Overfitting and its Solutions," in *Journal of Physics: Conference Series*, vol. 1168, p. 022022, 2019, doi: 10.1088/1742-6596/1168/2/022022.
11. M. Alhamid, "Ensemble models: What are they and when should you use them?," Built In, https://builtin.com/machine-learning/ensemble- model (accessed Sep. 14, 2023).
12. C. Makhijani, "Advanced Ensemble Learning Techniques," Medium, https://towardsdatascience.com/advanced-ensemble-learning- techniques-bf755e38cbfb (accessed Oct. 15, 2023).
13. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Dec. 2015. doi:10.1109/cvpr.2016.90.
14. K Simonyan, A Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", 2015 The International Conference on Learning Representations (ICLR), Sep. 2014., arXiv:1409.1556
15. A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," 2017. [Online]. Available: arXiv:1704.04861. (accessed: 2-September-2023) doi:10.48550/arXiv.1704.04861.
16. B. Soni, "Stacking to improve model performance: A comprehensive guide on ensemble learning in python," Medium, https://medium.com/@brijesh_soni/stacking-to-improve-model-performance-a-comprehensive-guide-on-ensemble-learning-in-python- 9ed53c93ce28 (accessed Nov. 1, 2023).
17. A. Singh, "A comprehensive guide to ensemble learning (with python codes)," Analytics Vidhya, https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide- for-ensemble-models/ (accessed Oct. 16, 2023).
18. S. Hussain, J. Keung, A. A. Khan, and K. E. Bennin, "Performance evaluation of ensemble methods for software fault prediction," Proceedings of the ASWEC 2015 24th Australasian Software Engineering Conference, 2015. doi:10.1145/2811681.2811699.
19. A. Mohammed and R. Kora, "A comprehensive review on Ensemble Deep Learning: Opportunities and challenges," Journal of King Saud University - Computer and Information Sciences, vol. 35, no. 2, pp. 757–774, Feb. 2023. doi:10.1016/j.jksuci.2023.01.014.

Content-Based Recommendation Model for Intelligent Resume Screening

Manimala Mahato
Associate Computer Engineering
Shah and Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ manimala.mahato@sakec.ac.in

Aagam Sheth
Associate Computer Engineering
Shah and Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ aagam.sheth_19@sakec.ac.in

Shubham Shah
Associate Computer Engineering
Shah and Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ shubham.shah_19@sakec.ac.in

Priyanshi Doshi
Associate Computer Engineering
Shah and Anchor Kutchhi Engineering College
Mumbai, Maharashtra
✉ priyanshi.doshi_19@sakec.ac.in

ABSTRACT

The digital transformation of recruitment practices has created new challenges in candidate selection processes. Human resource departments face the daunting task of processing vast quantities of unstructured resume data submitted through online platforms. This research presents an innovative automated system that streamlines the candidate evaluation process by implementing a Content-based Recommendation approach utilizing Cosine similarity metrics. Our framework analyzes the correlation between candidate resumes and specific job requirements, generating quantifiable similarity scores. This automation significantly reduces the manual effort required in resume processing while enhancing the precision of candidate shortlisting. The system's effectiveness was validated through comprehensive testing using authentic recruitment data, demonstrating its practical applicability in modern hiring workflows.

KEYWORDS *: Content-based recommendation, Cosine similarity, Recruitment, Screening.*

INTRODUCTION

Talent acquisition is an important, complex, and time-consuming function within Human Resources (HR). The sheer scale of India's market is overwhelming. The job market has created a huge turnover with over a million people applying for jobs.

The most difficult part is the absence of a standard resume format and version which enables short listing the profiles you would like to have in order to apply for required roles. Domain knowledge is needed to scrutinise the resumes and ascertain how relevant and applicable a profile is for the job role.

This paper proposes and implements a model by designing a website through which company administrators can upload descriptions of their job requirements and candidates can upload their resumes. The proposed Content-Based Recommendation model computes the similarity between the job requirement with each resume using the Cosine Similarity index and generates a list of potential candidates who are closely fit for the job. This will result in cost savings in terms of money as well as time for corporate organizations.

LITERATURE SURVEY

The landscape of modern corporate recruitment presents several critical challenges in the digital age. Through our comprehensive analysis, we identified three primary obstacles facing organizations:

Challenge 1: Volume Management The expansive Indian job market generates an overwhelming number of applications for each position. Traditional manual screening methods have become inadequate for processing this massive influx of candidates, resulting in prolonged hiring cycles and increased operational costs.

Challenge 2: Resume Heterogeneity A significant hurdle emerges from the lack of standardization in resume formats. Each candidate submits documentation with unique structures and organizational patterns, compelling hiring teams to adapt their evaluation approach for each submission. This variability not only consumes excessive resources but also introduces the risk of overlooking qualified candidates during the screening process.

Challenge 3: Qualification Alignment Perhaps the most complex challenge lies in accurately mapping candidate qualifications to specific job requirements. This crucial assessment determines whether a candidate's profile aligns with the position's demands, directly impacting hiring success rates.

Our investigation into existing solutions revealed various approaches to these challenges [3]. The following analysis examines current systems and their effectiveness in addressing these recruitment obstacles.

LinkedIn

Passive Candidates will be present and they won't be actively searching for a job. Many will not be interested in finding a new job. Having so many profiles will make it difficult for job seekers to get recognized. Too many candidates and their resumes lead to overloaded information.

Naukri.com

Employer Branding is hardly visible which is a major product flaw on the portal. Only trusted and verified sources can post their jobs on our website.

Internshala

There's a Low-Earning Potential. The website will consider more job postings than internships, therefore increasing earning potential

MonsterIndia

Passive Candidates will be present and they won't be actively searching for a job. Many will not be interested in finding a new job.

Killer Launch

It is a platform that redirects potential employees to job opportunities available in the market. Unfortunately, they are a membership type of platform which requires people to pay to gain access to this information which tends to demotivate people from creating accounts.

Existing recommendation systems can be classified into two types:

Item Based

Finding a similarity between items, and recommending an item based on the interest of a similar item. For example, if an employer views one resume for a specific job, then an item-based algorithm would suggest other resumes with similar skills or fields present in the viewed resume. This is also called "Content-based filtering" [4]. A benefit of using this approach is that you do not have to rely on a lot of user data since the similarity calculation is happening at the item level [5].

Customer Based

Finding similarities between customers, and recommending an item that a similar customer has purchased/watched [7]. This is also called "Collaborative filtering". A benefit of this approach is that it works better for discovery, since items that look unrelated initially, might be liked by similar customers [6]. Take Netflix: if I watch an episode of Marvel's Agents of Shield and other people who have watched Marvel's Agents of Shield also watched The Office, I will get recommended The Office, even though they don't seem to have anything to do with each other.

Additional Comparison

The content-based recommendation structure brings various benefits exceeding traditional keyword-based and collaborative filtering systems. The platform differs from Naukri.com and MonsterIndia because it measures job-description-resume relationships using cosine similarity for a more accurate match. The matching efficiency of our method compares well with more advanced deep learning techniques while maintaining equivalent accuracy levels. The analytical trade-off between system performance and efficiency makes our model a practical solution for big-scale resume selection procedures.

METHODOLOGY

The main objective is to check if the job seeker is good enough for the required job position. Hence, the similarity between the resume and the job requirements are calculated and candidates with a higher similarity index are shortlisted as potential candidates. The preprocessing of the resumes and job descriptions is carried out by following four steps as shown in Fig. 1. (i) Word Tokenization: Individual words are extracted from the document. (ii) Bad characters removal: Numbers, punctuations, and special characters are removed from the list of words obtained from the previous step. (iii) Stop words removal: Stop words like for, the, and am, are removed from further as these words don't carry any meaningful information required for candidate selection. (iv) Stemming or Lemmatization: The root words of each word are obtained for further processing.

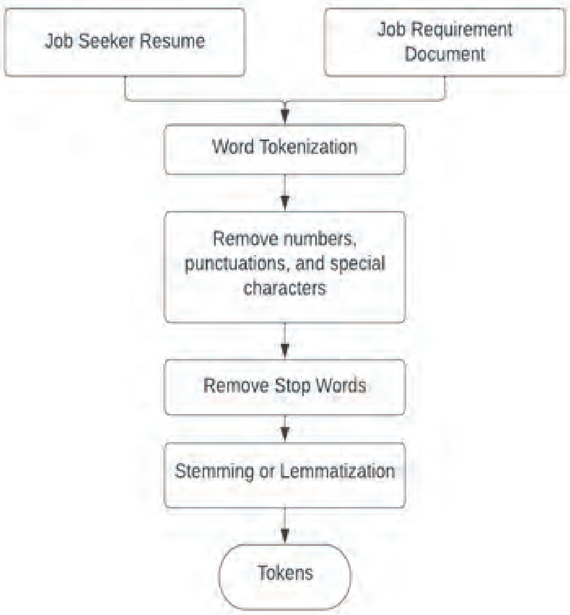


Fig. 1: Preprocessing of the resume and job description

After preprocessing, the frequency count of all root words in the resume and job requirements were saved in two separate lists each of length 'n' where 'n' represents the number of distinct words in both the resumes and job requirements combined. Each list is represented by a vector of 'n' dimensions, in which the vector value represents the frequency of each word in the resume or job requirement [2], [8]. For example, a resume has the statement 'Specialized in Web Developer' and the job requirement has the statement 'Looking for Web Developer'. The number of distinct words is 6 i.e. ['Specialized', 'in', 'Web', 'Developer', 'Looking', 'for']. After preprocessing the list will become ['Specialize', 'Web', 'Develope', 'Look']. So the vector for the resume will be represented as [1,1,1,0] and the job requirement will be [0,1,1,1] where each value represents the frequency of the corresponding word in the resume or job requirement. The similarity between the resume and job requirement is computed using cosine similarity as in equation (1).

$$cos(\Theta) = \frac{A.B}{|A||B|} = \frac{\sum_{i=1}^{n} A_i.B_i}{\sqrt{\sum_{i=1}^{n} A_i}.\sqrt{\sum_{i=1}^{n} B_i}} \quad (1)$$

where A is the vector representing the frequency list of the resume and B is the vector representing the frequency list of the job requirement. |A| is the magnitude of the vector A while |B| is the magnitude of vector B. The cosine similarity gives results in the interval between 0 to 1. If the value is closer to 1, it means vectors are closer to each other representing maximum similarity [1].

Algorithm: Content-Based Resume Recommendation System

1. Extract all tokens from resume using word tokenization
2. Remove numbers, punctuations and special characters
3. Remove words like (is, in, for, am), etc.

which are not significant

4. Find the root word using Stemming or Lemmatization
5. Remove duplicate words if any
6. Create 2 vectors one for job seeker's resume and one for job requirement of company of size equal to that of number tokens present after preprocessing. Place the frequency of each word in each document in the respective vector
7. Apply the cosine similarity formula as shown in equation (1)
8. The value obtained from cosine similarity is between 0 and 1. To convert it into percentage multiply the value by 100.
9. If the value is greater than 70 then the resume is perfect match for the job requirement else the resume is not a perfect match for the job requirement

Fig. 2: Algorithm

The proposed solution is provided by designing a website consisting of two modules: company module and job seeker module. In the Company module, each company can submit their job description. Similarly, in the candidate module, each candidate can submit their resume.

Our system works on a database model as shown in Fig 3 where relevant information of the users, companies and their resume data is stored in the database.

When a user enters the website they are asked if they are recruiters or job seekers as in Fig.4. From there they will be redirected to the login page. After successfully logging in, the job seeker will land on the profile page and from there they can update their profile, update their resumes and check out existing job opportunities present. The job seeker will also be able to view which all companies are interested in their profile.[4] Meantime the company will land on their profile page where they can post job openings and their requirements. They can also mark favorites if they are interested in a particular user.

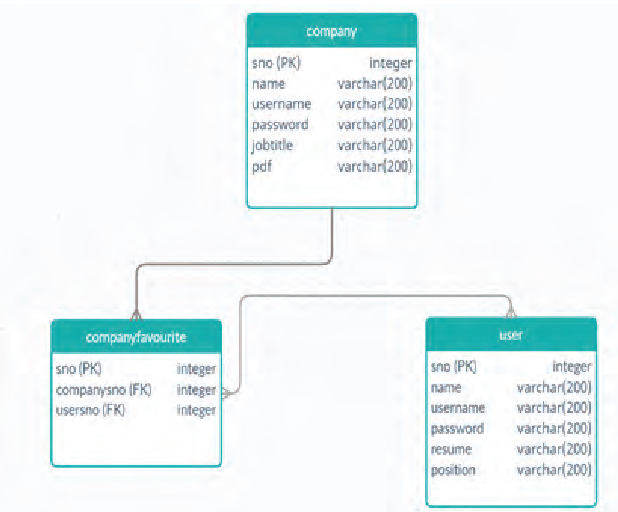


Fig. 3: Database schema

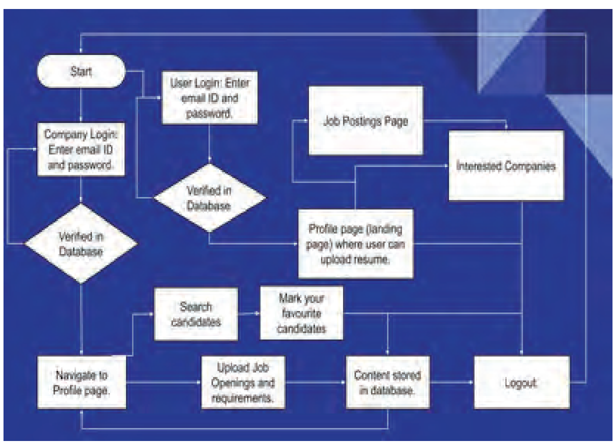


Fig. 4: Model Flow

RESULTS

The separate types of authentications were implemented successfully, leading to separate homepages and functions for both job seekers and companies.

The job seeker module was effectuated with the planned features, which included resume uploading, viewing the list of available jobs, and getting an idea of the potential jobs they may receive through the 'favorites' functionality that was implemented.

The company module was also implemented successfully with features such as uploading available jobs and viewing candidates that fit their requirements through our filtration system.

We have minimized the number of database calls using our algorithms to implement the above, and have an optimized website.

Two models have been built on the cleansed data:

Resume Classification System

The framework implements an algorithmic classification mechanism that systematically categorizes resumes into appropriate domains. This classification utilizes a Cosine-based mathematical model, with performance metrics carefully documented to validate the system's accuracy. The model's ability to correctly assign resumes to relevant categories serves as a fundamental step in the screening process.

Content-Based Recommender

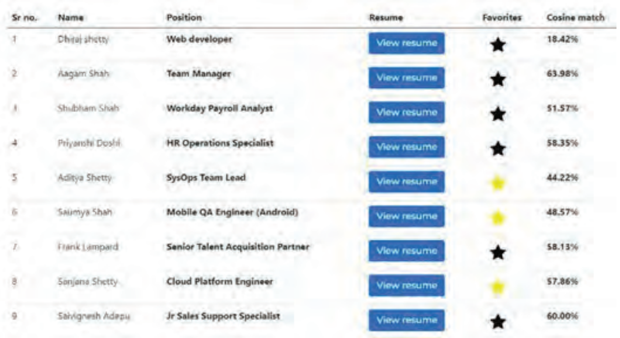
We regard the above case as document similarity identification similar to a Content based recommender work where the Job Description provided by the employer is used to match its content with the content of resumes in the document space and the n top(n configurable) matching resumes are recommended for the recruiter. Cleansed resume data and job description were taken together, and pushed into the model, followed by computing the cosine similarity between the job description and CVs.

Two datasets were used in the experimentation

Monster_india.csv

UpdatedResumeDataset.csv

Monster India dataset was used for the job requirements of the company. Updated Resume Dataset was required for the resumes of the user. The monster India had a job position column and the resume dataset also had a job position column. The model was given the entire dataset and it compared each resume with each job requirement. From the cosine similarity match, it could be concluded that the matches for the same job position of the user resume and the company job requirement were high and that for different job requirements were relatively low. If the job positions were similar the results that were governed were around 70% match while if the job positions were different the results fell as low as 10%.



Sr no.	Name	Position	Resume	Favorites	Cosine match
1	Dhiraj shetty	Web developer	View resume	★	18.42%
2	Aagam Shah	Team Manager	View resume	★	63.98%
3	Shubham Shah	Workday Payroll Analyst	View resume	★	51.57%
4	Priyanshi Doshi	HR Operations Specialist	View resume	★	58.35%
5	Aditya Shetty	SysOps Team Lead	View resume	★	44.22%
6	Saumya Shah	Mobile QA Engineer (Android)	View resume	★	48.57%
7	Frank Lampard	Senior Talent Acquisition Partner	View resume	★	58.15%
8	Sanjana Shetty	Cloud Platform Engineer	View resume	★	57.86%
9	Saivignesh Adapu	Jr Sales Support Specialist	View resume	★	60.00%

Fig. 5: [Company View] Cosine Match of Job Seeker Resume with their Job Requirement

In Fig.5. the Company can view the Cosine Percentage Match of their Job Requirements with the Resumes of the Job Seeker. The Company can also view the Resumes of all the Job Seekers. The Favorites Column shows if a particular job seeker is present in their Favorites section. The company can also view the job post the job seeker is searching for.

SYSTEM EVALUATION AND ANALYSIS

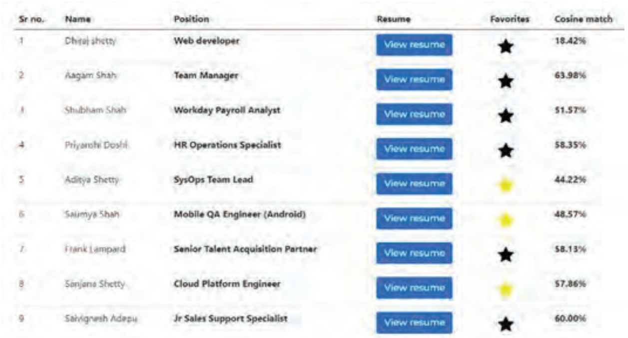
To demonstrate the effectiveness of the proposed content-based recommendation model, we conducted a comprehensive evaluation using multiple parameters. Inspired by the volunteer recommendation system analysis approach[9], we evaluated our system on various fronts, including accuracy, relevance, scalability, and user satisfaction. The following sections detail the evaluation process and findings.

Evaluation Parameters: Fifty participants, including HR professionals and recruiters, were engaged to assess the model's performance. They were tasked with evaluating the system on the following metrics, rated on a scale of 1 to 5:

1. Matching Accuracy: How accurately the system matches resumes with job descriptions.
2. Recommendation Relevance: The relevance of the recommended candidates to the job roles.
3. User Interface: Ease of use and clarity of the system interface.
4. Processing Time: Efficiency in processing large volumes of resumes.
5. Scalability: Ability to handle increasing data volumes without performance degradation.
6. Overall User Satisfaction: General satisfaction with the system's performance.

Evaluation Results: Based on participant feedback, Table 1 summarizes the ratings for each parameter:

Parameter	Rating (1/5)
Matching Accuracy	4.2
Recommendation Relevance	4.0
User Interface	3.8
Processing Time	4.3
Scalability	3.9
Overall User Satisfaction	4.1



Sr no.	Name	Position	Resume	Favorites	Cosine match
1	Dhiraj shetty	Web developer	View resume	★	18.42%
2	Aagam Shah	Team Manager	View resume	★	63.98%
3	Shubham Shah	Workday Payroll Analyst	View resume	★	51.57%
4	Priyanshi Doshi	HR Operations Specialist	View resume	★	58.35%
5	Aditya Shetty	SysOps Team Lead	View resume	★	44.22%
6	Saumya Shah	Mobile QA Engineer (Android)	View resume	★	48.57%
7	Frank Lampard	Senior Talent Acquisition Partner	View resume	★	58.15%
8	Sanjana Shetty	Cloud Platform Engineer	View resume	★	57.86%
9	Saivignesh Adapu	Jr Sales Support Specialist	View resume	★	60.00%

Fig. 6: [Job Seeker View] Recommended Companies to job seeker

In Fig.6. the Job Seeker is shown Job Offering of companies which has a Cosine Percentage match of above 60%. The Job Seeker can view the job requirements of those companies.

Our system provides efficiency improvements along with its accuracy functionality. A typical calculation of cosine similarity between job descriptions and resumes takes around 0.003 seconds thus providing quick assessments across all data sets. The sparse vector encoding method optimizes system memory consumption and parallel processing features enhance scalability significantly. The optimized framework makes the model work effectively with large datasets but does not reduce its operational speed.

CONCLUSION

Modern recruitment faces unprecedented challenges due to the massive volume of digital applications received for each position. Our research addresses the critical need for automating the candidate screening process through an innovative machine learning framework. The developed system demonstrates significant potential in revolutionizing traditional hiring practices through two key contributions: an intelligent resume classification mechanism and a precision-driven similarity matching algorithm.

By implementing content-based recommendation techniques, our system successfully bridges the gap between job requirements and candidate qualifications, offering recruiters a data-driven foundation for decision-making. The framework's ability to process unstructured resume data and generate quantifiable matching scores represents a significant advancement in recruitment automation.[9]

Future development will combine BERT (Bidirectional Encoder Representations from Transformers) and Sentence-BERT (SBERT) deep learning embeddings with our current cosine similarity matching system. The advanced approach has the ability to extract deeper contextual data which improves the semantic matching capabilities between job descriptions and resumes. The accuracy of our system can be enhanced by combining traditional TF-IDF vectorization with deep neural embeddings to keep computational complexity at manageable levels.

REFERENCES

1. A. R. Lahitani, A. E. Permanasari and N. A. Setiawan, "Cosine similarity to determine similarity measure: Study case in online essay assessment," 2016 4th International Conference on Cyber and IT Service Management, 2016, pp. 1-6, doi: 10.1109/CITSM.2016.7577578.
2. Z. Wang, X. Tang and D. Chen, "A Resume Recommendation Model for Online Recruitment," 2015 11th International Conference on Semantics, Knowledge and Grids (SKG), 2015, pp. 256-259, doi: 10.1109/SKG.2015.31.
3. V. Lai et al., "CareerMapper: An automated resume evaluation tool," 2016 IEEE International Conference on Big Data (Big Data), 2016, pp. 4005-4007, doi: 10.1109/BigData.2016.7841091.
4. A. Jaiswal, V. Jhawar, Y. Jadhav and M. Mahato, "TaskCO: Android App for Task Management," 2022 5th International Conference on Advances in Science and Technology (ICAST), Mumbai, India, 2022, pp. 358-361, doi: 10.1109/ICAST55766.2022.10039511.
5. Celma, O., 2010. Music recommendation, in: Music recommendation and discovery. Springer, pp. 43–85.
6. Das, A.S., Datar, M., Garg, A., Rajaram, S., 2007. Google news personalization: scalable online collaborative filtering, in: Proceedings of the 16th international conference on World Wide Web, ACM. pp. 271–280.
7. Diao, Q., Qiu, M., Wu, C.Y., Smola, A.J., Jiang, J., Wang, C., 2014. Jointly modeling aspects, ratings and sentiments for movie recommendation (jmars), in: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM. pp. 193–202.
8. Farber, F., Weitzel, T., Keim, T., 2003. An automated recommendation approach to selection in Personnel Recruitment.
9. P. P. Wadekar, S. K. Sachan, S. K. Patil, H. M. Wagh and M. Mahato, "V - Vibe: Efficient Volunteer Management using Mobile Technology," 2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2024, pp. 1236-1245, doi: 10.1109/ICSCSS60660.2024.10625339.

Time Series AQI Prediction using Sequence-Infused Autoencoders and Hybrid Deep Sequential Networks

Darakhshan Khan
Computer Engineering Department
Thadomal Shahani Engineering College
Mumbai University
Mumbai, Maharashtra
✉ darakshan.khan@thadomal.org

Juhi Janjua
Computer Engineering Department
Thadomal Shahani Engineering College
Mumbai University
Mumbai, Maharashtra
✉ juhi.ganwani@thadomal.org

Archana B. Patankar
Computer Engineering Department
Thadomal Shahani Engineering College
Mumbai University
Mumbai, Maharashtra
✉ archana.patankar@thadomal.org

Nabanita Mandal
Computer Engineering Department
Thadomal Shahani Engineering College
Mumbai University
Mumbai, Maharashtra
✉ nabanita.mandal@thadomal.org

ABSTRACT

Air pollution has emerged as a major international concern due to its hazardous impact on human health and the environment. Real-time AQI prediction becomes pertinent for timely management and safety protection. This study analyses the effectiveness of state-of-the-art deep learning architectures, which include sequence infused autoencoders and hybrid CNN-LSTM networks for three locations across Mumbai city. A feature vector with 21 variables was preprocessed using GAN based network for imputation and anomaly treatment. Models were evaluated using regression metrics such as, RMSE, MAPE and R2 score. Results demonstrate that Bi-LSTM AE and hybrid network has outshined on all metrics, indicating its robustness and generalization ability.

***KEYWORDS** : Autoencoder, Hybrid, Convolution, SVR, AQI.*

INTRODUCTION

With increasing air pollution and its negative impact on human health, AQI forecasting has emerged as a crucial area of research in recent years. The Air Quality Index (AQI) is determined based on the concentration of various air pollutants. These pollutants can be broadly classified into two types: primary air pollutants and secondary air pollutants. Direct emissions to the environment occur due to industrial processes, vehicles, and natural events. These include particulate matter, such as PM10 and PM2.5, nitrogen oxides, h dioxide, and carbon monoxide. Secondary air pollutants, however, are formed through chemical reactions within the atmosphere and result from exposure to hydrocarbons, NOx, and sunlight. An example of a secondary air pollutant is ozone (O3) [1]. Other factors affecting AQI include meteorological conditions, including temperature, wind speed, humidity, pressure, and dew point, which affect the dispersion and transformation of pollutants and spatial factors such population density.

AQI, being a sequential data, shows several characteristics which make it a perfect candidate for time series analysis and predictive modelling. It has a strong temporal relationship as values are recorded at different periods that capture changes in AQI levels over time. Recently, there is huge advancement on techniques used for analyzing time series data as there are abundant of applications [2] such as, finance, healthcare, weather, supply chain, retails, manufacturing, agriculture, etc.

Various machine learning regression algorithm such as support vector machine, random forest, KNN, etc., are used for AQI forecasting as they can capture non-linear dependency with less computational cost [3]. Advanced deep learning architectures have revolutionized AQI forecasting giving not only more reliable and accurate forecasts but also understanding hidden patterns and converting them into valuable insights. It includes hybrid architectures that use both Convolution Neural Networks (CNN) and LSTM [4] to capture spatial and temporal relationships. Other architectures are attention-based networks that give more weightage to certain parts of a time series which is difficult to model. For AQI forecasting, autoencoder and their variants are used as they can reduce dimensions while understanding latent patterns, remove noise, and extract effective features needed by AQI forecasters [5].

AQI being a time series data, this study focuses on devising advanced deep networks that can be used for time series analysis across domain while using AQI dataset as research object. The structure of this paper is as follows: The background research in Section 2 discusses work done utilizing deep networks for forecasting multivariate time series, such as AQI. Section 3 complete methodology adopted for this study including designs of proposed architectures. Section 4 highlights the results achieved on different metrics along with visual plots.

REVIEW OF LITERATURE

Various works on air quality prediction utilized Support Vector Regression and Support Vector Machines, where different experiments were performed in comparison with other models. The study [6] used the Environmental Protection Agency (EPA) data set for California concerning air and weather quality, applied Yeo-Johnson power transformation in the pre-processing phase, and then determined the hourly AQI and O_3, CO, SO_2 pollution levels for California by SVR model. In one of the related research efforts, an 11-year data set from EPA for three cities of Taiwan were compiled, where k-NN imputation was used to fill the missing values of the dataset. The result demonstrated that the popular algorithms of SVMs with its different kernel techniques, including the polynomial, RBF, and linear kernels; showed inferior performance and were dominated by that of the Adaptive Boosting along with Stacking Ensemble [7].

This further leads to a novel Grey Wolf Optimization with Decision Tree (GWO-DT) method being proposed for predicting AQI values in major Indian cities, like Delhi, Hyderabad, and Kolkata. The performance of this method was evaluated using performance metrics such as R-Square, RMSE, MSE, MAE, and accuracy, in comparison to traditional models, which involve k-NN, Random Forest, and SVR. The results reported a high accuracy and showed competitive performance of the models based on SVR with results such as 97.68% accuracy for Visakhapatnam, 97.66% for Hyderabad, and 88.98% for New Delhi [8]. Collectively, these findings underlie the relevance of support vector machines in air quality prediction across various datasets and regions.

Autoencoders have been presented as a good choice for air quality forecasting by capturing patterns in the data and increasing the accuracy of predictions. In an article [9], introducing an Autoencoder-based Air Quality Predictor of outdoor air quality over Beijing, used sequence-to-sequence architecture for prediction of future AQI values with input of hourly measurements of air quality and weather. To address the aforementioned problems, namely slow training and error accumulation, the authors have replaced the classical RNN encoder with a fully connected encoder added with position embedding. Moreover, an n-step recurrent prediction approach has been utilized, which has resulted in significant performance improvement in terms of long-term forecasting.

Hybrid models combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks have been found to be significantly promising for air quality forecasting because they can efficiently capture spatial and temporal features. One study [10] proposed a novel PM2.5 forecasting model for metro stations using ensemble by combing CNN-GRU and CNN-LSTM networks. This is a three-stage model. It first decomposes the PM2.5 time series into several subseries with VMD, and then predicts the regular part of components with CNN-GRU while CNN-LSTM predicts irregular components. In the end, these networks together increase the precision of the entire forecast. This research [11] compares prediction of the AQI values of ten large Chinese cities for different deep learning models, RNN, Bi-GRU, BiLSTM, CNN-BiLSTM, Conv1D-BiLSTM, and the XGBoost of the machine learning model. Between these neural networks, CNN-BiLSTM and Conv1D-BiLSTM were proven to be highly capable in understanding the complex non-linear patterns within the air quality data. These findings indicate the merits of hybrid CNN-LSTM models in providing strong, accurate forecasts by combining convolutional feature extraction with LSTM's sequential learning.

METHODOLOGY

This section elaborates on methodology used for this research, which involved data acquisition and pre-processing, model development and performance evaluation.

Data Cleansing

The pollutant parameter data has been taken from the CPCB portal [12], and the climatic parameters are sourced from OpenWeatherMap [13] for three locations in Mumbai, namely, Bandra, Colaba, and Kurla. The dual source would ensure full coverage which includes total of 21 features including air pollutants and meteorological factors. Data preprocessing is an important aspect that ensures a good prediction accuracy in machine learning models. For this study, missing values in the dataset were imputed using a GAN-based network [14], thereby maintaining realistic patterns and distributions for the imputed data. Upon filling the missing values, the dataset was subjected to anomaly detection using statistical inter-quantile ranges method. It was quite effective in picking up outliers in the dataset. These outlying data points also were replaced via the same GAN-based network to keep uniformity of the data. After the cleaning, the dataset was fed to standardization and normalization module to transform values between 0 and 1 for maintaining equalization, which helps in avoiding overfitting by decreasing effects of differing scales across features. Lastly, the entire data set is split into three subsets, namely, training, validation, and testing set with a ratio of 70:20:10.

Model Development

In this section, multiple techniques use for AQI forecasting is explained. This research uses support vector regression (SVR) as baseline model and results are compared with advance deep leaning architectures such as sequence infused autoencoders and hybrid sequential convolution and recurrent architecture (HSCRA).

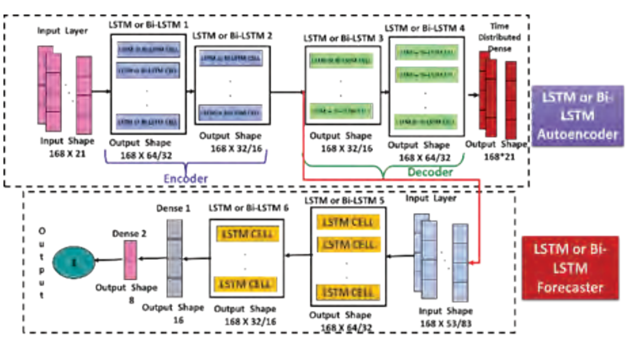


Fig. 1. Sequence Infused Autoencoder Architecture

a) Sequence Infused Autoencoders: Autoencoders are a powerful model for air quality prediction, which manages complex datasets by minimizing noise and offering robust, accurate forecasts. In the two-stage proposed deep learning architecture, Fig. 1 presents the LSTM/Bi-LSTM Autoencoder and LSTM/Bi-LSTM Forecaster for AQI time-series prediction. The auto-encoder processes the input data 168 × 21 through the stacked LSTM or Bi-LSTM layers to encode the input into a latent space (168 × 32/16). The decoder is similar to the encoder which is complemented by additional LSTM/Bi-LSTM layers reconstructing the data back into its original form via a Time Distributed Dense layer (168 × 21). The forecaster takes the latent representation and combines it with the actual input to form a feature space of shape 168 × 53/83. This is passed to stacked sequential layers with 64/32 and 32/16 cells, followed by Dense layers with 16 and 8 neurons, finally producing single AQI predictions. This architecture ensures denoising, compression, and accurate sequence forecasting.

b) Hybrid Sequential Convolution and Recurrent Architecture (HSCRA): This architecture uses convolution layers to gather spatial components from input data, whereas LSTM layers capture sequential temporal relationships. Fig. 2. represents a Hybrid CNN-LSTM Network used for multivariate AQI forecasting. It uses 1 or 3 CNN layers accompnaied by 1 or 3 LSTM layers, using ReLU activation and a kernel size of 3 for CNNs. The flatten layer transfoems the outpur into a 1D tensor and finally passed to 1 or 3 dense layerswhich have [32, 16, 1] neurons. Thus, this hybrid architecture ensures the robust learning of spatial correlations as well as temporal dynamics and maintains natural time-series information flow. For different locations different architectural configuration is used as shown in table 1.

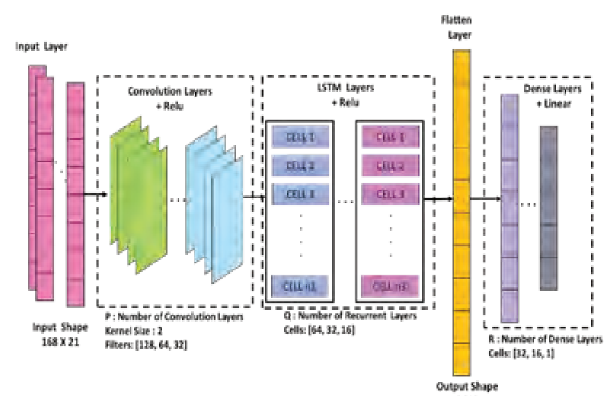


Fig. 2. Hybrid Sequential Convolution and Recurrent Architecture

Performance Evaluation

Evaluating performance means testing prediction accuracy of deep learning models on unseen data to understand how well a trained model will generalize on unseen observations. Three metrics used are, Root Mean Squared Error (RMSE) which provides a measure of the difference between predictions and observations. Another metric is Mean Absolute Percentage Error (MAPE) that shows the percentage accuracy of predictions. Lastly, the R^2 score indicates how good is the fitting model in capturing variance in the target variable [15, 16S].

Table 1. Station-wise model configuration used by HSCRA architecture

Models Configuration Parameters	Stations		
	Bandra	Colaba	Kurla
P, Q, R	3, 3, 3	3, 3, 3	2, 2, 3
CNN Filters	128, 64, 32	128, 64, 32	64, 32
LSTM Cells	64, 32, 16	64, 32, 16	64, 32
Neurons in Dense Layer	32, 16, 1	32, 16, 1	32, 16, 1

RESULTS AND DISCUSSION

In this work, authors have implemented three advance DL architectures for three different locations to prove robustness of proposed approach and compared with baseline SVR method. A complete feature vector with 21 variables, in particular, NOx, SO2, CO, Ozone, PM2.5, and PM10 temperature, dew point, pressure, humidity, wind speed, distances from shore, population density etc., were utilized for accurate forecasting. Past one-week hourly frequency (24*7) data was used to make a single prediction of AQI on next day. Autoencoder models were trained for 150 epochs whereas HSCRA models were trained for 100 epochs with 256 batch size.

Table 2. Station-Wise RMSE Values for all architectures

Models	Stations		
	Bandra	Colaba	Kurla
SVR	46.35	38.25	63.06
LSTM AE	25.34	30.38	39.27
Bi-LSTM AE	23.96	21.03	30.88
HSCRA	26.63	20.67	26.66

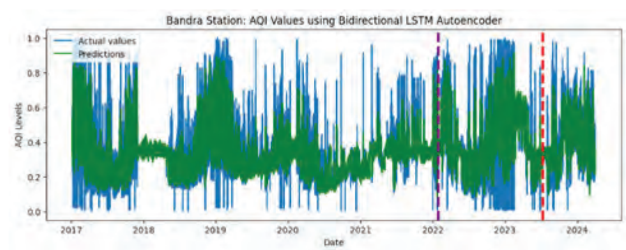
Table 3. Station-Wise MAPE Values for all architectures

Models	Stations		
	Bandra	Colaba	Kurla
SVR	0.316	0.203	0.225
LSTM AE	0.122	0.192	0.208
Bi-LSTM AE	0.115	0.129	0.159
HSCRA	0.138	0.121	0.135

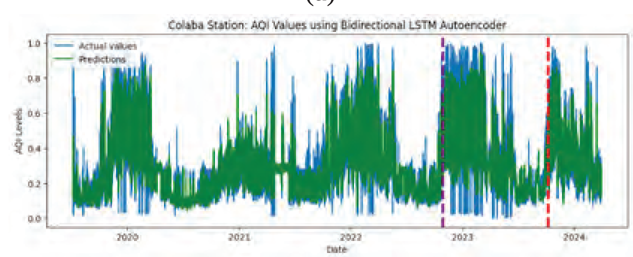
From tables 2 to 4, it is evident that based on metrics values achieved if we perform cumulative analysis, the HSCRA is the best performing model at all metrics, with lowest aggregated RMSE and MAPE, and the highest mean R^2 score, so this is the most resilient model in producing AQI forecasts in this investigation. The Bi-LSTM AE model also performed consistently well, especially in terms of RMSE and MAPE values. Although SVR produced a decent R^2 score, its RMSE and MAPE values which are very high indicate it is less accurate in making predictions than the deep learning models. The performance of LSTM AE was mixed, with some stations performing well while others performed poorly, showing negative R^2 scores, which is clear sign of model's inability in generalization.

Table 4. Station-Wise R2 Score Values for all architectures

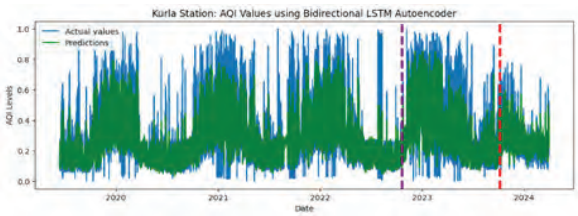
Models	Stations		
	Bandra	Colaba	Kurla
SVR	0.417	0.590	0.501
LSTM AE	0.588	0.111	-0.87
Bi-LSTM AE	0.643	0.769	0.058
HSCRA	0.45	0.74	0.56



(a)



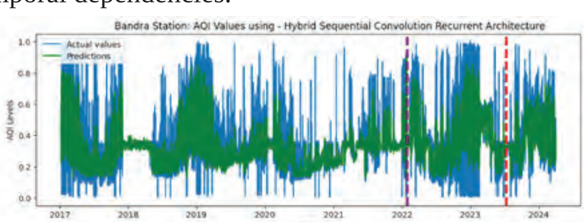
(b)



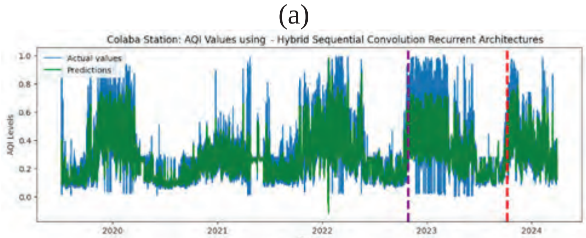
(c)

Fig. 4. Station-wise AQI predicted and Actual values for Bi-LSTM infused Autoencoder Architecture

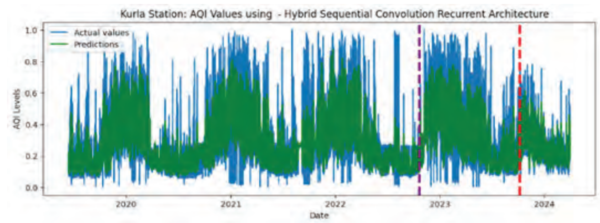
Based on tables values we have depicted the visual representation of best performing models for all stations in Fig. 4 and 5. It presents station-wise AQI forecast using Bi-LSTM AE and HSCRA architecture respectively for three locations, namely, Bandra, Colaba, and Kurla. In all plots, blue color indicates actual values and green values are predictions made by models. Data points before purple dashed lines are used as training set, between purple and red dashed lines are used as validation set, and after red lines are used as test set. The plots further illustrate claims that Bi-LSTM AE model closely traces the AQI levels with variations over time; thus, in addition to detecting short-term patterns, it is able to capture long-term trends as well. Similarly, HSCRA models shows good predictive performance with accurately forecasting AQI values closely following the actual values for all three stations. This captures seasonal patterns as well as significant swings in the levels of AQI. The consistent alignment indicates the robustness and generalization capability of the model in multivariate time-series forecasting. The use of hybrid architecture combining both convolutional and recurrent layers is successful in learning both spatial and temporal dependencies.



(a)



(b)



(c)

Fig. 5. Station-wise AQI predicted and Actual values for HSCRA Architecture

CONCLUSION

This research shows the capability of advanced deep learning models in AQI forecasting, and the effectiveness of hybrid architectures in learning complex spatio-temporal patterns in multivariate time-series data. Among the models tested, HSCRA proved to be the best architecture with the lowest RMSE and MAPE values and the highest R^2 score for all three stations, which indicates its superior predictive capability. Consistent performance was also demonstrated with the Bi-LSTM Autoencoder in terms of capturing short-term variability and long-term trends. While a reasonable R^2 score for SVR is reported, higher error metrics indicate less reliability for accurate AQI prediction. The study concludes that hybrid models combining convolutional layers with recurrent layers are better suited for AQI forecasting as they can handle both spatial correlations and temporal dynamics. Future work may include enhancing hybrid models in new ways, extending the real-time forecasting capabilities to other domains related to time series. Additionally, ensemble or attention based network can further enhnace and improve models prediction capability.

ACKNOWLEDGMENT

The National Environmental Engineering Research Institute (NEERI), Nagpur, the Central Pollution Control Board, State Pollution Control Boards, and Pollution Control Committees are all acknowledged by the authors.

REFERENCES

1. D. Khan, A. B. Patankar, R. Motwani and A. Ailsinghani, “Investigating Efficacy of RNN and its Variants for AQI Forecasting,” 2023 International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT), Faridabad, India, 2023, pp. 1303-1307, doi: 10.1109/ICAICCIT60255.2023.10465781.

2. F. Dama and C. Sinoquet, "Analysis and modeling to forecast in time series: a systematic review," *CoRR*, vol. abs/2104.00164, 2021. [Online]. Available: https://arxiv.org/abs/2104.00164.

3. S. Srikamdee and J. Onpans, "Forecasting Daily Air Quality in Northern Thailand Using Machine Learning Techniques," 2019 4th International Conference on Information Technology (InCIT), Bangkok, Thailand, 2019, pp. 259-263, doi: 10.1109/INCIT.2019.8912072.

4. V. -D. Le, T. -C. Bui and S. -K. Cha, "Spatiotemporal Deep Learning Model for Citywide Air Pollution Interpolation and Prediction," 2020 IEEE International Conference on Big Data and Smart Computing (BigComp), Busan, Korea (South), 2020, pp. 55-62, doi: 10.1109/BigComp48618.2020.00-99.

5. R. Xu, D. Wang, J. Li, H. Wan, S. Shen, and X. Guo, "A Hybrid Deep Learning Model for Air Quality Prediction Based on the Time–Frequency Domain Relationship," Atmosphere, vol. 14, no. 2, p. 405, 2023, doi: 10.3390/atmos14020405.

6. M. Castelli, F. M. Clemente, A. Popovič, S. Silva, and L. Vanneschi, "A Machine Learning Approach to Predict Air Quality in California," Complexity, vol. 2020, Article ID 8049504, 23 pages, 2020. doi: 10.1155/2020/8049504.

7. Y.-C. Liang, Y. Maimury, A. Chen, and J. Juarez, "Machine Learning-Based Prediction of Air Quality," Applied Sciences, vol. 10, p. 9151, 2020. doi: 10.3390/app10249151.

8. S. K. Natarajan, P. Shanmurthy, D. Arockiam, et al., "Optimized Machine Learning Model for Air Quality Index Prediction in Major Cities in India," Scientific Reports, vol. 14, no. 6795, 2024. doi: 10.1038/s41598-024-54807-1.

9. B. Liu et al., "A Sequence-to-Sequence Air Quality Predictor Based on the n-Step Recurrent Prediction," IEEE Access, vol. 7, pp. 43331-43345, 2019, doi: 10.1109/ACCESS.2019.2908081.

10. Yu, C., Yan, G., Ruan, K. et al., "An ensemble convolutional reinforcement learning gate network for metro station PM2.5 forecasting," Stoch. Environ. Res. Risk Assess., vol. 37, no. 8, pp. 2359-2372, 2023. doi: 10.1007/s00477-023-02564-4.

11. I. Ayus, N. Natarajan, and D. Gupta, "Comparison of machine learning and deep learning techniques for the prediction of air pollution: a case study from China," Asian J. Atmos. Environ., vol. 17, no. 4, 2023. doi: 10.1007/s44273-023-00005-w.

12. Central Pollution Control Board: CPCB Home page. https://app.cpcbccr.com/AQI_India/ last accessed on 2024/10/24.

13. OpenWeather Homepage, https://openweathermap.org, last accessed 2024/10/02.

14. C. Fang and C. Wang, "Time Series Data Imputation: A Survey on Deep Learning Approaches," arXiv preprint arXiv:2011.11347, Nov. 2020. Available: https://arxiv.org/abs/2011.11347

15. Rizwan, K.D., Patankar, A.B, "Comparative Investigation of Machine Learning and Deep Learning Methods for Univariate AQI Forecasting," In: Balas, V.E., Semwal, V.B., Khandare, A. (eds) Intelligent Computing and Networking. IC-ICN 2023. Lecture Notes in Networks and Systems, vol 699. Springer, Singapore, 2023. https://doi.org/10.1007/978-981-99-3177-4_14.

16. Nabanita, M., Tanuja, S, "Prediction of Dew Point Temperature and Relative Humidity for Nashik Region Using LSTM," In: Sarkar, D.K., Sadhu, P.K., Bhunia, S., Samanta, J., Paul, S. (eds) Proceedings of the 4th International Conference on Communication, Devices and Computing. ICCDC 2023. Lecture Notes in Electrical Engineering, vol 1046. Springer, Singapore, 2023. https://doi.org/10.1007/978-981-99-2710-4_39.

Smart IoT Based Pollution Monitoring System for Air and Water

Nilesh Shimpi
Assistant Professor
Atharva College of Engineering
Malad, Mumbai, Maharashtra
✉ nileshshimpi@atharvacoe.ac.in

Snehal Bhelkar
Assistant Professor
Atharva College of Engineering
Malad, Mumbai, Maharashtra
✉ snehalbhelkar@atharvacoe.ac.in

Prajakta Borole
Assistant Professor
Atharva College of Engineering
Malad, Mumbai, Maharashtra
✉ prajaktaborole@atharvacoe.ac.in

ABSTRACT

The growth of pollution is increasing day by day. Some things can affect the environment and cause loss in biodegradation. This can be due to rapid economic growth and urbanization. Toxic substances in the environment can be harmful to human health and require special precautions. The model continuously measures water levels and calculates them online to the host computer through the Internet of Things. With internet access, survey data can be obtained from remote locations without the need for easy access. The structure of this analysis model depends on the relationship or interaction between the distribution of the transformed data and the data structure. It works with new ideas using Internet of Things, cloud, air and water computing. This allows the government to monitor air and water pollution in different regions and take precautions. Water pollution is any change in the biological, physical or chemical components of water that ultimately harms any organ.

KEYWORDS *: Internet of Things (IoT), Arduino, Sensor, Cloud.*

INTRODUCTION

In this paper, the model uses air actuators to detect the presence of pollutants and compounds in the air and water. The model continuously transmits this information while tracking scaled sound levels and cycles like [1]. The main purpose of this paper is to create and make a useful assessment model that can be used and written by guidelines for air and water pollution. The actuators and sensors are stored in the cloud and drive the estimated deviation of the web browser, so it is important to monitor it [3]. The monitoring model uses actuators and sensors to detect pollutants such as carbon dioxide (CO2), sulfur dioxide (SO2) and others. The products in the environment equipped with actuators, microcontrollers and various software applications serve as personal protection and environmental monitoring [4], also known as smart environment. It helps to understand many things and transfers information and stored regularly and continuously on cloud. Monitoring environmental most of them do not need remote monitoring technology. These IoT systems have many features that can be easily integrated into devices and work in IoT. IoT technology supports the development of equipment in real time with information about performance. The Internet of Things is an innovation that connects sensors to the system and allows data from sensors to be sent to the network. We create models that can detect the variability of parameters such as weather, noise, temperature and humidity. Water quality for commercial purposes has different characteristics depending on the specific job. Some freshwater resources, such as groundwater and surface water, are low-use resources. However, commercial/ human activities and other natural processes can pollute these resources. This method also has its limitations. For example, if a farmer wants to check the water quality of his crops, he has to wait a week or more for each test. However, it is possible to manually estimate water quality using sensors to get differential results. This will take a lot of time, so IoT with cloud technology can be a solution that will provide predictability.

LITERATURE REVIEW

In today's generation, many types of pollution have been created and developed, including different types of work done with them [6]. It uses the wireless sensor model developed by various networks to monitor and control physical and ecological conditions with many applications in various places. These are different types of commercial scales available in the market, including the Fortuity CO-220 carbon monoxide scale, the amprobe CO2 meter specially designed for CO2 leakage detection, the forbix Semicon LPG alarm. Researchers[7] have been developing air and noise pollution models over the years using Wireless Sensor Networks (WSN), Global Mobile System (GSM), Geographic Information Systems, etc. According to the functional purpose of Zigbee, it is implied that it will work with cross receivers and Bluetooth. Here, the actuator element communicates directly with the IoT, eliminating the use of complex operating procedures. We use RFID to collect and receive data. RFID system has two main components called tags and readers [8]. Tags consist of unique identifiers (IDs) and store additional information regarding temperature, humidity, etc. information for the tag. To track tags, tags are fixed or molded into objects in complex RFID applications.

In 2019, Umair Ahmed et al. As water quality is deteriorating at an alarming rate , the rate of disease stress, which is a direct impact of rapid urbanization and industrialization, also continues to increase . Their study tracked and applied machine learning algorithms to calculate Water Quality Index (WQI) and Water Quality Class (WQC); the former is a measure that describes the overall water quality and the latter is derived from WQI and is specific to the group. In 2022, Manisha Koranga et al. Machine learning algorithms were analyzed and eight regression algorithms and nine classification algorithms were used. Random forest, SVM, and stochastic gradient descent are considered as the best machine learning algorithms.

PROPOSED SYSTEM

The proposed structure is used to control the air in the environment, to make the environment suitable and to provide more interaction with the objects through wireless communication. This model is designed to meet the needs and motives of the public to analyze the air and water pollution in an area in real time through the Internet of Things. The working flow model is shown in Figure 1. The working flow model has a 4-layer architecture that includes all the functions in one module and is designed to monitor air and water pollution. Here, each layer has its own job and task. Tier 1 provides information on whether there is a responsibility to monitor and control air and water. Tier 2 includes the actuators and sensors with their operating characteristics and features. The structure is more flexible and integrated into the ecological management framework environment. Actuator ownership information and construction options will occur in Tier 3, while appropriate interaction will occur in Tier 4. In Tiers 2 and 3, Identification and monitoring functions are usually selected on a case- by-case basis, such as setting the threshold, changing alarms, audible alerts or LEDs, etc.

This is based on Tier 3 and Tier 2 performance information, and differs from the previous article by including the initial price model taken at the main plant and the work site. This ultimately represents data for a specific parameter. Layer 4 has appropriate mechanisms that can recognize changes in extreme prices. It is based on the determination of the CO degree or the water level. The information at this level will be analyzed and stored in the cloud.

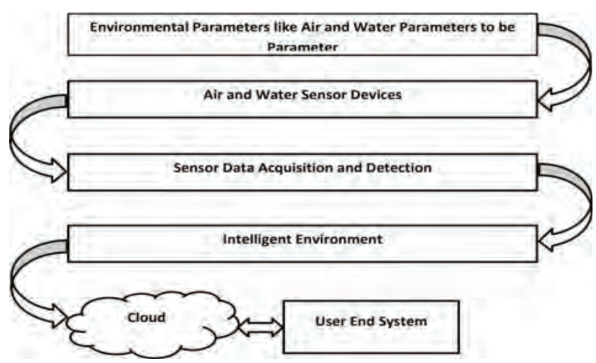


Fig. 1 Working Flow Model

In other words, we can say that it is a treasure trove of Google spreadsheets that demonstrate the practice of understanding specific data. Users can now search or information via mobile phones, computers, etc. We know that air pollution and water pollution are new problems at present and are also the main targets of the Internet of Things The purpose of modeling is to control and minimize the pollution.

The goal of air and water pollution management in the Internet of Things (IoT) is to enhance the monitoring and control of air and water quality to prevent or mitigate pollution, ensuring clean and safe water sources for communities. This can be achieved through various key components and applications, including real-time water quality monitoring, automated response systems, and efficient water management applications.

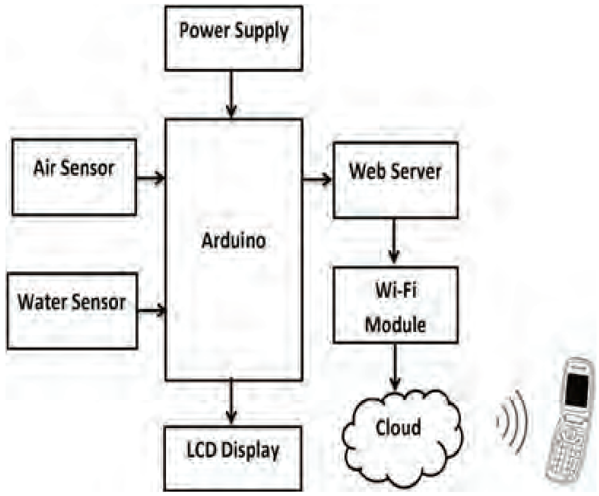


Fig. 2 Proposed System Architecture

River and marine water pollution detection sensors are devices designed to monitor and identify pollutants in water bodies like rivers, oceans, and lakes. These sensors are critical for maintaining water quality, protecting ecosystems, and ensuring the safety of aquatic life and human populations. Here an overview of how these sensors work and the types commonly used:

Types of Water Pollution Sensors

Chemical Sensors

PH Sensors: Measure the acidity or alkalinity of the water. A sudden shift in pH can indicate pollution. Dissolved Oxygen Sensors: A drop in dissolved oxygen levels could suggest the presence of organic pollutants or algal blooms.

Turbidity Sensors: Measure the cloudiness of water, which can be caused by suspended particles like sediments or pollutants.

Nitrate and Phosphate Sensors: High levels of nitrates and phosphates can indicate agricultural runoff or wastewater discharge.

Biological Sensors

Biological Oxygen Demand (BOD) Sensors: Measure the amount of oxygen needed by microorganisms to decompose organic matter in the water, which can indicate pollution.

Microbial Detection Sensors: These sensors detect the presence of harmful bacteria like E. coli, which can indicate sewage contamination.

Heavy Metal Detection Sensors

Detect toxic metals such as mercury, lead, cadmium, and arsenic, which are commonly found in industrial waste and mining runoff.

Optical Sensors

Fluorescence Sensors: Measure the fluorescence emitted by organic compounds, such as oil spills, which are harmful to marine life.

UV-Vis Absorption Sensors: Used to detect certain types of contaminants like pesticides and oils.

Temperature and Salinity Sensors

Temperature: Water temperature fluctuations can indicate pollution, as polluted water often has a different thermal profile.

Salinity: Sudden changes in salinity levels can also indicate freshwater runoff or industrial discharges. Deployment Methods:

Fixed Monitoring Stations: These are set up in key locations like riverbanks or near industrial areas to continuously monitor water quality.

Mobile Sensors: Drones, boats, or even submerged devices can be used to sample and monitor larger areas of water, including remote or offshore locations.

Wearable Sensors: Portable sensors used by researchers or field teams to take on-the-go measurements at different points along rivers or marine sites.

RESULT ANALYSIS

This system uses IoT technology to monitor and analyze air and water pollution. We measured various pollutants and water contaminants that can cause air and water pollution and used models for real-time monitoring, control and analysis. Real-time monitoring enables us to work on time and prevent a major incident from occurring, both by displaying alerts on LCD screens and uploading them to the cloud using IoT. The results can be divided into three parts. It should have functions such as displaying data on LCD monitors, displaying data on IoT platforms for remote access, and warning users when the environment is too polluted. Temperature, humidity, CO, and dust density values are displayed on the LCD screen.

Fig. 3 Output on LCD screen

As seen in the figure, the results of various parameters like CO, NH3 tested and displayed on the LCD screen. The relationship between the PPM value of NH3 and date is shown in the figure.

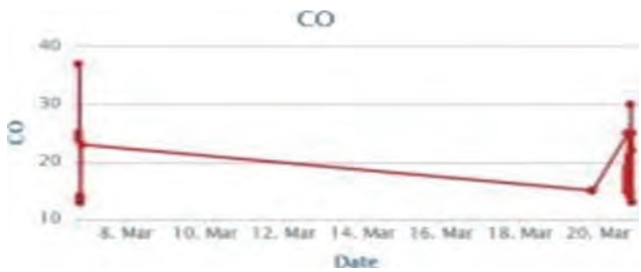


Fig. 4 CO in ppm / date

Fig. 5 NH3 in ppm / date

Fig. 6 Sending SMS

It also sent the SMS to the administrator as shown in figure 6 & 7.

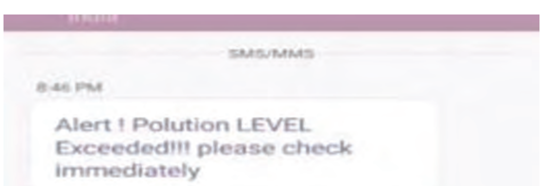


Fig. 7 Generated Messages

CONCLUSION

The proposed IoT-based architecture for weather and pollution monitoring offers a powerful solution for cities dealing with the challenges of pollution and urbanization. Through the combination of cloud computing, edge and fog computing, and predictive analytics, the system provides real-time monitoring, enhances decision-making, and offers a platform for innovative applications. By empowering policymakers with data-driven insights, Such systems can play an important role in improving air and water quality , and environmental sustainability in smart cities.

REFERENCES

1. L.Ezhilarasi, K.Sripriya, A .Suganya , 4K.Vinodhini .: A System for Monitoring Air and Sound Pollution using Arduino Controller with IOT Technology Vol. 3 Issue 2 (2017) Pages 1781 – 1785.

2. Navreetinder Kaur, Rita Mahajan, Deepak Bagai, “Air Quality Monitoring System based on Arduino Microcontroller,” International Journal Innovative Research in Science, Engineering and Technology (IJIRSET), Vol 5, Issue 6- June 2016

3. Umair Ahmed, Rafia Mumtaz, Hirra Anwar, Asad A. Shah, Rabia Irfan and Jose Garcia-Nieto (2019). Efficient Water Quality Prediction Using Supervised Machine Learning.

4. Uppuguduru Anil Kumar, G.Keerthi, G.Sumalatha, M.Sushma Reddy “ Iot Based Noise and Air Pollution Monitoring System” International Journal of Advance Technology.

5. Arusi Singh, Divya Pathak, Prachipandit, Shrutipatil, Prof. Priti. C. Golar.” IoT Based Air and Sound Pollution Monitoring System” International Journal Of Advanced Research In Electrical, Electronics And Instrumentation Engineering Volume 6, Issue 3, March 2017.

6. Haghiabi, A. H., Nasrolahi, A. H., &Parsaie, A. (2018). Water quality prediction using machine learning methods.

7. Palghat Yaswanth Sai: An IoT Based Automated Noise and Air Pollution Monitoring System Vol. 6, Issue 3, March 2017.

8. Mohammed Al-Yaari, Hasan Alkahtani and Mashael Maashi (2020). Water Quality Prediction using Artificial Intelligence Algorithms.

Design and Development of Contextual Framework for Translation and Interpretation of Vedas

Ravi Prakash
Professor
Computer Engineering Dept
KCCEMSR
Thane, Maharashtra
✉ jravi54@gmail.com

Umair Malim
Student
Computer Engineering Dept
KCCEMSR
Thane, Maharashtra
✉ malimumair2@gmail.com

Rashi Sharma
Student
Computer Engineering Dept
KCCEMSR
Thane, Maharashtra
✉ rashisharma2204@gmail.com

Sushmita Shukla
Student
Computer Engineering Dept
KCCEMSR
Thane, Maharashtra
✉ sushmitashuka@kccemsr.edu.in

ABSTRACT

"Vedify" is a unique application created to translate the four Vedas from Sanskrit to Hindi, making these ancient scriptures more accessible to today's generation. With the growing disconnect from Sanskrit, many people struggle to understand the wisdom of the Vedas, which hold immense cultural and spiritual significance. Vedify bridges this gap by providing accurate and meaningful translations that preserve the original context and essence of these texts.The application relies on advanced Natural Language Processing (NLP) and Machine Learning (ML) technologies to achieve this. It processes Sanskrit texts using techniques like tokenization and syntactic analysis, while ML models, trained on curated datasets, ensure translations are not just literal but contextually accurate. Tools like Python, TensorFlow, and Hugging Face Transformers power the backend, while a simple and intuitive interface ensures a smooth user experience. Vedify is not just a translation tool—it's a step toward preserving and promoting India's ancient knowledge in a modern, relatable way. By combining cutting-edge technology with cultural heritage, it aims to help more people connect with the timeless teachings of the Vedas. This project highlights how technology can bridge the gap between the past and the present, making ancient wisdom accessible to all.

***KEYWORDS** : Vedas translation, Natural language processing (NLP), Machine learning (ML), Contextual translation, Ancient text interpretation, Indian scriptures, Text semantics and syntax.*

INTRODUCTION

The project "Vedify" is an initiative to translate the four Vedas—Rigveda, Samaveda, Yajurveda, and Atharvaveda— from Sanskrit to Hindi using advanced Natural Language Processing (NLP) and Machine Learning (ML) techniques. The Vedas are among the oldest texts in human history, containing profound knowledge on spirituality, philosophy, science, and rituals. However, the diminishing familiarity with Sanskrit has created barriers to understanding these ancient scriptures. Vedify seeks to bridge this gap by making the Vedas accessible to a broader audience while retaining their essence and depth.

WHY VEDIFY?

The Vedas: Origins and Significance

The Vedas, composed between 1500 BCE and 500 BCE, are the cornerstone of Indian culture and spiritual tradition. They were originally passed down orally before being documented in Sanskrit, a language renowned for its precision and structure. Each Veda has a unique focus: Rigveda: Contains hymns praising various deities and cosmic principles.

Samaveda: Focuses on melodic chants and their application in rituals.

Yajurveda: Provides details about sacrificial ceremonies and their procedures.

Atharvaveda: Offers insights into practical aspects of life, including health, medicine, and daily practices.

Despite their historical and cultural significance, the intricate linguistic structure of Sanskrit makes the Vedas challenging to interpret without specialized knowledge.

Importance of Preserving Cultural Heritage

The Vedas are not just religious texts but also a repository of wisdom that has shaped Indian civilization. However, with Sanskrit losing its prominence as a commonly spoken or studied language, a significant portion of this heritage risks being forgotten. Preserving the Vedas through translations into accessible languages like Hindi and English ensures that their teachings remain relevant and relatable in the modern era. This initiative contributes to cultural preservation, allowing individuals to connect with their heritage and fostering a deeper understanding of ancient Indian traditions.

Proposed Mechanism

Vedify employs a combination of modern technologies to deliver accurate and meaningful translations:

Natural Language Processing (NLP): Techniques like tokenization, lemmatization, and syntactic analysis are used to process the structured Sanskrit text. These methods break down complex sentences, enabling accurate translation.

Machine Learning (ML): The project utilizes machine learning models trained on carefully curated Sanskrit, Hindi, and English datasets. These models are designed to ensure translations preserve the context, meaning, and cultural nuances of the original texts.

Technologies Used: Frameworks such as Python, TensorFlow, and Hugging Face Transformers provide the backbone for NLP and ML tasks, ensuring precision in language processing and context-aware translation.

Dual-Language Outputs: The application generates translations in both Hindi and English, catering to diverse user groups.

User-Friendly Interface: The platform features an intuitive interface, enabling users to explore translations, search specific sections, and gain insights into the meanings of the Vedas.

Why Translate Sanskrit to Hindi

Hindi, as the most widely spoken language in India, serves as a bridge for millions to access the teachings of the Vedas. Translating Sanskrit into Hindi makes these texts comprehensible to a broad audience while retaining their cultural context. By using Hindi, Vedify ensures that the philosophical and spiritual messages of the Vedas reach individuals who might otherwise find Sanskrit difficult to grasp. This approach also supports national efforts to promote indigenous languages and cultural knowledge.

Why Translate Sanskrit to English

While Hindi is pivotal for reaching Indian audiences, translating the Vedas into English is equally important for connecting with global readers. English, as an international language, enables Vedify to share India's spiritual and philosophical heritage with the world. Many scholars, researchers, and enthusiasts worldwide have a deep interest in understanding the Vedas, but the lack of accessible translations poses a barrier. Providing English translations ensures that the universal messages of the Vedas—on harmony, self-realization, and the nature of existence— transcend cultural and geographical boundaries. By making the Vedas available in English, Vedify fosters global appreciation for India's ancient wisdom and promotes intercultural dialog.

LITERATURE SURVEY

Table 1: Comparative analysis of literature survey

Sr no.	SURVEY		
	Title And Author	**Year**	**Remarks**
1	Exploring the Role of Vedic	2023	Advantages:
	Literature in Education: A Comparitive Study of Acient Vedic Texts and Modern Educational System-Ambia Nargis Banu		-Moral Development -Linguistic Skills -Cultural Awareness Disadvantages: -Rigidity -Limited Scope -Cultural Exclusivity

2	Vedic literature- A significant literature of ancient India: An introduction- Sangeeta Roy	2016	Advantages: Rich Historical Insight -Foundation of Hindu Philosophy -Diverse Literary Forms -Cultural Significance -Preservation of Ancient Knowledge Disadvantages: -Limited Archaeological Evidence -Complexity of Interpretation -Accessibility Issues -Potential for Misinterpretation
3	Text to Speech Synthesis:A Systematic Review, Deep Learning Based Architecture and Future Research Direction- Fahima Khanam, Farha Akhter Munmun, Aloke Kumar Saha	2022	Advantages: -Naturalnessof Speech -Accessibilityfor Users -Versatile Applications -Deep Learning Performance -Comprehensive Taxonomy Disadvantages: -Data Dependency -Language-Specific Limitations -Evaluation Challenges -Computational Resources -Implementation Complexity
4	Text to Speech Conversion- D.B.K. Kamesh, Radhikla Rani	2016	Advantages: -Versatile Functionality -Open Source Tools -Multi-Language Support -Accessibility Enhancement -Cost-Effective Solution Titles of Disadvantages: -Accuracy Limitations -Environmental Dependencies -Setup Complexity -Performance Variability Hardware Constraints

PROPOSED MODEL

This section contains a diagram which has the proposed model of our application vedify. The activity diagram for our application is shown in this section.

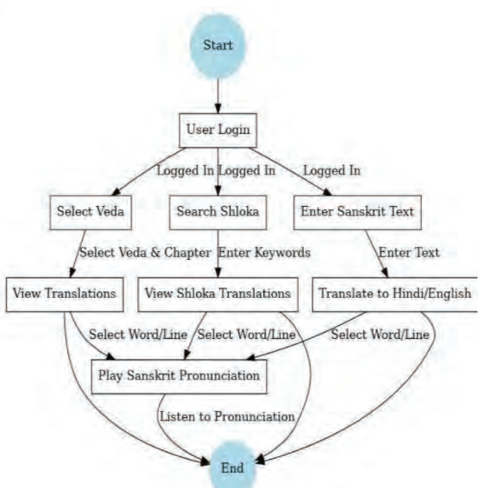


Fig 1. Activity Diagram for Vedify

Reading Translated Vedas

At the start, the user is prompted to choose what they'd like to do. If they opt to read a translated Veda: They are presented with the four Vedas: Rigveda, Samaveda, Atharvaveda, and Yajurveda.The user selects one of the Vedas, which is then provided in a translated form (likely in English or another preferred language).Once the translated content is delivered, the process ends.This feature makes the ancient Vedic texts accessible to users who might not be familiar with Sanskrit. It's a valuable tool for anyone interested in exploring the wisdom of these scriptures in a language they understand.

Sanskrit Pronounciation

If the user selects the Sanskrit pronunciation option. They can input any word or phrase in Sanskrit that they want help pronouncing. The system, using a specially trained dataset, processes the input and provides an audio pronunciation. After the pronunciation is delivered, the process concludes. This functionality is incredibly useful for those learning Sanskrit or aiming to recite shlokas and verses accurately. It ensures that users can focus on mastering the correct pronunciation, an essential aspect of engaging with Sanskrit texts authentically.

Shloka Translation

If the user chooses the shloka translation feature: They can enter a shloka, which is a verse or couplet written in Sanskrit. The system then performs a detailed contextual analysis using Natural Language Processing (NLP) techniques. This step is critical since Sanskrit is a complex language, where meanings can change depending on the context.After analysis, the system translates the shloka into English (or another preferred language) and provides the output. The process concludes once the translation is delivered.

This feature is designed to help users better understand the deeper meanings of Sanskrit verses. It ensures that the translation considers the context, making it as accurate and meaningful as possible.

WORKING OF THE PROPOSED MODEL

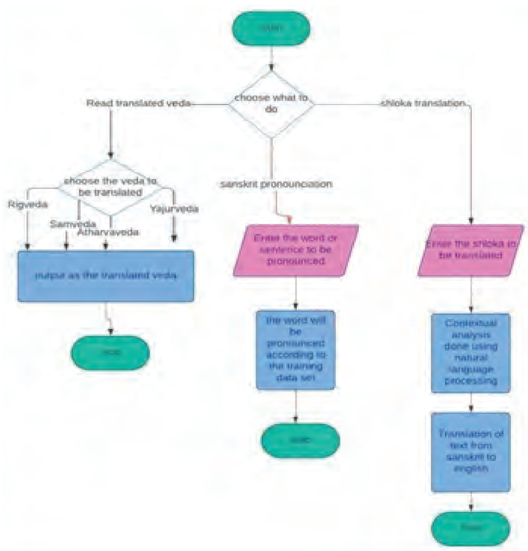


Fig 2: Flowchart for Vedify

Algorithm

User Authentication & Profile Management

To provide a secure and personalized experience, the app enables users to either sign up or log in.

- New Users:

Select the "Register" option and input your name, email address, and password.

Once registered successfully, you'll receive a confirmation message that reads, "Registration Successful!"

- Returning Users

Choose "Login" and enter your username and password. If the credentials match, you'll be granted access to explore the app.

If there's an error, the system will display a helpful message, "Invalid credentials," prompting you to check your details.

Veda Selection and Reading Interface

Immerse yourself in the wealth of Vedic knowledge with an easy-to-navigate reading feature.

- At the start, you'll see a list of the four Vedas: Rigveda, Samaveda, Atharvaveda, and Yajurveda. Select the one you wish to explore.
- Once chosen, the chapters in the selected Veda will be displayed. Pick a chapter of interest to proceed.
- The selected chapter, along with its translation, will then load on the screen.

You can enhance your experience with additional features like:

- Switching the language (e.g., Sanskrit, Hindi, or English).
- Bookmarking sections for quick access in the future.
- Highlighting important parts of the text to make your reading more engaging and tailored to your preferences.

Contextual Sanskrit to Hindi/English Translation

Unlock the meaning of Sanskrit texts effortlessly with this advanced translation feature.

- Enter or paste the Sanskrit text you wish to translate.
- The app uses Natural Language Processing (NLP) to analyze the context and meaning of the text before translating it.
- If the requested translation already exists in the app's database, it will appear instantly.
- For new texts, the app's translation engine will process the input and provide a precise result.
- Alternative translations may also be presented, offering a deeper understanding of the text.

This feature ensures that translations go beyond literal meanings, preserving cultural and contextual nuances.

Pronunciation Guide for Sanskrit Words

Master the pronunciation of Sanskrit words and phrases with ease!

- Enter the word or phrase you want to learn how to pronounce.
- The app will search its database for the corresponding audio file.
- If an audio file is available, you can:
  - o Play it as many times as needed.
  - o Adjust the playback speed to match your learning pace.
- If the pronunciation isn't in the database, the app will notify you politely with a message: "Audio not available."

This feature is an excellent resource for anyone looking to refine their pronunciation skills, whether for reciting verses, studying, or personal interest.

RESULTS AND COMPARISON

Fig 3

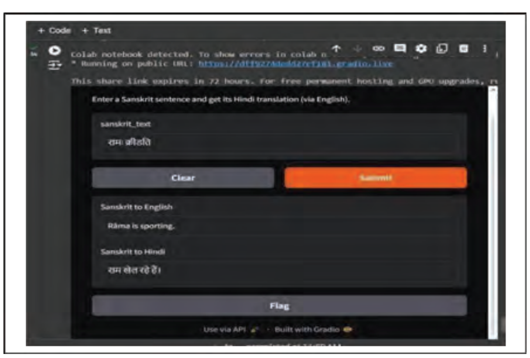


Fig 4

Fig 5

Table 2: Comparison with Existing systems

Sr no.	COMPARISON		
	Tool/application	*ACCURACY*	*SPEED*
1.	Vedalingo	Accuracy is 80%, struggles with classic grammar.	Moderate, requires for detailed parsing.
2.	Google Translate	Moderate for basic texts; struggles with classical grammar and idioms.	Very fast (instant translations).
3.	Microsoft Translator	Moderate; better with context in basic sentences.	Very fast (instant translations).
4.	Indic Transliteration Tools (e.g., Aksharamukha)	Low for translations; focuses on script conversion.	Very fast for transliteration

Future work

1. Our app provides a range of tools designed to help users explore Sanskrit texts effortlessly.
2. It enables reading and translating all four Vedas into Hindi and English while maintaining the contextual accuracy of even the most complex verses.
3. Advanced algorithms ensure that subtle meanings are preserved, staying true to the original essence of the texts.
4. A pronunciation feature allows users to hear the correct pronunciation of Sanskrit words, making it easier to learn and recite accurately.

CONCLUSION

From a technical standpoint, the project is highly feasible, utilizing advanced technologies to meet its objectives effectively. Economically, it creates opportunities for revenue through premium features, collaborations with educational institutions, and partnerships with cultural organizations. The system is designed for scalability and adaptability, ensuring long-term sustainability. Legal concerns are addressed with proper content licensing and robust data privacy measures, making it a trustworthy and dependable platform.

The Vedic Translation and Pronunciation System is a groundbreaking digital initiative that simplifies access to ancient Sanskrit texts, especially the Vedas. Featuring a user-friendly interface, it provides accurate Sanskrit-to-Hindi and English translations, a pronunciation guide, and an intuitive Shloka search feature. By leveraging advanced NLP and machine learning technologies, the platform ensures translations are both authentic and contextually accurate.

This system empowers users, including those unfamiliar with Sanskrit, to delve into the timeless wisdom of the Vedas while fostering proper phonetic recitation. Scalable, economically sustainable, and compliant with legal standards, it also supports educational and cultural partnerships. Beyond bridging the gap between ancient knowledge and modern audiences, the platform lays the groundwork for future expansions, such as additional texts and advanced learning tools, encouraging global appreciation of the spiritual, philosophical, and historical heritage of the Vedas.

ACKNOWLEDGMENT

We extend our heartfelt gratitude to the faculty of the Computer Engineering Department at K C College of Engineering, Management Studies, and Research, Thane, for their invaluable support and encouragement throughout this project. Their active involvement and expert guidance were instrumental in the successful completion of this paper. We are especially thankful for their dedication to fostering a collaborative and inspiring learning environment.

REFERENCES

1. K. U. Rao, P. V. Viswanath, and S. Sharma, “Real-Time Conversion of Vedic Sanskrit Texts into Multilingual Formats for Cultural Tourism,” presented at the IEEE International Conference on Tourism and Culture, vol. 2, pp. 112–118, April 2023.
2. Smith, Machine Learning for Ancient Languages: A Survey, 1st edition, vol. 1. Cambridge, MA: MIT Press, 2021, pp. 50–75.
3. Gupta, R. Mehta, and A. Kapoor, “Recognizing Vedic Texts with OCR and SVM Algorithms for Translation,” published in the International Journal of Research in Modern Engineering and Technology Sciences, vol. 5, no. 2, pp. 345–360, May 2024.
4. Kumar and D. Roy, “Framework for Translating Ancient Texts Using Artificial Intelligence,” unpublished manuscript.
5. B. Patel, “The Digital Evolution of Sanskrit Studies: From Grammar to Algorithms,” Journal of Ancient Studies and Algorithms, forthcoming.
6. J. Taylor and L. Watson, “Using Deep Neural Networks to Restore and Attribute Ancient Texts,” Nature, vol. 415, no. 7, pp. 123–130, February 2022.
7. P. Chakrabarti, Vedic Sanskrit: The Path to Digitization, 1st edition. New Delhi: Academia Press, 2021, pp. 20–45.
8. R. Sharma and A. Dubey, Sanskrit to Hindi Translation through Computational Linguistics, 1st edition. New Delhi, India: Prentice Hall, 2022, pp. 75–105.
9. P. N. Mishra, Bridging the Gap: Translating Sanskrit Texts into Modern Indian Languages, Varanasi, India: Banaras Hindu University Press, 2020, pp. 120–160.
10. K. Raj and S. Das, “Machine Translation of Sanskrit Texts to Hindi: Challenges and Progress,” Journal of Computational Linguistics in Indian Languages, vol. 14, no. 3, pp. 45–58, August 2023.
11. V. N. Iyer and M. Kapoor, “Translating Sanskrit Using Hybrid AI Models,” International Journal of Sanskrit Computing, vol. 8, no. 2, pp. 10–25, June 2022.
12. R. Gupta and H. Joshi, “Preservation of Vedic Texts: Hindi Translations Using Rule-Based Techniques,” Journal of Indian Heritage Studies, vol. 9, no. 1, pp. 85–100, January 2024.
13. S. Verma and P. Sharma, “AI-Driven Translation of Sanskrit Manuscripts into Modern Indian Languages,” presented at the 5th International Conference on Heritage and AI, vol. 3, pp. 89–94, October 2023.
14. M. Jain and K. Rao, “Neural Machine Translation for Sanskrit to English and Hindi,” in Proceedings of the 12th Workshop on Indian Languages and AI, pp. 112–119, December 2023.
15. T. Roy, “Multilingual AI Models for Sanskrit Translation,” Journal of Indian AI Research, forthcoming.
16. P. Narayan, “Translating Sanskrit to English: Decoding Ancient Texts,” Journal of Language and AI Studies, forthcoming.

PID Controller for a Process that Doubles in Size using the Coefficient Diagram Method (CDM)

Sowmya Kini M
Department of Electrical Engineering and
Atharva College of Engineering
Malad, Mumbai, Maharashtra
✉ sowmyakini@atharvacoe.ac.in

Gauri Vaidya
Department of Electronics and Telecommunication Engg
Atharva College of Engineering
Malad, Mumbai, Maharashtra
✉ gaurivaidya@atharvacoe.ac.in

ABSTRACT

Because of their inherent reaction qualities, double integrating processes are difficult to regulate. The Ball & Beam system, a common benchmark piece of laboratory equipment, is an example of a double-integrating process with strong nonlinear unstable properties that offers important insights into the nature of response. A PID controller design is one method that is frequently used in process control to stabilize the closed-loop control system. However, there aren't many methods for a double-integrating process where a PID can be easily applied. One such strategy that was developed using algebraic equations is the Coefficient Diagram Method (CDM). It is rather easy to create a controller that is reliable enough to stabilize the system and enhance time domain response characteristics by integrating CDM.

This straightforward but unique technique is applied to the Ball & Beam system in order to analyze the system dynamics and determine a CDM-PID control structure. The Euler-Lagrange approach is used to create the mathematical model of the benchmark laboratory system, and the control algorithm is then applied to it. Both the setpoint tracking and disturbance rejection capabilities of the system are tested. In order to compare the outcomes and assess the controller's performance, the Ziegler-Nichols conventional PID tuning approach is also used. It is demonstrated that in terms of stability and steady-state error reduction, the CDM-PID controller exhibits superior response characteristics.

***KEYWORDS** : Ball & beam system, CDM-PID, Double-integrating process, ZN-PID.*

INTRODUCTION

Process controller design has its own set of difficulties and restrictions. PID controllers are commonly utilized for First Order Processes with Dead Time (FOPDT), but applying the same tuning technique to higher order processes presents considerable difficulties. When a basic PID structure was introduced, the process of creating controllers improved by incorporating sophisticated tuning methods for higher order systems. To minimize or eliminate steady state mistakes and adapt to the dynamics of higher order systems, the controller can be fine-tuned utilizing sophisticated techniques. The process used to put these algorithms into practice either uses model-based approaches or entirely relies on non-model-based approaches.

The benefit of employing these techniques is that they offer tuning values directly, making them easy to use. It is clear from the literature that model-based tuning techniques are more dependable than their non-model-based equivalents. Therefore, model-based controller design has always been the main focus for industrial processes and equipment, whether they are integrated in nature or unstable. Nevertheless, there is a dearth of research on model-based unstable systems, including the Ball & Beam Process, which is a double integrating transfer function.

A strong control technique can be used to stabilize such a system and improve control response. The CDM Method is one such method that comes close to a robust control strategy. From classical control theory, this has been widely used to generate contemporary control methods. The method outlined in the literature that is currently available outlines a simple implementation. The controller and the process's characteristic polynomial are first developed separately, and the coefficients for the controller design are established in accordance with the intended characteristics response. The system's stability, transient response characteristics, and resilience can all be inferred from the coefficient diagram. A balance can be struck as a trade-off to modify the controller's parameters using this information.

Since the Coefficient Diagram Method (CDM) depends on the system's polynomial, cancellations of poles and zeroes won't be a problem. This method combines the state-space and transfer function representation of a system. It is possible to arrive at a controller with zero overshoot in the prescribed settling time since the controller design primarily depends on the stability index and the equivalent time constant of the system. In order to verify setpoint tracking and disturbance rejection capabilities on benchmark process control equipment, including industrial heat exchangers, temperature control systems, inverted pendulums, and various other systems, CDM-based controllers have been used.

Since the outcomes in each of the aforementioned situations were satisfactory, an effort has been undertaken to create a CDM-based controller for an unstable double integrating process. A double integrating system with unstable characteristics is accurately represented by a ball and beam laboratory benchmark system. As a result, the CDM-based PID controller has been implemented using this approach.

The following structure has been used to arrange the work that is discussed in the following sections: The CDM-PID controller design is explained in Section II. The simulation results are highlighted in Section III, and the work's conclusions are wrapped up in Section IV.

Mathematical Modelling of the system

As mentioned in the previous system, the ball and beam system is shown in Figure 1. A base, load gear connected by a lever arm, a stainless steel ball, and a supporting foundation make up the system. The base is connected to one end of the fixed beam, while the movable lever arm is connected to the other end. A DC motor coupled with a pulley belt system regulates this movement. The ball's speed is dictated by inclination and gravity, and it rolls lightly on the beam. It is possible to change the angle of the beam by giving the DC motor electrical input.

The inclination of the beam is altered by the rotation of the DC motor, and by modifying the signal transmitted to the DC motor, we may alter the ball's necessary stay position. In each instance, the tilted beam accelerates the ball in a precise proportion to the tilt angle. This acceleration causes the ball to roll over the entire length of the beam. The positioning data of the ball over the entire length of the beam is continuously updated in the controller through feedback. The feedback element of the system is an incremental rotary optical encoder, which updates the location of the motor shaft.

To determine the ball's current position over the beam, a potentiometer is also employed as a sensor. The controller receives the two measurements and takes the appropriate tracking and error-reduction measures.

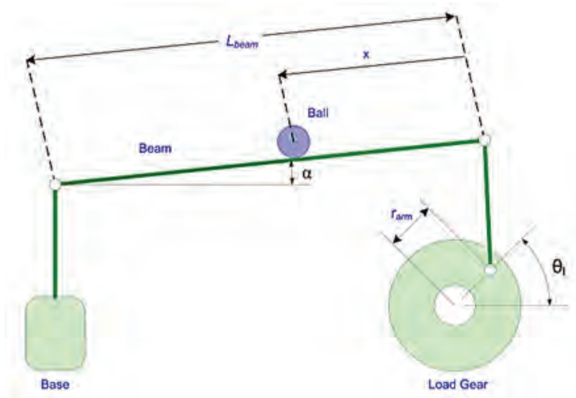


Fig. 1: Benchmark Laboratory Ball & Beam System

Maintaining the ball location at a single point over the beam on which it rolls freely is the major goal of the controller's design. To keep the ball's position at a single spot along the length of the beam, a controller automatically modifies the inclination angle. Its acceleration will alter to correct the position when the beam angle is changed by the controller.

However, this change in acceleration still allows us to indirectly control the ball's position. If the system is observed, we can conclude that the structure is an unstable open loop, meaning that the system output—in this case, the ball's position—varies without regard to a fixed type of input, such as the beam's inclination. In order to keep the ball in the desired location for the whole length of the beam, this system requires a feedback element. A controller must be created based on the dynamics of the system in order to continuously monitor the target position and the existing position.

Table 1: Ball and Beam System Notations

Notation	Description	Quantity/Value
L_{beam}	Length of Beam Ball Mass	42.55 cm
m_b	Mass of the ball	0.064 kg
r_b	The ball's radius	1.27 cm
r_{arm}	The separation between the coupled joint and the output gear shaft	2.54 cm

g	Gravity-induced acceleration	9.8 m/s^2
J_b	The ball's moment of inertia	$\frac{2}{5} m_b r_b^2$ Kgm^2
α	Angle of beams	-
θ	Angle of the load gear	-
R	The ball's position coordinate	-

The system dynamics are subjected to forces due to gravity, moment of inertia and the centrifugal motion. The equation can be simplified using the Lagrangian approach for the motion of the ball over the beam:

$$(m_b + \frac{J_b}{R^2})\ddot{r} + \frac{J_b}{R\ddot{\alpha}} - mr\dot{\alpha}^2 + mgsin\alpha = 0 \quad (1)$$

Now the equation (1) shall be linearized by taking the beam angle as sinα = α for the smaller values of α. The system would be linearly approximated using this approach.

$$(m_b + \frac{J_b}{R^2})\ddot{r} = mg\alpha \quad (2)$$

Now the arc distance can be equated to the gear angle as specified in the equation (3).

$$\alpha = (\frac{d}{L})\,\theta \quad (3)$$

Hence by substitution in equation (2), we get:

$$(m_b + \frac{J_b}{R^2})\ddot{r} = mg(\frac{d}{L})\,\theta \quad (4)$$

Using Laplace Transform, we can get an equation for the input as well as for the output which is given by:

$$(m_b + \frac{J_b}{R^2})s^2 r(s) = mg(\frac{d}{L})\,\theta \quad (5)$$

$$\frac{r(s)}{\theta(s)} = \frac{mgd}{L(m_b + \frac{J_b}{R^2})}\frac{1}{s^2} \quad (6)$$

The transfer function depicted in equation (5) is a function of the angle of gear θ(s) and the position of ball r(s). From the parameters mentioned in Table 1, we can substitute the value in equation (5) to arrive at:

$$\frac{r(s)}{\theta(s)} = \frac{0.7}{s^2} \quad (7)$$

By having a look at equation (7), its clear that the process is of double-integrating type which clearly depicts that its an unstable system which can pose significant challenges in designing a controller.

CDM-PID Controller Design

Figure 2 shows the block diagram that illustrates the CDM technique. Often referred to as a SISO process, the system represents a single input and a single output. Figure 2 illustrates this with y standing for the output, r for the setpoint, u for the controller output, and d for the process disturbance. A transfer function with the numerator (N(s)) and the denominator (D(s)) is how the plant has been represented. The forward denominator function is denoted by the term A(s), the reference function by the term F(s), and the feedback function polynomial by the term B(s).

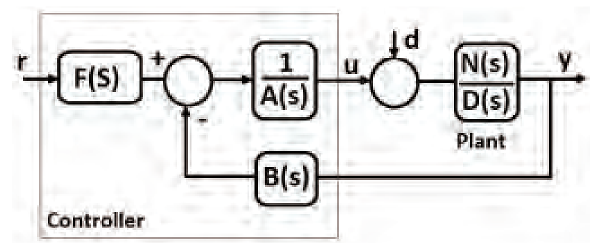


Fig. 2: Block Diagram of CDM

From the figure (2), the CDM based control system can be depicted as:

$$y = \frac{N(s)F(s)}{P(s)}r + \frac{A(s)N(s)}{P(s)}d \quad (8)$$

In the equation (8), P(s) is a characteristic polynomial given by the equation:

$$P(s) = A(s)D(s) + B(s)N(s) = \sum_{i=0}^{n} a_i s^i \,, a_i > 0 \quad (9)$$

Where A(s) and B(s) are defined as:

$$A(s) = \sum_{i=o}^{p} l_i s^i \quad \& \quad B(s) = \sum_{i=o}^{p} k_i s^i \quad (10)$$

The polynomial function F(s) can be determined as:

$$F(s) = \frac{P(s)}{N(s)} \; for\ s = 0 \quad (11)$$

Hence the closed loop system becomes of Type-1 process and a good overall response can be expected from the system. From the equation (7), the system under consideration is of a double integrating type. The generalized model of the system can be represented as:

$$G_p(s) = \frac{Y(s)}{U(s)} = \frac{K}{s^2} \quad (12)$$

The transfer function of the process is constituted from independent polynomials given by numerator (N(s)) and the denominator (D(s)) as CDM is a polynomial based approach.

$$N(s) = K \quad \& \quad D(s) = s^2 \tag{13}$$

In this method, the system is designed for step inputs, which can be represented using the controller polynomial as:

$$A(s) = I_1 s + I_0 \quad \& \quad B(s) = k_2 s^2 + k_1 s + k_0 \tag{14}$$

In equation (14), the controller design parameters are described by the terms l1, l0, k1, k2, & k0. Here the consideration is such that the disturbance is a step signal which is having an impact on the system. Thus, the controller is selected by choosing the term with l0=0.

$$G_c(s) = \frac{B(s)}{A(s)} = \frac{k_2 s^2 + k_1 s + k_0}{I_1 s} = \frac{k_2 s}{I_1} + \frac{k_1}{I_1} + \frac{k_0}{I_1 s} \tag{15}$$

Substituting (13) & (14) in equation (9) results in:

$$P(s) = (I_1 s)s^2 + (k_2 s^2 + k_1 s + k_0)K \tag{16}$$

By taking the values of time constant τ & stability indices $\gamma_i = \{\gamma_1, \gamma_2\}$, we get the characteristic polynomial as:

$$P_{target}(s) = \frac{\tau^3}{\gamma_2 \gamma_1^2} s^3 + \frac{\tau^2}{\gamma_1} s^2 + \tau s + 1 \tag{17}$$

By equating equation (17) & (16), we can find the controller parameters given by:

$$I_1 = \left(\frac{\tau^3}{\gamma_2 \gamma_1^2}\right), k_2 = \left(\frac{\tau^2}{\gamma_1 K}\right), k_1 = \left(\frac{\tau}{k}\right) \& \, k_0 = \left(\frac{1}{k}\right) \tag{18}$$

F(s), which represents the pre-filter element can be computed as given by:

$$F(s) = \frac{P(s)}{N(s)} \; for \; s = 0 = \frac{P(0)}{N(0)} = \frac{Kk_0}{K} = k_0 \tag{19}$$

With the above stated approach, the CDM-PID control system can be developed as shown in Figure 3. The PID controller incorporated using the CDM approach is given by C(s) which acts as the main controller and a feed forward controller has been depicted as Cf(s). The main controller depicted by C(s) can be expressed as:

$$C(s) = K_c\left(1 + \frac{1}{T_i s} + T_d s\right) \tag{20}$$

For determining the CDM variables, the CDM-PID block diagram given in Figure 3 has to be further modified to equivalent as shown in Figure 4.

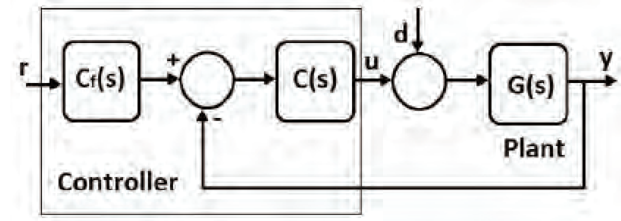


Fig. 3: Block Diagram of CDM-PID Control System

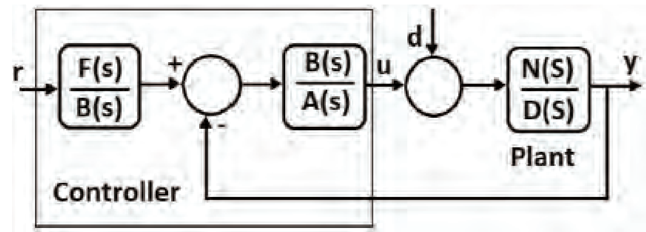


Fig. 4: Equivalent Block Diagram of CDM-PID System

In comparison between Figure 3 and Figure 4, the following connections can be established.

$$C(s) = \frac{B(s)}{A(S)} \tag{21}$$

$$C_f(s) = \frac{F(s)}{B(S)} \tag{22}$$

By substituting equation (14) & (20) in equation (21) leads to:

$$\frac{k_2 s^2 + k_1 s + k_0}{I_1 s} = K_c\left(1 + \frac{1}{T_i s} + T_d s\right) \tag{23}$$

After comparing the like power terms in the equation, the controller parameters of the CDM-PID controller can be obtained as given by:

$$K_c = \frac{k_1}{l_1}: \; T_i = \frac{k_1}{k_0}: T_d = \frac{k_2}{k_1} \tag{24}$$

After substituting equation (14) and (19) in equation (22), we get the forward controller as:

$$C_f(s) = \frac{F(s)}{B(S)} = \frac{k_0}{k_2 s^2 + k_1 s + k_0} \tag{25}$$

Using the above equations, we can arrive at the CDM-PID controller for a ball and beam system. Using the model parameters described in equation (7) and by selecting the time constant τ=0.13 and index of stabitlity γ_1=2.5 & γ_2=2, the controller parameters can be obtained by substitution in equations (24) & (25).

$K_c = 1055:\ T_i = 0.14\,s : T_d = 0.58\,s$ &

$$C_f(s) = \frac{F(s)}{B(S)} = \frac{1.4}{0.01s^2 + 0.19s + 1.5} \tag{26}$$

SIMULATION RESULTS AND DISCUSSION

The response of the controller parameters that were derived in the preceding section using the CDM-PID approach was simulated, and Figure 5 illustrates this for various ball position operating points. Additionally, the graphic compares the traditional ZN-PID method with the CDM-PID method. Table 2 displays the PID values that were acquired using the ZN-Method.

Table 2: PID controller parameters using ZN Method

$K_c = 0.6K_u$ ($K_u = 30$)	$T_i = 0.5P_u$	$T_d = 0.125P_u$
19	2.12	0.54

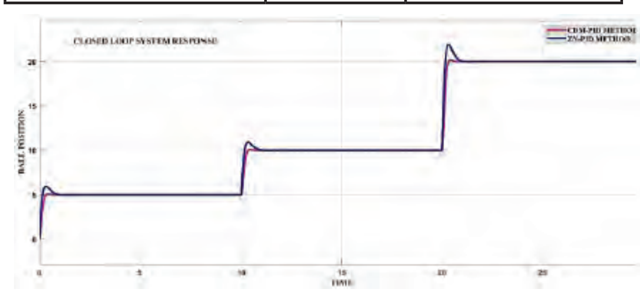


Fig. 4: Closed Loop Response of the System at 5 cm, 10 cm and 20 cm ball-position

It is clear by examining the closed loop response that the CDM-PID outperforms the traditional ZN-PID approach in terms of response characteristics. Better setpoint tracking capabilities for servos and regulatory tracking are represented by the CDM-PID controller. When compared to the ZN-PID Control approach, the closed loop response showed that the system had the fastest rise time and the time needed for settling, together with a very low overshoot. It is also evident that the system's robust response to the CDM-PID controller approach stayed constant across various operating points. Table 3 displays the error indices that were computed for the various ball positions from the response.

Performance Measure	CDM PID			ZN-PID		
	05 cm	10 cm	20 cm	05 cm	10 cm	20 cm
Rise Time (s)	0.33	0.33	0.33	.029	.029	.029
Settling Time (s)	1.32	1.35	1.41	3.26	3.35	3.47
% Overshoot	0.5	0.5	0.5	14.2	14.2	14.2

From the table it is clear that the error indices are minimum for the CDM-PID controller. From the values of error indices, we can confirm that the CDM-PID controller is capable of handling the disturbance rejection to provide reasonable consistent performance. This was verified by introducing a disturbance into the system at 20th second as shown in Figure 7.

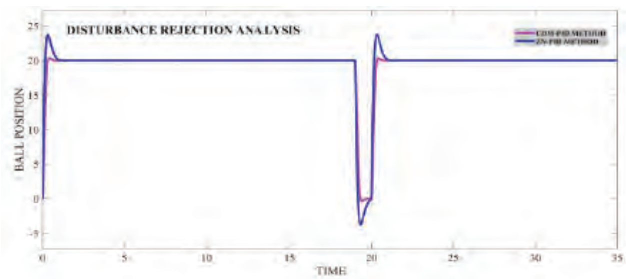


Fig. 5: Disturbance rejection plot for ball position at 20 cm

It is evident from the reaction shown in Figure 7 that the CDM-PID modified the controller output to reduce the disturbance and return the system to its stable state with a modest peak in a shorter period of time. Conversely, the non-linearities significantly suppress the ZN-PID approach, demonstrating the CDM-PID method's ability to manage control system disturbances.

CONCLUSION

For the ball and beam system, which constituted an unstable double-integrating system, a CDM-PID controller was created. Servo and regulatory response mathematical modeling was used to test the controller's effectiveness. In contrast to the traditional ZN-PID method, the controller was able to deliver comparatively superior performance by properly tuning it by fixing on a time constant and stability index. In contrast to the ZN-PID approach, the CDM-PID based controller's response for the closed loop system offered the quickest rise time, settling time, and least amount of overshoot. In order to improve the overall system response, it is suggested that the CDM-PID approach be applied to the unstable double integrating process.

REFERENCES

1. B. Meenakshipriya, K. Kalpana, Modelling and control of ball and beam system using coefficient diagram method (CDM) based PID controller, IFAC Proc. Vol., 47 (2014), pp. 620-626
2. Aidan O'Dwyer (2006), Handbook of PI-PID controller tuning rules, Imperial college Press.
3. Astrom,K.L., & Hagglund.T., (2001). The future of PID control, Control Engineering Practice, 9,1163-1175.
4. Erkan IMAL (2009), CDM based controller design for nonlinear heat exchanger process, Turk J Electrical Engineering and Computer Science, 17(2).
5. Meenakshipriya.B., Saravanan.K., Krishnamurthy.K., & Bhaba.P.K., (2012), Design and implementation of CDM-PI control strategy in pH neutralization system, Asian Journal of Scientific Research, 5(3), 78-92.
6. Mohammad Keshmiri., Ali Fellah Jahromi., Abolfazl Mohebbi., Mohammad Hadi Amoozgar., & Wen Fang Xie., (2012), Modeling and Control of Ball and Beam system using Model Based and Non-Model Based Control Approaches, International Journal on Smart Sensing and Intelligent system.
7. S.U. Din, Q. Khan, F.-U. Rehman, R. Akmeliawanti A comparative experimental study of robust sliding mode control strategies for underactuated systems IEEE Access, 5 (2017), pp. 10068-10080.
8. Saeed Zaare, Mohammad Reza Soltanpour, The position control of the ball and beam system using state-disturbance observe-based adaptive fuzzy sliding mode control in presence of matched and mismatched uncertainties, Mechanical Systems and Signal Processing, Volume 150, 2021, 107243, ISSN 0888-3270.
9. W. Yu, Nonlinear PD regulation for ball and beam system, Int. J. Electr. Eng. Educ., 46 (2009), pp. 59-73.
10. P.M. Maalini, G. Prabhakar, S. Selvaperumal Modelling and control of ball and beam system using PID controller in 2016 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT) (2016), pp. 322-326.
11. O. Castillo, E. Lizárraga, J. Soria, P. Melin, F. Valdez, New approach using ant colony optimization with ant set partition for fuzzy control design applied to the ball and beam system Inf. Sci., 294 (2015), pp. 203-215.
12. R. Moezzi, V.T. Minh, M. Tamre, Fuzzy logic control for a Ball and Beam system Int. J. Innov. Technol. Interdiscip. Sci., 1 (2018), pp. 39-48

Optimizing EMG Signal Analysis with Advanced Machine Learning Techniques

Shashidhar R
Dept of Electronics and Communication Engineering
JSS Science and Technology University
Mysuru, Karnataka
✉ shashidhar.r@jssstuniv.in

Rujwal K
Dept of Electronics and Communication Engineering
JSS Science and Technology University
Mysuru, Karnataka
✉ rujwalprince@gmail.com

Natasha Naik
Department of Electronics and Computer Science
Atharva College of Engineering
Mumbai Maharashtra
✉ natashanaik@atharvacoe.ac.in

Ruchitha Y
Dept of Electronics and Communication Engineering
JSS Science and Technology University
Mysuru, Karnataka
✉ ruchithay163@gmail.com

ABSTRACT

In this paper, electro-myography is used to study human muscle movement and bionic hands. It comprises complicated signal processing, involving acquisition, pre-processing, feature extraction, and classification. A new feature extraction method was proposed for the study using correlation heat maps, which, in conjunction with PCA and SVM feature elimination, showed superior classification accuracy compared to traditional methods. The study investigated the possibility of using a Decision Tree algorithm to classify electromyography signals on feature extraction with correlation heat maps. Compared to traditional techniques, decision trees demonstrated better accuracy in classification. The findings have implications for describing its effectiveness in enhancing the performance of both EMG signal analysis and gesture recognition. The Decision Tree classifier achieved a training accuracy of 94 percent with 6 percent of error, using features extracted with the Correlation Heat Map method. In comparison, the model trained with KNN-extracted features achieved a training accuracy of 75.4 percent.

KEYWORDS *: KNN, EMG, Correlation heat map, PCA and SVM.*

INTRODUCTION

Electromyography is one of the kinds of bioelectric signals which are extracted from human muscle contraction by a certain method from the surface of the skin. EMG signals control the interface and the information devices obtained in two ways: embedded and surface-based. In the existing methodology, they used different models like SVM, KNN, Random Forest for training purposes. PCA, CHM had been used for feature extraction and gaining 81.6 percentage of precision with feature extraction using the correlated heat map and the training model in the Random Forest method. The features of EMG signals can be extracted generally using time frequency analysis, generating more detailed information. STFT and Wavelet Transform and Wavelet Packet Transform are techniques that create high dimensional feature vectors. This approach increases the complexity and hence is quite difficult to apply in real-time use in controlling prosthetic devices.

Here, the most important thing will be to extract the features from EMG signals and enhance the classification accuracy from multiple channels of complex hand motions. Hence, the feature extraction method of the correlation heat map method is introduced here, which may determine the correlation be- tween variables by the correlation coefficient between them. If it is smaller, the degree of correlation between variables is also small. These can be classified with different techniques such as K-Nearest Neighbors(KNN), Random Forest(RF), and Support Vector Machines(SVM), each guaranteeing a certain accuracy. In the KNN case, classification is made by the proximity of the data points in the feature space. Random Forest improves classification performance by using an ensemble of decision trees. SVM finds the optimal boundary separating the different gestures. Of course, each has its strengths, and different levels of accuracy can be achieved subject to the specifics of a data-set and applied feature extraction processes. The datasets has been taken from UC Irvine Machine learning repository. The problem involves classifying hand gestures using EMG signals, which includes distinguishing between states such as hand at rest, wrist flexed, wrist extended, and extended palm. Achieving high classification accuracy requires employing various machine learning methods, includ- ing Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Random Forest (RF), and Decision Tree methods to enhance performance. The challenge lies in optimizing these techniques to accurately classify gestures, with a particular focus on applications like converting sign language into text. Addressing this problem has significant implications for im- proving communication and accessibility for individuals with hearing impairments, underscoring the importance of precise EMG signal analysis. Objectives of our paper are as follows:

1. Enhancing EMG signal processing to improve the analysis of muscle movements and applications like bionic hands and prosthetic limbs. 2. Develop and compare feature extraction methods like PCA and Correlation Heat Map to identify the most accurate and representative signal channels. 3. Evaluate and compare the accuracy of the Correlation Heat Map versus PCA in extracting meaningful features from EMG signal. 4. Apply classification algorithms, including Decision Tree and Regression methods,to classify EMG signals and assess their performance. 5. Identify the optimal combination of feature extraction methods and classification algorithms to surpass traditional methods in classification accuracy. 6. Integrate EMG signals with virtual reality to enhance human-computer interactionby accurately identifying gestures and emotional responses.

RELATED WORK

This literature survey examines advancements in hand gesture classification using Electromyography (EMG) signals, focusing on a novel approach with the MYO Thalmic bracelet. The study highlights the successful application of an ensemble of decision trees to classify static hand gestures with high accuracy (98.0 percent validation). It reviews methods for capturing and analysing EMG signals, emphasizing the significance of accurate gesture recognition in human-computer interaction. Key aspects include dataset characteristics, recording techniques, and model performance, underscoring ongoing innovations in EMG- based gesture classification [1]. Prototype systems utilizing surface EMG (sEMG) for facial expression recognition demonstrate the efficacy of integrating multivariate classifiers with cloud-based processing for real-time data analysis. To provide remote visualization and classification of EMG Signals the system controls eight-channel measurement devices and Wi-Fi modules [2]. By using electromyogram signals by exploring literature on hand gesture recognition, and explore how the correlation heat map make use of feature selection and visualization of signal relationships, and as well as decision tree employed to classification [3]. Upper limb amputation restricts significantly the amputees to do daily duties. The myoelectric prosthesis employs the impulses from surviving muscles in the stump to slow revival function to such severed limbs [4]. Modern algorithms and machine learning techniques enhance the detection of signal abnormalities and enable timely notifications for better patient management [5]. Here introduction to feature extraction method for selecting the best EMG channels from eight channel recordings, using support correlated heat maps. The proposed approach demonstrates superior classification accuracy compared to traditional methods, emphasizing the effectiveness of the new feature extraction technique [6]. Recent research highlights the effectiveness of integrating Virtual Reality (VR), utilizing physiological data and convolutional neural networks for detection. VR interventions, including pleasant videos and music, have been shown to significantly reduce stress levels [7]. EMG of humans are used for various uses such to control devices via Bluetooth or controlling robot arm [8]. This paper evaluates a proposed EMG sensor for IoT and exoskeleton applications by comparing its Signal to Noise Ratio (SNR) with commercial sensors it results in less appropriate for musculoskeletal monitoring or diagnosis [9].Muscles and nerve cells are examined with electromyography signals. Electrical signals are transmitted by these cells which cause muscle contraction and relaxation [10]. This is necessary to improve the performance of recent research into EMG-based hand gesture classification in terms of classifying accuracy, generalization ability, and robustness. Some recent advances have even dealt with the augmentation of techniques through the fusion of convolutional autoencoders with convolutional neural networks and windowing and majority voting data-processing [11]. Muscle activation signals form an important part of biomechatronics systems for control purposes; therefore, EMG-based motion classification can prove to be quite a challenge due to signal characteristic variability. In this study, machine learning methods with regard to the classification of wrist and hand gestures were evaluated and compared[12]. Electromyography signals, which are of prime importance in applications such as prosthetic control and motion detection, are normally classified using techniques such support vector machines and artificial neural networks. In this paper, a CNN is used to improve the accuracy of the classification of hand movements by applying deep learning methods to learn intricate relationships in data[13]. Effective characterization of the EMG signal in neuromuscular disorders is much dependent on machine learning-based pattern classification. In this paper, classifiers for EMG signal characterization will be reviewed considering both accuracy and computational efficiency. This paper summarizes state-of-the- art techniques and their main findings related to the advance- ment of neuromuscular pathology [14]. This paper contrasts the performance of ”Trainlm” and ”Trainscg” algorithms and finds that both of them have processing times less than 100 ms, while the ”Trainlm” algorithm is more accurate. Testing on five healthy subjects gave results that suggest these algorithms are efficient for hand motion recognition and could enhance volitional control of robotic assistive devices [15].

PROPOSED METHOD

This is the process of an experiment carried with pre-processing methods, feature extraction methods, and training model and finally we carried out validation process for the result and performed the methodology of project.

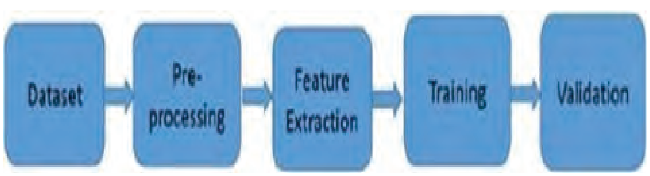


Fig. 1. Block Diagram of methodology

Feature Extraction

This refers to the process of extracting the specific features of the provided datasets which carried by the training model of machine learning. Correlation Heat map: A correlation heat map, also known as a correlation heated map (CHM), is a visual tool used to analyse and understand the relationships between multiple variables in a dataset and also useful for dimensionality reduction and improving classification accuracy.

The correlation coefficient X1X2 between two variables X1 and X2 is calculated using

$$\rho_{X_1X_2} = \frac{Co(X_1, X_2)}{\sqrt{D_{X_1}D_{X_2}}} = \frac{E[X_1X_2] - E[X_1]\cdot E[X_2]}{\sqrt{D_{X_1}D_{X_2}}} \quad (1)$$

Classification algorithm

Decision tree: They work on the basis of splitting of the data into subsets based upon the values of given input features, this leads to a tree-like structure where each node is represented as a decision based on a feature and every node in the tree represent a class label.

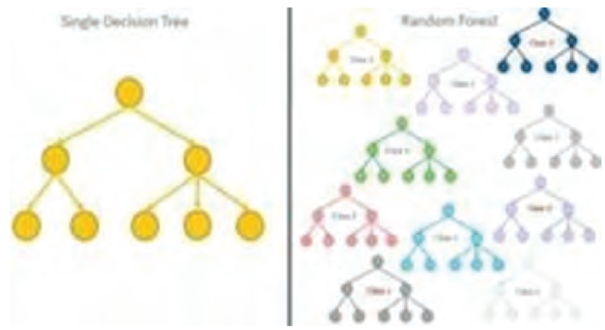


Fig. 2. Decision tree method

1. Start at the Root Node: The entire dataset is represented by the root node. The algorithm will make decisions based on the feature values to split the data into subsets. 2. The decision tree algorithm selects the feature that best separates the data into different classes. This is done using criteria such as Gini impurity, information gain, or variance reduction. 3. Based on the selected feature, the data is divided into subsets. Each subset is then used to create child nodes. 4. For each child node the process is repeated choosing the best feature to split on and dividing the data further, until stopping criteria are met. 5. Stopping criteria: The tree building process continues until one or more stopping criteria are reached, such as a maximum tree depth is reached, minimum number of samples is reached in a node and no further information gain can be achieved. 6. Once the tree is fully grown, each leaf node will represents class label. The class label assigned to a new instance is determined by following the path to leaf node from the root node, based on the feature values of the instance.

RESULTS AND DISCUSSIONS

The Decision Tree algorithm was used to classify hand gestures from electromyography (EMG) signals and trained the classification model on a dataset of EMG signals representing different hand gestures including hand at rest, wrist flexed, wrist extended, and extended palm. Then, split the dataset into training and testing sets to assess the Decision Tree classifier's performance.

Discussion

Performance comparison: By classifying hand gestures Decision Tree classifier demonstrate dynamic performance, and utilizing the Correlation Heat Map method for feature extraction with a 94 percentage of higher accuracy. This shows that the Correlation Heat Map arrives more effective in capturing and representing the closer features of EMG signals compared with PCA. B. Feature Extraction Impact: The better representation of relationships between different EMG signal is provided by Correlation Heat Map method, leading to improved classification accuracy. By analysing and visualizing the correlations between variables, the Correlation Heat Map approach facilitate reduced redundancy and better feature selection. PCA, despite the fact that effective in reducing dimensionality, it has removed some of critical information similar to the patterns and variability in the EMG signals, which can also explain the slightly lower accuracy.

C. Decision Tree Effectiveness: Decision Tree algorithm's high accuracy emphasize its relevant for complex patterns in EMG data by classifying . Its ability to handle non- linear relationships and make conclusion based on hierarchical splits contributed its effectiveness in difference with different hand gestures. The interpretability of Decision Trees proves especially advantageous in contexts requiring transparency in decision-making, as it facilitates insight into how feature splits influence classification outcomes. D. Implement for applications: The high accuracy achieved with the Correlation Heat Map method and Decision Tree classifier suggests that this approach is promising for applications requiring precise gesture recognition, such as prosthetic control and human-computer interaction. Future work could explore further optimization of feature extraction and classification methods to enhance performance, particularly in real-time or complex scenarios. The above one visualize the result of the experiment with descent accuracy of training and validation of datasets. This explains the approachable feature extraction method that is correlated heat map which extracts the features of different hand gestures from 8 different hand gestures. The confusion matrix explains the result of the trained model to predict the different hand gestures and classify them.

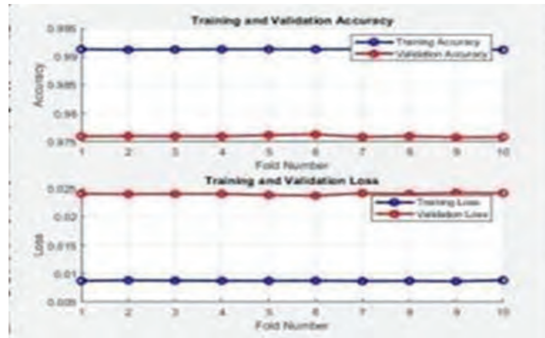


Fig. 3. Training and validation plot

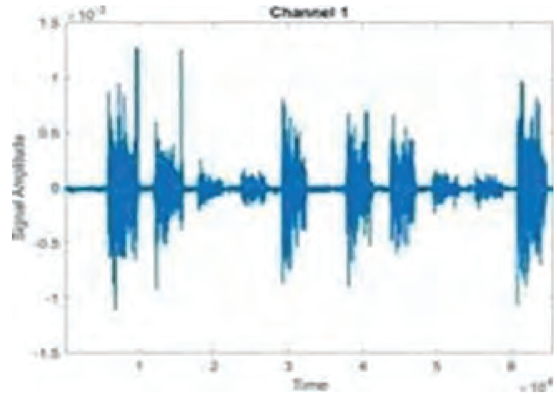


Fig. 4. EMG plot from channel

Fig. 5. Correlation Heat Map

Comparison Table

FUTURE SCOPE

From this study, at least three more promising leads were found that could further optimize in their pursuit for efficiency in EMG-based gesture recognition systems. The most immedi- ate direction where further fine-tuning is needed is with regard to the general approach in feature extraction techniques, or the Correlation Heat Map, per se, so that it can be optimized better and more effectively catch and represent the most relevant and discriminative EMG data. More advanced or hybrid methods of feature extraction may be attempted by researchers for increased accuracy on classification and the performance of the system overall. Further work would also include applying other machine learning models or even ensemble methods that may perhaps achieve better accuracy or more generalization in a variety of real-world conditions. Promising directions are along lines such as combining multiple classifiers or an approach through deep learning in order to further strengthen gesture recognition systems. Another challenge is to apply real-time, more practical uses of the research into prosthetics, rehabilitation, and human-computer interaction. The integration of this real-time classification along with feedback mechanisms may allow for immediate response to a detected gesture, further increasing user interactivity with assistive devices as well as interactive systems. It could also expand the scope of real-time applications concerning telemedicine, remote monitoring, and long-distance rehabilitation, if they are made via communication protocols. This brings about the possibility of the development of new systems that can offer particular functions based on each gesture rendered. In this regard, newly designed systems would now be used for recognizing any specific gesture and then activating the previously defined actions in the system. Such systems can be implemented in prosthetic devices, rehabilitation applications, and other interactive systems where personalized functionality becomes more important. Further development and refinement of these technologies will be essential for further progress in future EMG-based applications along with ideas about innovating neuromuscular rehabilitation, human-computer interaction, and more.

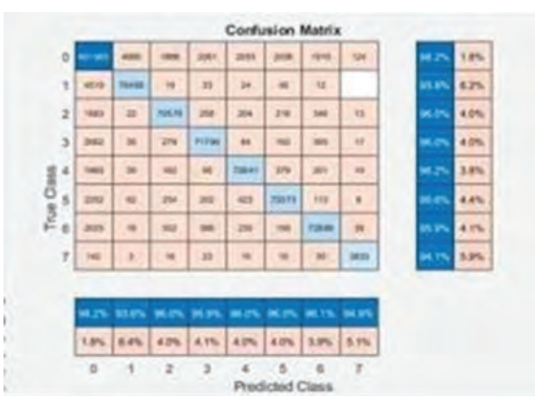


Fig. 6. Confusion Matrix

Table 1 Comparison of Proposed Work with Related Work

	Classification Random Forest [1]	Classification method KNN[1]	Proposed Decision Tree
Data set	*EMG datasets* forgestures from UCIrvine machine learning repository.	*EMG datasets* for gestures from UC Irvine machine learning repository.	*EMG datasets* for gestures from UC Irvine machine learning repository
Accuracy in Percent	*81.6*	*75.4*	*94*
Error Rate in Percent	*18.4*	*24.6*	*6*

REFERENCES

1. Chandran, D., Devaraj, S., Ashok, S., Shaji, D., Saju, S. K. (2024). Hand gesture classification using sEMG signals and ensemble learning. 2024 Third International Conference on Intelligent Techniques in Con- trol, Optimization and Signal Processing (INCOS). DOI: 10.1109/IN- COS59338.2024.1052763.

2. Mingzhe, J., Tuan, N. G., Arman, A., Rahmani, A.-M., Westerlund, T., Salantera, S., Liljeberg, P., Tenhunen, H. (2024). IoT-based remote facial expression monitoring system with sEMG signal. UTC from IEEE Xplore. DOI: 10.1109/CATCON52335.2021.967048.

3. Amlan, D. M., Amol, A., Rajesh, K. T. (2024). Automated recognition of hand gestures from multichannel EMG sensor data using time–frequency domain deep learning for IoT applications. IEEE Sensors Letters, 8(6).

4. Ramkumar, S., Dhanusha, Rema, Archana Devi, Elavarasi, K., Sel- vaganapathi, T., Gokila, S. (2023). A review on EMG-based pattern identification methods for effective controlling of hand prostheses. 3rd International Conference on Innovative Mechanisms for Industry Application. DOI: 10.1109/ICIMIA60377.2023.10426249.

5. Arunsankar, G., Rajeshkumar, S., Vanathi, A., Gopi, B., Sasikala, S., Srinivasan, C. (2023). IoT controlled device to manage the EMG signals of the patient and alert in real time. Second International Conference on Smart Technologies for Smart Nation. DOI: 10.1109/SmartTech-Con57526.2023.1039176.

6. Xiaoyan, Z., Mengru, Z., Ranjeesh, R. (2023). Study on the methods of feature extraction based on electromyographic signal classification. Medical and Biological Engineering and Computing.

7. Ahilan, M., Anline, R., Kumar, S. N., Muthu Kumar, B. (2023). Virtual reality sensor- based IoT embedded system for stress diagnosis. IEEE Sensors Journal, 23.

8. Minh, N., Tuan, N. G., Westerlund, T. (2021). EMG-based IoT system using hand gestures for remote control applications. 2021 IEEE 7th World Forum on Internet of Things (WF-IoT). DOI: 10.1109/WF-IoT51360.2021.959595.

9. Arturo Gonza´lez-Mendoza, Alberto Isaac Pe´rez – SanPablo, Ricardo Lopez-Gutierrez, Ivette Quinones "Validation of an EMG sensor for Internet of Things and Robotics" 2018 15th International Conference on Electrical Engineering, 5/7/2018

10. Shubham, S., Abhishek, K., Srijan, V., Sumathi, G. (2021). IoT enabled EMG monitoring health system with SASS based signal denoising . 2021 IEEE 5th International Conference on Condition Assessment Techniques in Electrical Systems.

11. Guangyu, J., Hak-Keung , L., Junkier, L., Rong, W. (2024). Classification of electromyographic hand gesture signals using machine learning techniques. Department of Engineering, King's College London.

12. Chingiz, S., Fedi, S., Erhan, A. (2020). Comparison of machine learning algorithms for EMG signal classification. Periodicals of Engineering and Natural Sciences, 8(2).

13. Kaan, B., Nalan, O. (2020). Classification of EMG Signals using convolutional neural network. International Open Access, 8(4).

14. Jamileh, Y., Andrew , H.-W. (2014). Characterizing EMG data using machine – learning tools. School of Computer Science, University of Guelph.

15. Yang, Z., Chaoyang, C., Juan, N., Guoxin, G., John, C.,Mark, C., Stephen, L. (2020). EMG signal processing for hand motion pat- tern recognition using machine learning algorithms. Archives of Orthopaedics Research Article.

Analyzing Network Intrusion Detection System Based on Machine Learning Algorithms

Kiruthika. B
Assistant Professor
Department of Electronics and Computer Science
Atharva College of Engineering
Mumbai, Maharashtra
✉ drkiruthika@atharvacoe.ac.in

Kiruba Balakrishnan
Project Manager
TECH Mahindra Limited India
Chennai, Tamilnadu
✉ kiruba.b@gmail.com

K.S. Raja
Chief Manager
Bank of Baroda
Mumbai, Maharashtra
✉ ksraja007@gmail.com

Kavita Bani
Assistant Professor
Department of Electronics and Computer Science
Atharva College of Engineering
Mumbai, Maharashtra
✉ Kavitabani@atharvacoe.ac.in

ABSTRACT

New security issues have arisen as billions of smart devices are connected via the Internet of Things (IoT). The proliferation of cyberattacks has rendered the privacy of network traffic an urgent issue over the past decade. The intrusion detection system (IDS) is the most promising method for identifying anomalous network behavior patterns and preventing hostile intrusions. Despite the tremendous efforts of the researchers, IDS continues to experience challenges in identifying new intrusions and enhancing detection accuracy while simultaneously reducing false alarm rates. Intrusion detection systems that are based on machine learning (ML) have recently been implemented as a viable method of identifying network breaches. We examine the efficacy of several machine learning techniques, such in the detection of network intrusions. This part is devoted to the evaluation of these techniques.

KEYWORDS *: Intrusion detection system (IDS), SVM, RF, KNN, DT, LR.*

INTRODUCTION

The world is shifting toward intelligence, networking, and digitalization. The internet's rapid development has sped up the flow of data assets. Specifically, the proliferation and interchange of data has been facilitated by the Internet of Things (IoT), which consists of a number of networked gadgets. People will find this transformation convenient, but there are also unspoken risks to digital data. As a result, safeguarding data assets from theft or infringement is a crucial problem that has drawn the attention of security researchers. against the Internet and the Internet of Things, there are innumerable attacks against commercial gaming and private interests. Both the assault technique and the quantity of attacks have significantly grown in recent years. This poses unforeseen hazards to the Internet's and the Internet of Things' steady and secure operation. Thus, there is an immediate need for a strong tool to recognize different network security risks and fend them off. IDS is a type of security technology that may thwart unauthorized external assaults and stop unauthorized connections. IDS can offer the Internet and IoT usability, secrecy, and integrity. IDS is unquestionably necessary to safeguard the Internet and the Internet of Things [1].

In response to the growing number of cyberattacks that are being launched against governments and commercial companies all over the world, fast development of intrusion detection systems (IDS) has taken place in both research and industrial institutions. There has been a consistent rise in the annual expense of combating cybercrime. Cybercrimes that are caused by malevolent insiders, denial of service attacks, and web-based assaults are the most catastrophic instances of cybercrime. In the event that these hostile assaults are launched against the system, industries or corporations may suffer the loss of their intellectual property. Organizations implement a firewall, antivirus software, and an intrusion detection system in order to combat the occurrence of such activities [2].

Intrusion detection is the secondary safeguard for a secure network. Intruder assaults are not a problem for an intrusion detection system (IDS) since it can learn from past attempts and strengthen its defenses accordingly. Based on the information gathered, intrusion detection systems may be grouped into three types: (1) Intrusion detection systems that run on individual hosts without access to external network data and instead rely on an internal library of suspicious data to identify potential threats. Not only do these systems not function well for real-time network data collecting, but they also squander a lot of CPU resources, making them unsuitable for tiny dispersed devices. (2) a network-based intrusion detection system may build an associated intrusion detection library and gather data packets in real-time from the network. (3) DISDS, which can detect both host and network data, are intrusion detection systems. Reliable and inexpensive, WSNs are distributed devices. To make sure the whole system is secure, they may be extensively used on the edge of the Internet of Things [3].

All network traffic, both incoming and outgoing, from Internet of Things devices is monitored by the Intrusion Detection System in order to identify any intrusions. Two methods exist for identifying malicious activity: signature-based and anomaly-based. Potential security risks that might jeopardize the communication connection between IoT devices include data integrity attacks, man-in-the-middle attacks, session hijacking, and Distributed Denial of Service (DDoS) assaults. The increase of Internet of Things (IoT) vulnerabilities is directly linked to the fast advancement of IoT hardware, software, and services. Hackers and thieves may readily exploit IoT devices due to their insufficient security measures. An instance is the use of botnets to execute distributed denial of service (DDoS) assaults against the Internet of Things. Cyberattacks have increased due to the extensive deployment of IoT devices on the internet. The primary purpose of the technology is to ensure that IoT devices operate with little processing burden [4].

Network Intrusion Detection Systems (NIDS) have advanced considerably with the use of machine learning (ML) methods, allowing enhanced detection of both recognized and unidentified threats. Conventional techniques like as signature-based detection are ineffective against zero-day assaults, but anomaly-based approaches often encounter elevated false positive rates. Machine learning methodologies tackle these difficulties by discerning trends from network traffic data and extrapolating to identify fresh assaults. The integration of machine learning methods with feature selection, data preprocessing, and hybrid methodologies has shown encouraging outcomes in enhancing the dependability of network intrusion detection systems.

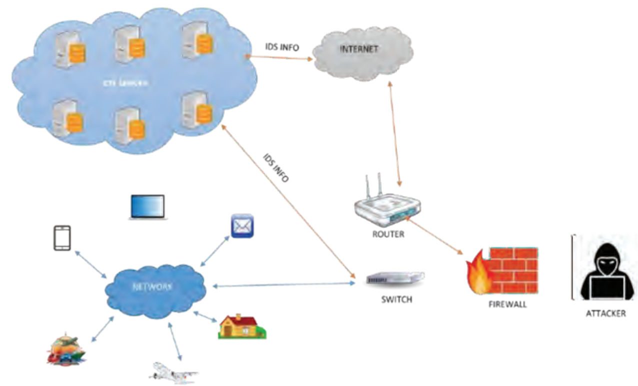


Fig. 1. Architecture of Intrusion Detection System

LITERATURE SURVEY

To prevent attacks on networks, Network Intrusion Detection Systems (NIDS) are essential. To keep up with complex and new threats, traditional methods like signature- based and anomaly-based detection often break down. One viable alternative is machine learning (ML), which improves the ability of Network Intrusion Detection Systems (NIDS) to detect intrusions by analyzing patterns in network data. This investigation takes a look at how NIDS has been improved using machine learning techniques and how well it can handle the growing number of cyber threats.

Decision Tree

Dasheng Chen et al. [5] created a better decision tree classifier to identify intrusions in networks. The data normalization procedure was enhanced via a series of tests, and an improved MaxAbs-DT classifier model was suggested by integrating the PCA feature selection approach. We train and validate our model using the widely- used NSL-KDD dataset, which is considered a gold standard in intrusion detection research, to see how well it performs.

Ranjit Panigrahi et al. [6] provide a host-based intrusion detection system that makes use of a C4.5-based detector. This system is based on the well-known Consolidated Tree Construction (CTC) methodology. This method consistently yields satisfactory results, regardless of whether or not the data are slanted toward a certain class.

For the purpose of generating a balanced sample from a dataset that was very skewed, an enhanced version of the random sampling method known as Supervised Relative Random Sampling (SRRS) was used during the pre-processing stage of the detector. For the purpose of enhancing the intrusion detection capabilities of the IDS dataset, we developed an enhanced multi-class feature selection technique and used it as a filtering component.

K-Nearest Neighbour

Using a machine learning strategy based on the K-Nearest Neighbors (KNN) method, Ghanshyam et al. [7] presented an intrusion detection system. The system effectively detects and categorizes network irregularities by preprocessing extensive data streams, choosing pertinent features, and enhancing classification processes. the model has shown a notable enhancement in its capacity to adjust to different network conditions and identify anomalies while maintaining a low rate of false alarms. The results highlight the significance of incorporating sophisticated machine learning methods into intrusion detection systems, setting the stage for enhanced security in industrial settings.

An approach to NIDS based on HOA is presented by Reza Ghanbarzadeh et al. [8]. The novel method takes use of herd- level horse behavior analysis to determine effective intrusion detection traits and the interplay between those traits. The new method begins with using the floor function to transform HOA into a discrete algorithm. The binarized algorithm is transformed into a quantum-inspired optimizer to improve the social dynamics of the herd's horses by combining HOA with quantum computing concepts. A novel technique called MQBHOA is used to identify intrusions in computer networks; it represents the multi- objective optimization challenge. The K-Nearest Neighbor (KNN) classifier is used to carry out the classification method.

Logistic Regression

An Ensemble Classifier based on Logistic Regression (LREC) is introduced by Silpa Chalichalamala et al [9] with the objective of improving the implementation of Intrusion Detection Systems. The LREC employs an iterative ensemble methodology to construct a robust classifier by integrating Random Forest (RF) and AdaBoost. The adaptive synthetic sampling (ADASYN) method is employed to resolve the issue of data imbalance. Furthermore, the process of recursive feature elimination (RFE) is employed to eliminate features that are not appropriate. The proposed RFE-LREC method is analyzed using two distinct datasets, BoT-IoT and TON-IoT.

Burak Kolukisa et al. [10] introduced a very effective approach using the LR model, which was trained with a parallel ABC algorithm. This method's training algorithm enables the learning of complex nonlinear and high-dimensional data by allowing local and global searches inside the solution space. Notwithstanding its remarkable performance results, the LR-ABC algorithm is impeded by protracted training durations. This is a significant issue, particularly for cybersecurity. Rapid training facilitates regular model updates and iterations, enabling the system to swiftly adapt to the constantly evolving landscape of intrusion techniques and network traffic. In response to evolving threats, it is essential to maintain current protections by routinely updating models. LRABC models are now available in both CPU and GPU variants, significantly reducing training time and mitigating the problem of extended computational duration.

Random Forest

Mhamad Bakro et al. [11] proposed a hybrid feature selection methodology that integrates the genetic algorithm (GA) with the grasshopper optimization algorithm (GOA). By integrating these two methodologies, we can expedite the identification of optimal solutions. An optimal set of attributes is used to train a random forest classifier. The approach employs a hybrid technique to address imbalanced data: it utilizes an adaptive synthetic (ADASYN) algorithm for over-sampling minority classes and random under- sampling (RUS) for the majority class as required. This integrative technique improves the system's overall performance by increasing the true positive rate (TPR) and decreasing the false positive rate (FPR) across all categories.

Muhammad Azhar et al. [12] provide an innovative approach to intrusion detection called "IDRandom-Forest." This technique utilizes a random forest model. Improving real-time detection capabilities while optimizing accuracy and reducing testing length is the goal of this technique. This method use stratified feature sampling to extract the optimal sub-ensemble from the conventional random forest model by integrating an accuracy sliding window with feature weighting.

A method was introduced by Majidian, et al [13] that enhances the effectiveness of intrusion detection systems by leveraging the benefits of software-defined networking, including its adaptability and centralized management. Two stages comprise the proposed methodology: At the outset, the network's topology is divided into numerous subdomains, with a controller node assigned to each subdomain. In the subsequent phase, a random forest-based ensemble classification model is implemented to detect intrusions within each subdomain. A genetic algorithm (GA) is used to fine-tune each element of this learning model, which is composed of a collection of classification and regression trees (CARTs). This classification model can be employed by controller nodes to identify intrusions independently or collaboratively. This work is distinguished by its collaborative pursuit of intrusion detection objectives through the enhancement of a variety of learning models. The efficacy of this model in detecting intrusions was assessed in an experimental setting using MATLAB software, with a particular emphasis on two databases: NSW-NB15 and NSLKDD.

Support Vector Machine

Chao Liu and associates [14] Created a model that integrates deep learning methodologies with machine learning approaches for intrusion detection. The model utilizes distributed computing on the Spark platform, using k-means and random forest algorithms for binary classification, efficiently distinguishing between normal and attack events. Consequently, several deep learning methodologies, are used to systematically categorize the identified anomalies into various attack classifications.

Mohammed Amin Almaiah et al. [15] studied an intrusion detection system model that uses a number of Support Vector Machine kernel classifiers in combination with the Principal Component Analysis feature selection approach. Then the performance is analysed on different kernel functions.

To enhance the detection capabilities and decrease the number of false positive alerts in host-based intrusion detection systems, Elsedimy et al. [16] created a new model for intrusion detection called QSVM-IGWO. Quantum Support Vector Machine (QSVM) performance is the target of this model, which incorporates IGWO algorithm-derived parameters. The idea of IGWO is to improve upon GWO by mimicking the social structure seen in grey wolves, which improves the search process. Binary classification also makes use of the QSVM concept, with the right kernel function chosen to provide the best possible outcome.

The bar chart in fig.2 illustrates the accuracy levels of five different machine-learning methodologies. Among the methods evaluated, K Nearest Neighbour achieved the highest accuracy at 99.8%, with Logistic Regression closely following at 99.44% and Random Forest at 98.98%. The SVM demonstrates the least accuracy, recording a value of 85.24%.

Table 1. Machine Learning based algorithms for intrusion detection system with NSL-KDD dataset

REF	Algorithm	IDS	Accuracy
[1]	Random Forest	HIDS	97.5
[8]	K Nearest Neighbour	SIDS	99.27
[14]	Support Vector Machine	HIDS	85.24
[17]	Decision Tree	SIDS	98.4
[18]	Logistic Regression	HIDS	99.44

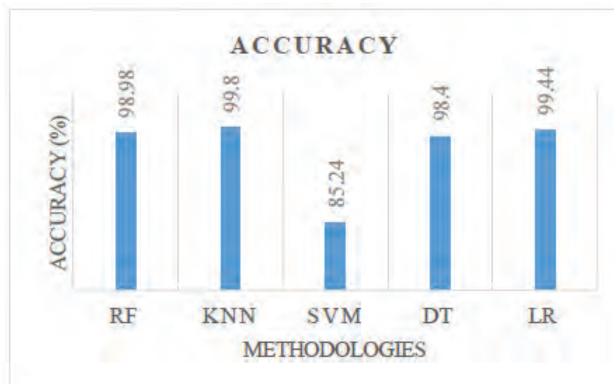


Fig. 2. Machine Learning based state-of-the-art techniques for intrusion detection system.

Table 2. Precision Recall and Fscore comparison of Machine Learning based algorithms for intrusion detection system

Algorithm	Precision	Recall	Fscore
RF[1]	97.85	94.67	97.48
KNN [8]	98.41	97.95	97.57
SVM [14]	82.45	84.12	83.84
DT [17]	97.64	97.89	98.21
LR [18]	99.61	99.32	99.12

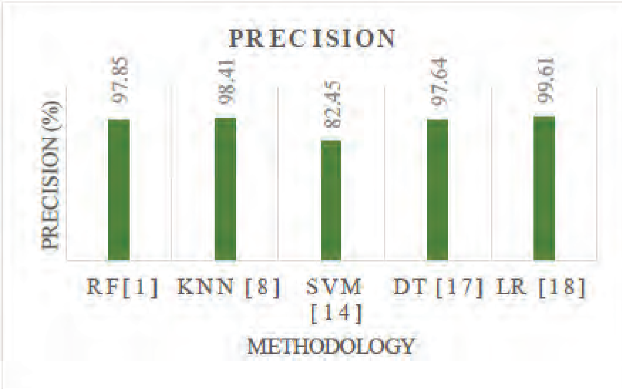


Fig. 3. Precision comparison of state-of-the-art techniques for intrusion detection system with NSL-KDD

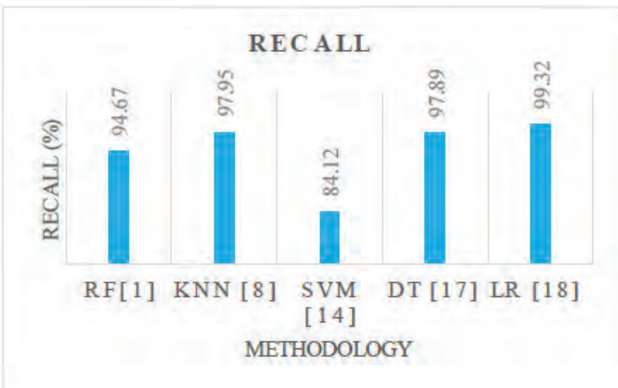


Fig. 4. Recall comparison of state-of-the-art techniques for intrusion detection system with NSL-KDD

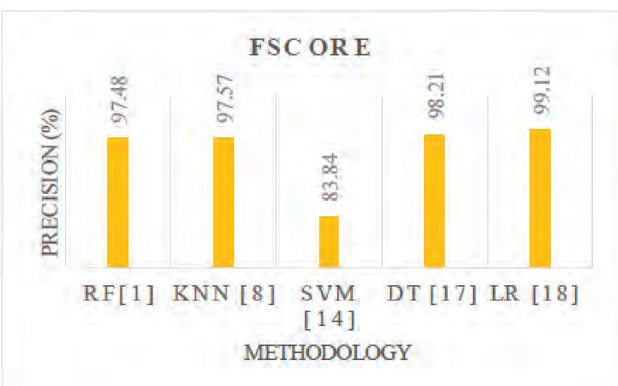


Fig. 5. Fscore comparison of state-of-the-art techniques for intrusion detection system with NSL-KDD

The algorithms for intrusion detection systems are compared in this table II and fig. 3-5. Overall, Logistic Regression (LR) has the highest Precision (99.61) and Recall (99.32). DT, which excels in F1-score (98.21), comes in second. Additionally, K-Nearest Neighbors (KNN) performs well and in balance. While Random Forest (RF) is competitive, Support Vector Machine (SVM) performs the worst.

CONCLUSION

Intrusion detection systems serve as a critical security element within today's technology-driven organizations. Nonetheless, delivering an effective and high-performance intrusion detection system to address a diverse range of security threats presents a significant challenge. The study showcases various advanced methods utilized within the realm of security for computer networks, specifically focusing on intrusion detection. These methods aim to enhance recognition rates in intrusion detection; however, it is widely acknowledged that the false positive rate remains a significant issue to be tackled in all research efforts. The methods evaluated for comparison include SVM, DT, KNN, RF, and LR using the NSL-KDD dataset, with KNN and LR demonstrating superior performance relative to the other methods.

Future research might focus on scalability to huge datasets, adaptation to changing assault patterns, and real-time detection efficiency. Lastly, resolving issues with unbalanced datasets and lowering false positive rates will improve the system's dependability even further.

REFERENCES

1. Balyan, A. K., Ahuja, S., Lilhore, U. K., Sharma, S. K., Manoharan, P., Algarni, A. D., ... & Raahemifar, K. (2022). A hybrid intrusion detection model using ega-pso and improved random forest method. Sensors, 22(16), 5986.
2. Mahbooba, B., Timilsina, M., Sahal, R., & Serrano, M. (2021). Explainable artificial intelligence (XAI) to enhance trust management in intrusion detection systems using decision tree model. Complexity, 2021(1), 6634811.
3. Liu, G., Zhao, H., Fan, F., Liu, G., Xu, Q., & Nazir, S. (2022). An enhanced intrusion detection model based on improved kNN in WSNs. Sensors, 22(4), 1407.
4. Chalichalamala, S., Govindan, N., & Kasarapu, R. (2023). Logistic Regression Ensemble Classifier for Intrusion Detection System in Internet of Things. Sensors, 23(23), 9583.
5. Chen, D., Song, Q., Zhang, Y., Li, L., & Yang, Z. (2023). Identification of network traffic intrusion using decision tree. Journal of Sensors, 2023(1), 5997304.

6. Panigrahi, R., Borah, S., Bhoi, A. K., Ijaz, M. F., Pramanik, M., Kumar, Y., & Jhaveri, R. H. (2021). A consolidated decision tree-based intrusion detection system for binary and multiclass imbalanced datasets. Mathematics, 9(7), 751.
7. Pandey, P. (2024). A KNN-Based Intrusion Detection System for Enhanced Anomaly Detection in Industrial IoT Networks. International Journal of Innovative Research in Technology and Science, 12(6), 1-7.
8. Ghanbarzadeh, R., Hosseinalipour, A., & Ghaffari, A. (2023). A novel network intrusion detection method based on metaheuristic optimisation algorithms. Journal of ambient intelligence and humanized computing, 14(6), 7575-7592.
9. Chalichalamala, S., Govindan, N., & Kasarapu, R. (2023). Logistic Regression Ensemble Classifier for Intrusion Detection System in Internet of Things. Sensors, 23(23), 9583.
10. Kolukisa, B., Dedeturk, B. K., Hacilar, H., & Gungor, V. C. (2024). An efficient network intrusion detection approach based on logistic regression model and parallel artificial bee colony algorithm. Computer Standards & Interfaces, 89, 103808.
11. Bakro, M., Kumar, R. R., Husain, M., Ashraf, Z., Ali, A., Yaqoob, S. I., ... & Parveen, N. (2024). Building a cloud-IDS by hybrid bio-inspired feature selection algorithms along with random forest model. IEEE Access, 12, 8846-8874.
12. Azhar, M., Perveen, S., Iqbal, A., & Lee, B. (2024). IDRandom- Forest: Advanced Random Forest for Real-time Intrusion Detection. IEEE Access.
13. Majidian, S. Z., TaghipourEivazi, S., Arasteh, B., & Ghaffari, A. (2024). Optimizing random forests to detect intrusion in the Internet of Things. Computers and Electrical Engineering, 120, 109860.
14. Liu, C., Gu, Z., & Wang, J. (2021). A hybrid intrusion detection system based on scalable K-means+ random forest and deep learning. Ieee Access, 9, 75729-75740.
15. Almaiah, M. A., Almomani, O., Alsaaidah, A., Al-Otaibi, S., Bani-Hani, N., Hwaitat, A. K. A., ... & Aldhyani, T. H. (2022). Performance investigation of principal component analysis for intrusion detection system using different support vector machine kernels. Electronics, 11(21), 3571.
16. Elsedimy, E. I., Elhadidy, H., & Abohashish, S. M. (2024). A novel intrusion detection system based on a hybrid quantum support vector machine and improved Grey Wolf optimizer. Cluster Computing, 1-19.
17. Singh, K. P., & Kesswani, N. (2022). An anomaly-based intrusion detection system for IoT networks using trust factor. SN Computer Science, 3(2), 168.
18. Kunhare, N., Tiwari, R., & Dhar, J. (2022). Intrusion detection system using hybrid classifiers with meta-heuristic algorithms for the optimization and feature selection by genetic algorithm. Computers and Electrical Engineering, 103, 10838.

Carbon Quantum Dots: Synthesis, Properties and Applications- A Future Perspective

Pinky Steffi Alexander
Assistant Professor
Humanities and Applied Sciences Department
Atharva College of Engineering
Mumbai, Maharashtra

Abhilasha Saini
Assistant Professor
Humanities and Applied Sciences Department
Atharva College of Engineering
Mumbai, Maharashtra

Pravin Nemade
Professor, Director
Humanities and Applied Sciences Department
Atharva College of Engineering
Mumbai, Maharashtra

Kiruthika B
Assistant Professor
Humanities and Applied Sciences Department
Atharva College of Engineering
Mumbai, Maharashtra

ABSTRACT

Traditionally fragile carbon nanoparticles (less than 10 nm in size) with a variety of special features, carbon quantum dots (CQDs, C-dots, or CDs) have been widely used in an increasing number of disciplines in recent years. With an emphasis on their synthetic techniques, size control, modification tactics, photoelectric capabilities, luminous mechanism, and applications in medical science, catalytic processes, and sensor concerns, Inherent many uses are made possible by these basic qualities, which enable accurate tweaking of their optical and electrical characteristics. Together with their water the ability to dissolve, biocompatibility, and different surface functional categories, CQDs' photoluminescence makes them essential elements in the domains of materials science, nanotechnology, and medicine. This article so emphasizes the wide range of applications in which CQDs have contributed significantly in the twenty-first century. we outline the latest developments in the area of CQDs in this section of the paper.

KEYWORDS *: Novel materials, Carbon, Quantum dots, Nanomaterials, Engineering applications.*

INTRODUCTION

A new type of zero-dimensional (0D) fluorescent carbon nanomaterials, Another name for carbon dots (CDs) is carbon quantum dots (CQDs), are often less than 10 nm (1-3). The production of these CQDs nanoparticles was accidentally invented in the year 2004 while purifying single-walled carbon nanotubes (SWCNTs). The optical qualities of CQDs were then improved in 2006 by laser ablation of a carbon target made of a cement and graphite powder combination. Sun and his group came up with the acronym "carbon quantum dots." Since then, CQDs have attracted a lot of interest because of their unique structure and intriguing qualities in a variety of domains. The unique characteristics of CQDs nanoparticles include exceptional photoluminescence and optical characteristics with a high quantum yield (QY), simple and economical production from biobased sources, adjustable excitation and emission characteristics, substantial thermal and optical stability, simple chemical functionalization, and negligible cytotoxicity with excellent biocompatibility. Thus, fluorescent CQDs are useful in many different areas, including energy storage and conversion, photocatalysis, electrocatalysis as, and bio-applications. The basic ideas that underpin CQDs and their top-down and bottom-up synthetic processes are summarized in this article. Additionally, it underlines the key characteristics of CQDs that set them apart from semiconductor QDs. Their uses in a variety of research projects, including drug/gene delivery, energy conversion and storage, electrochemical sensing, biosensing, bioimaging, and more recently, corrosion inhibitions, are being emphasized.

CQDs are a unique class of photoluminescence nanoparticles that are distinguished by their tiny size (less than 10 nm) and a structure that is dependent on the particular synthesis techniques used. They have a quasi-spherical form, a low molecular weight, and the quantum confinement effect. The surface of CQDs usually contains a range of functional moieties, including -OH, -NH2, -COOH, etc. Their exceptional water solubility and biocompatibility are due to these functional moieties. A network structure made up of active moieties, especially oxygen and amino-based groups, is visible on the surface of CQDs. CQDs are suitable for a variety of chemical alterations because to the large number of functional moieties on their surface. Furthermore, CQDs have exceptional polymerization properties that enable interactions with a variety of organic and inorganic components that are physiologically active. Surface passivation is an approach that may improve the physical and fluorescent characteristics of CQDs. Atom doping may be used to accurately control the optical characteristics of CQDs by changing their surface states. The luminescence characteristics of CQDs are mostly determined by their chemical structure. CQDs have been effectively characterized using a variety of analytical methods, including X-ray photoelectron spectroscopy (XPS), ultraviolet visible (UV-vis) spectroscopy, and Fourier transform infrared (FTIR) spectroscopy, among others. CQDs have a competitive edge over other carbon-based nanomaterials thanks to their unique characteristics.

Assessments by Lee et al., Zhu et al., and Baker et al. [3], [4] support this.[5] In the past few years, substantial progress has been made in the production, properties, and applications of carbon-based quantum dot particles. The excellent (aqueous) dissolution, robust chemical inaction, simplicity of alterations, and superior susceptibility to sunlight exposure make photoluminescent carbon-based quantum dots superior to both organic dyes and traditional semiconductor quantum dots. Carbon-based QDs are potential for application in biosensors, biomolecule/drug delivery, and biophysical imaging because of their exceptional biological properties, which include minimal toxicity and exceptional biological suitability. Because of their remarkable electronic properties as electron donors and acceptors, which produce the technique of chemical and electrochemical luminescence, carbon-based quantum dots have a wide variety of potential uses in optronics, catalysis, and sensing.[1],[2]

Since GQDs' characteristics and uses have already been thoroughly outlined elsewhere, we will concentrate on CQDs because of their special qualities and enormous potential for a wide range of uses. With an emphasis on their synthetic techniques, size control, modification tactics, optical characteristics, luminescent mechanism, and applications in biomedicine, optronics, catalysis, and sensing concerns, we outline the latest developments in the area of CQDs in the aforementioned articles.

SYNTHESIS APPROACHES

Over the last ten years, a variety of techniques have been put out to construct CQDs; they may be broadly categorized as "Top-down" and "Bottom-up" techniques, and each of them can be altered either before or after intervention (Fig. 1). Three issues with CQD planning should be noted: The use of electrochemical synthesis, restricted pyrolysis, or solution chemistry techniques can prevent (i) carbonaceous aggregation during carbonization; (ii) size control and uniformity are crucial for consistent characteristics and technological research and can be optimized through following treatment techniques like gel electrophoresis, centrifugation, and dialysis; and (iii) surface properties are crucial for solubility and specific applications and can be adjusted during preparation or post-treatment. The primary techniques for creating CQDs, controlling their size by restricted pyrolysis, and modifying them through functionalization, doping, and nanohybrids will all be covered. Table 1 provides an overview of the characteristics of the various artificial methods used to produce CQDs.

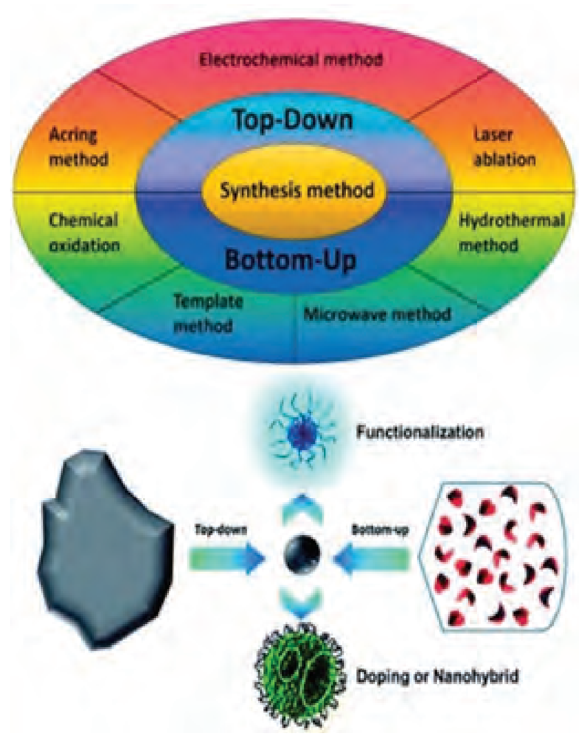


Fig.1. Synthesis techniques of Carbon QDs

Table 1: Various Synthesis approaches of Carbon QDs

Synthetic Approaches	Pros	Cons
Chemical ablation	Most easily available, several sources	extreme circumstances, extreme procedures, many stages, and inadequate size control
Electrochemical carbonization	Nanostructure and dimension are stable, controlled, and a single-step	A modest number of precursors for tiny molecules
Laser ablation	Quick, efficient, and adjustable states of surface	Insufficient QY, inadequate size control, and the necessity for modification
Microware irradiation	Quick, scalable, economical, and environmentally friendly	Inadequate management of sizes
Hydrothermal/ solvothermal treatment	economical, environmentally benign, and non-toxic	Inadequate management of sizes

PROPERTIES AND APPLICATIONS

CQDs are often divided into three categories: GQDs, CQDs, and CPDs. These are defined by their "core-shell" nanostructure, which is made up of surface groups with functions and a nanoscale carbon centre.

Fluorescent and electrochemical monitors represent both of the main types of sensors created recently using CQDs or CQD-based nanomaterials. CQDs' exceptional optical and electrical qualities may explain why they are so widely used in the research profession as fluorescence and electrochemical analysers.

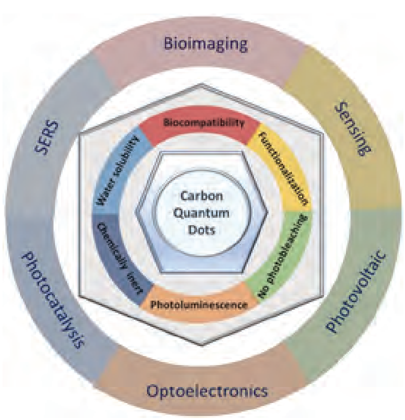


Fig.2. Benefits of CQDs in a variety of domains

For therapeutic reasons, it is crucial to quantify the biomolecules present in extracellular fluids like blood and urine samples. In order to determine the quantity of biomolecules present in these fluids at the lowest feasible cost, researchers have developed a number of instruments that can come into direct contact with the fluid.

It is important to recognize the role that carbon quantum dots (CQDs) play in preventing corrosion in addition to their achievements to technology development. The ability of various metal substrates to withstand corrosion in extreme environments has been greatly enhanced by CQDs and heteroatom-doped CQDs . It's interesting to note that doping CQDs with metal nanoparticles and heteroatoms like N and S has been identified as one of the elements influencing how well CQDs prevent corrosion.

CONCLUSION AND FUTURE PERSPECTIVE

CQDs seem to have limitless promise as we look to the future. They are attractive prospects for developments in industries including materials science, energy technology, and medicine because of their flexibility and versatility. The secret to opening up even more options lies in the continuous study into improving their qualities and broadening their uses. Furthermore, the environmentally benign method of synthesizing CQDs from renewable and sustainable sources is in line with the rising focus on sustainability and environmental awareness throughout the world. This feature guarantees that CQDs will continue to be a focus of study and innovation in the years to come, especially when combined with their affordability and scalability. Future medical research may concentrate on improving the ability to target certain illnesses, reducing toxicity, and investigating novel treatment and diagnostic approaches. New opportunities in material science and nanotechnology may arise from developments in the synthesis and functionalization of CQDs. Additionally, new uses are developing, such as in the creation of nanocomposites, sophisticated coatings, and smart detecting materials. Additionally, future studies in energy conversion and storage could focus on improving materials derived from CQDs for increased catalytic efficiency and storage capacity. To sum up, the fascinating development of CQDs from their basic ideas to their current uses is evidence of the astounding advancements in nanoscience and technology. We believe that as we go forward, CQDs will remain essential in determining the direction of technology, materials science, and many other scientific fields.

REFERENCES

1. A. P. Steffi, R. Balaji, , and C. Narendhar, “Rational Construction of SiO2/MoS2/TiO2 Composite Nanostructures for Anti-Biofouling and Anti-Corrosion Applications,” ChemistrySelect, vol. 6, no. 5, pp. 917–927, 2021, doi: 10.1002/slct.202004263.

2. A. P. Steffi et al., “Incorporation of SiO2 functionalized gC3N4 sheets with TiO2 nanoparticles to enhance the anticorrosion performance of metal specimens in aggressive Cl− environment,” Chemosphere, vol. 290, no. December 2021, p. 133332, 2022, doi: 10.1016/j.chemosphere.2021.133332.

3. S. N. Baker and G. A. Baker , ” Luminescent carbon nanodots: emergent nanolights”Angew. Chem., Int. Ed., 2010, 49 , 6726 .

4. H. Li , Z. Kang , Y. Liu and S.-T. Lee , “Carbon nanodots: synthesis, properties and applications“J. Mater. Chem., 2012, 22 , 24230.

[5]. J. Shen , Y. Zhu , X. Yang and C. Li , “Graphenequantum dots: emergent nanolights for bioimaging, sensors, catalysis and photovoltaic devices“Chem. Commun., 2012, 48 , 3686.

Enhancing Security and Privacy in Book Exchanges with Blockchain Technology

Renuka Nagpure
Information Technology
Atharva College of Engineering
Mumbai, Maharashtra
✉ renukanagpure@atharvacoe.ac.in

Divya Gajangi
Information Technology
Atharva College of Engineering
Mumbai, Maharashtra
✉ divyagajangi@atharvacoe.ac.in

Dhanshree Bhanushali
Information Technology
Atharva College of Engineering
Mumbai, Maharashtra
✉ dhanashree@atharvacoe.ac.in

ABSTRACT

Trust and transparency are paramount in the digital era. The process of exchanging books with people you meet on the internet always carries the risk of repudiation of origin. In a digital age where trust and transparency are at a premium, blockchain emerges as a game-changer. Blockchain provides a decentralized, tamper tamper-proof ledger which can help to foster trust by creating a secure and transparent ecosystem. With blockchain, every book exchange becomes a verifiable transaction, recorded permanently on the ledger. Users can confidently trade books, knowing that ownership history is beyond dispute. Smart contracts are self - executing agreements that automate and enforce book exchanges. By eliminating intermediaries and establishing a trust layer, smart contracts play a pivotal role in enhancing security, accountability, and efficiency within our book exchange platform. To facilitate the development and testing of these smart contracts, we've integrated Ganache, a local Ethereum blockchain emulator. Ganache streamlines the development process by providing a lightweight and customizable local blockchain environment, ensuring that our platform's smart contracts function seamlessly. The Truffle Suite complements our local blockchain network development by offering a suite of tools that simplifies smart contract development and deployment. Our project aims to provide users with secure, cost-effective, and transparent book exchanges by harnessing blockchain's transformative power and utilizing tools like Ganache and Truffle Suite.

KEYWORDS *: Blockchain, Trust, Transparency, Book exchange, Ethereum, Smart contract, Ganache.*

INTRODUCTION

In the traditional book exchanges, people who wanted to exchange books with one another often relied on manually searching people who would be willing to exchange books with them. In this system there may not be a robust system for tracking book ownership or the history of exchanges.

Book-Exchange is a web application, where users can purchase from the site and exchange their books with other users on the platform. Book enthusiasts can discover like- minded individuals in their vicinity who are open to book exchanges. The power of blockchain ensures that each exchange is transparently recorded, establishing an immutable history of book ownership. [1].

Existing book exchange platforms are susceptible to encounter cases of repudiation of origin. Books offered for exchange have no real track of ownership. Blockchain's tamper-proof record-keeping empowers users to confidently trade books while maintaining a reliable record added transactions.[2].

By conducting a blockchain transaction before the user adds their book for exchange, a tamper-proof immutable record ofthe book being put for exchange is created with the help which, once an exchange is displayed, all proof of it is maintained.[1]. Ethereum is a decentralized blockchain platform that works with smart contracts. [3]. These smart contracts are programs that automatically execute a transaction based on the criteria that specified conditions are met. This allows smart contracts to conduct transactions without the need of any central authority.[4]. By allowing participants to conduct transactions between each other when they want to initiate an exchange, we can create a tamper proof ledger on the ethereum blockchain. Ganache is a tool useful for creating ethereum based blockchain networks and smart contracts that the ethereum blockchain can deploy can be built using Remix IDE[5].

The users need a Metamask wallet to conduct the transaction.[8].

The application also allows the users to make use of a user friendly interface that can allow them to buy books and initiate transactions. The backend of the application is handled by PHP and MySQL database and Xampp server is used to host the site on the local machine.[9].

LITERATURE REVIEW

Kushwaha, Satpal Singh, et al [2] highlight the vulnerabilities faced by an Ethereum Smart Contract. As these Smart Contracts hold huge amount of cryptocurrency, it can lead to huge losses. Real life attacks and threats are explored which can be faced by systems employing these contracts. Afterward, preventive measures are given to avoid being exposed to the various vulnerabilities.

Hassan HS, Hassan R, et al [3] explore how an Ethereum Blockchain can be used for developing a Voting system. The immutability of blockchain allows it to host fair elections where the votes cannot be tampered. The paper explores how a local Ethereum network using Ganache can be configured and used with smart contracts deployed on Remix IDE.

Jain, SM [5] elaborates on the functionality of Remix IDE. Remix IDE provides an Integrated development environment to develop smart contracts and then manage the lifecycle of those contracts—like compilation, testing, deployment, and updates. Remix comes with a full suite of development and deployment tools integrated for developing and managing the lifecycle of smart contracts. And, as Remix is web-based no software installation is required. To use the smart contracts, we can make use of the contract address and contract ABI and make use of them in the web3,js driver file.

OBJECTIVES

1. The primary objective is to establish a cutting-edge blockchain-driven system tailored specifically for book exchanges, promoting transparency through a meticulously designed framework that guarantees users a seamless experience. This innovative solution will revolutionize how books are exchanged by providing an unequivocal, tamper-proof record of all transactions, setting a new standard in reliability and accountability within the exchange process.
2. Furthermore, a key aspect of this venture involves the development of an intuitive and user-friendly interface that empowers users to effortlessly list their books for exchange, execute transactions with ease, and engage with smart contracts in a straightforward manner. By streamlining the entire exchange process, this interface aims to enhance user convenience and satisfaction, ultimately making book exchanges more efficient and engaging for all participants.
3. To fortify the system's integrity and safeguard users' data, a robust and secure platform will be built leveraging the advanced security features inherent in blockchain technology. By harnessing these cutting-edge security mechanisms, the platform will ensure that sensitive information remains protected and shielded from unauthorized access, instilling confidence in users and reinforcing the platform's reliability.
4. Additionally, a vital functionality will be implemented to empower administrators with the ability to access and update the website's data promptly and efficiently. This administrative capability will serve as a pivotal tool in maintaining the platform's accuracy and relevance, enabling seamless management of content and information to meet the evolving needs of users and stakeholders effectively.
5. In line with enhancing user experience and facilitating secure transactions, the integration of the Metamask wallet into the website will enable users to leverage this trusted tool for blockchain transactions seamlessly. By providing users with this familiar and reliable means of conducting transactions, the platform seeks to streamline the exchange process, promoting user adoption and confidence in utilizing blockchain technology for their book exchanges efficiently.

METHODOLOGY

Proposed System

If a person wishes to engage in a book exchange the first step is to identify a community of like-minded individuals who share an interest in sharing and exchanging literature. However even within such communities concerns about security and the authenticity of exchanges can often arise.

To address these pressing issues the Blockchain-based Book Exchange emerges as an innovative solution offering a seamless and user-friendly interface that caters specifically to book enthusiasts. This platform allows users to buy books and facilitate exchanges directly with one another fostering a vibrant community of readers and collectors. Utilizing an Ethereum-based blockchain network not only enhances the overall experience but also ensures the integrity of the transactions. When users decide to engage in an exchange they can list their books by executing a transaction on the platform. This transaction process guarantees that the exchange listing is legitimate tamper-proof and immutable significantly alleviating concerns about fraud. By leveraging the power of blockchain technology the Book Exchange creates a secure environment where users can confidently explore and exchange their literary treasures fostering a culture of sharing and collaboration among book lovers. Ultimately this innovative approach revolutionizes the way people connect over books enriching the literary community as a whole.

Block Diagram for Blockchain functionality

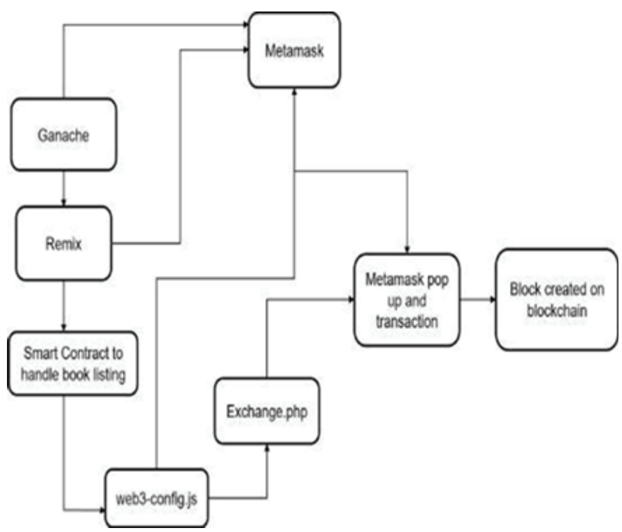


Fig. 1. Block Diagram for Blockchain functionality

1. First, the user sets up a workspace on Ganache, a personal blockchain for Ethereum development, and configures the RPC server to communicate with their local environment.
2. Next, the user links their Metamask wallet to the local host and imports their Ganache account into Metamask, enabling them to interact with the blockchain.
3. The user then creates a smart contract using Remix, an online web-based Integrated Development Environment (IDE), to handle the listing of books on the blockchain. This smart contract will contain the necessary functions for adding, updating, and retrieving book information.
4. After creating the smart contract, the user compiles it and deploys it to the blockchain network, making it ready for interaction.
5. The user copies the ABI (Application Binary Interface) and contract address from the deployed smart contract and utilizes it in their web3-config.js file. This allows their frontend application to communicate with the smart contract.
6. Within the web3-config.js file, the user creates an onClick function that triggers interactions with the smart contract when a form submission occurs on their Exchange.PHP webpage. This function handles the processing of book listings and transactions.
7. Upon submission of the form, a Metamask pop-up window appears, prompting the user to confirm the transaction. Once confirmed, the transaction is executed on the blockchain network through Metamask.
8. The transaction details are then stored on the Ganache blockchain, providing a secure and immutable record of the book listing activity. This ensures transparency and traceability of all interactions within the blockchain ecosystem.

User Interface

1. In the backend of the website, PHP is utilized to handle the server-side operations and manage the database interactions effectively.
2. On the front end, the user interface is developed using a combination of HTML, CSS, and Javascript to create a visually appealing and interactive experience for the users.
3. Upon visiting the website, individuals have the option to register either as an Admin or a regular user, based on their specific needs and requirements.
4. As an Admin, a dedicated dashboard is provided that displays all the vital activities and operations of the website, offering a comprehensive overview of the platform.
5. The Admin has the ability to perform key tasks such as adding new products, reviewing customer feedback

messages, verifying payments, and monitoring the active user base to ensure smooth operations.

6. Users logging into the platform are greeted with a user-friendly home screen that features intuitive navigation elements, enabling seamless access to various functionalities and resources available on the website.
7. The integration between the frontend and backend systems is facilitated through PHP, allowing for the efficient transfer of data generated on HTML pages to the MySQL database and enabling seamless data retrieval as needed.
8. The website is currently hosted locally using the Xampp server environment, ensuring a stable and secure platform for users to access and interact with the website's features and functionalities seamlessly.
9. The website is hosted on local host using Xampp server.

REQUIREMENTS

1. The objective is to design and implement a blockchain-based system that ensures transparency and accountability in book exchanges. By leveraging the immutable nature of blockchain technology, users can have a clear and tamper-proof record of books put up for exchange. This will instill trust among users and promote fair and secure transactions within the platform.
2. In order to enhance user experience, a user-friendly interface must be developed that simplifies the process of listing books for exchange, initiating transactions, and interacting with smart contracts. This interface should be intuitive and easy to navigate, making it convenient for users to engage with the platform without encountering any complications.
3. Security is paramount in any online platform, and leveraging blockchain's inherent security features is essential to build a secure environment for book exchanges. By utilizing encryption techniques, decentralization, and consensus mechanisms, the platform can effectively safeguard user data and transactions from unauthorized access and fraudulent activities.
4. A key requirement is to create a functionality that allows the Admin to access and update the data of the website. This will enable the Admin to manage and oversee the platform effectively, ensuring that the information displayed is accurate and up-to-date. By providing the Admin with the necessary tools and permissions, they can maintain the integrity of the platform and address any issues that may arise promptly.
5. Integration of the Metamask wallet into the website is crucial to enable users to conduct blockchain transactions seamlessly. By incorporating this feature, users can securely store their digital assets, manage their transactions, and interact with smart contracts directly through the website. This integration will streamline the exchange process and provide users with a convenient and secure way to engage with the platform.

RESULTS

To create a smart contract, go to the Remix IDE website. There we can create our smart contract defining the functionality to add the details of the book exchange on our blockchain with the help of a transaction. Once we have coded the smart contract we can then deploy the smart contract. When we deploy the contract, we get our contract address and the contract ABI. To integrate the smart contract into our website, we need to note down the contract address and contract ABI of our smart contract on Remix IDE. We can also test the smart contracts working on remix IDE by conducting transactions to verify. Then in ouir project directory we need to create a web3.js file which will connect the smart contract to our project using the contract address and contract ABI. We also create the instance to load up Metamask when we want to conduct a transaction.

The website uses PHP and MySQL for the backend and HTML, CSS, javascript are used to create the user interface.

Lastly we set up a workspace on ganache and define a connection to the local network. The RPC server address from the workspace can be used in the web3.js file to connect ganache to our local network.

Xampp server is used to host the project on the local machine. With the help of Xampp we also get access to phpMyadmin, we can access our database and all the tables of the database on it.

The exchange page allows users to enter the details of the book they want to put up for exchange. When the user presses the button to perform the action, the web3.js driver will create the instance that prompts the user to log in to their mettamask wallet account. Once logged in the user can perform the transaction to list the book for exchange. The details of the listing will be stored in the blockchain where they will be immutable and transparent.

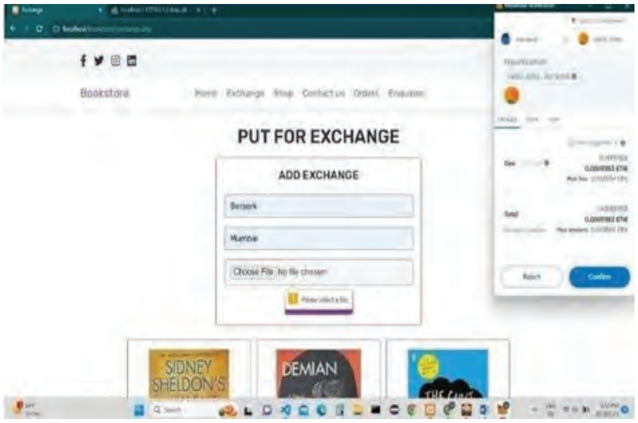


Fig. 2. Metamask Wallet Popup during transaction

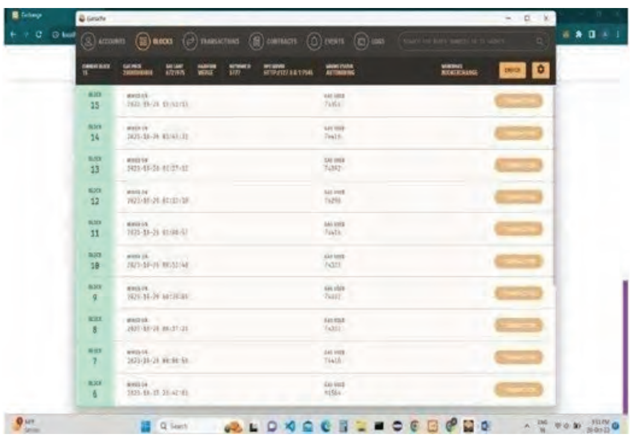


Fig. 3. Blockchain Network on Ganache

This is the ganache workspace which shows the blockchain network that includes all the blocks for the various transactions conducted on our website. Using an Ethereum blockchain network allows us to incorporate trust and transparency into our website.

CONCLUSION

Books are valuable sources of knowledge and at times when people are done with reading them, they might serve no use to the reader and so they might end up lying around collecting dust. To foster the creation of a community who is willing to aid one another and keep the cycle of gaining knowledge ongoing, exchanging books with one another is a valuable technique. Blockchain based book exchange aids the creation of this community. By providing an easy to use user interface, people can have more success in finding people who are willing to conduct book exchange with them. With the help of Ethereum based blockchain network and smart contracts, we managed to create a system that can keep an immutable ledger of the transaction which contains details of the book exchange, leading to better trust and transparency between the users.

REFERENCES

1. Ahmad D, Lutfiani N, Ahmad AD, Rahardja U, Aini Q. Blockchain technology immutability framework design in e- government. Jurnal Administrasi Publik (Public Administration Journal). 2021 Jun 16;11(1):32-41.
2. Kushwaha SS, Joshi S, Singh D, Kaur M, Lee HN. Systematic review of security vulnerabilities in ethereum blockchain smart contract. IEEE Access. 2022 Jan 4;10:6605-21.
3. Oliva, Gustavo A., Ahmed E. Hassan, and Zhen Ming Jiang. "An exploratory study of smart contracts in the Ethereum blockchain platform." Empirical Software Engineering 25 (2020): 1864-1904.
4. Hewa T, Ylianttila M, Liyanage M. Survey on blockchain based smart contracts: Applications, opportunities and challenges. Journal of network and computer applications. 2021 Mar 1;177:102857.
5. Hassan HS, Hassan R, Gbashi EK. E-Voting System based on Ethereum Blockchain Technology Using Ganache and Remix Environments. Engineering and Technology Journal. 2023 Mar 3;41(4):1-6.
6. Jain SM. Introduction to Remix IDE. In A Brief Introduction to Web3: Decentralized Web Fundamentals for App Development 2022 Dec 22 (pp. 89-126). Berkeley, CA:Apress.
7. Ahamed NN, Vignesh R. A Build and Deploy Ethereum .-Smart Contract for Food Supply Chain Management in Truffle-Ganache Framework. In 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS) 2023 Mar 17 (Vol. 1, pp. 36-40). IEEE.

[8] Lee WM. Using the MetaMask Crypto-Wallet. In Beginning Ethereum Smart Contracts Programming: With Examples in Python, Solidity, and JavaScript 2023 Apr 8 (pp. 111-144). Berkeley, CA: Apress.

[9] Welling, Luke, and Laura Thomson. PHP and MySQL Web development. Sams Publishing, 2003.

[10] Jain, Shashank Mohan. "Introduction to Remix IDE." A Brief Introduction to Web3: Decentralized Web Fundamentals for App Development. Berkeley, CA: Apress, 2022. 89-126

Timeless Strategies: Applying Chanakya's Arthashastra to Modern Governance and Leadership

Ritu Sharma
Dept. Humanities and Science
Atharva College of Engineering
Mumbai, Maharashtra
✉ drritusharma@atharvacoe.ac.in

Rajendra Mahajan
Dept. Humanities and Science
Atharva College of Engineering
Mumbai, Maharashtra
✉ rajendramahajan@atharvacoe.ac.in

Balaji Shinde
Dept. Humanities and Science
Atharva College of Engineering
Mumbai, Maharashtra
✉ balajishinde@atharvacoe.ac.in

Shobha D'mello
Dept. Humanities and Science
Atharva College of Engineering
Mumbai, Maharashtra
✉ shobhadmello@atharvacoe.ac.in

ABSTRACT

This paper explores the enduring significance of Chanakya's Arthashastra in relation to modern management principles, particularly in the areas of leadership, strategy, governance, ethics, and crisis management. Although originally crafted for rulers, Chanakya's teachings align closely with contemporary management frameworks, offering practical insights for today's business leaders. His focus on transformational and servant leadership highlights the importance of ethical, empathetic, and adaptive leadership. Likewise, his emphasis on strategic foresight and meticulous planning parallels modern methodologies such as SWOT analysis. Chanakya's principles on governance stress accountability, transparency, and delegation, reflecting the pillars of modern corporate governance. His pragmatic yet ethical stance on decision-making underscores the importance of values in business ethics and corporate social responsibility. Additionally, his foresight-driven crisis management strategies find parallels in modern risk management practices. The Tata Group is presented as a case study, demonstrating the practical application of Chanakya's principles in handling the 2008 financial crisis and challenges surrounding the Tata Nano project. The group's adaptability, ethical governance, and strategic foresight illustrate how Chanakya's teachings can guide organizations in navigating complex challenges and achieving long-term success. The findings highlight the enduring wisdom of the Arthashastra and its profound influence on modern management practices, offering timeless lessons for leadership, strategy, and governance in today's dynamic business environment.

***KEYWORDS** : Leadership, Strategy, Governance, Business ethics, Crisis management.*

INTRODUCTION

The *Arthashastra*, written by the ancient Indian scholar and strategist Chanakya (also known as Kautilya) in the 4th century BCE, remains one of the most influential texts on governance, economics, military strategy, and statecraft. Although primarily aimed at kings and rulers, its core principles have enduring relevance in today's rapidly evolving global business environment. Chanakya's teachings on leadership, strategy, governance, and ethics are not confined to the political sphere but provide invaluable insights for modern management practices. This paper explores how the wisdom contained in Chanakya's *Arthashastra* aligns with contemporary management practices, particularly in the areas of leadership, strategy, governance, and ethics.

LEADERSHIP IN CHANAKYA'S *ARTHASHASTRA* AND MODERN MANAGEMENT

Leadership in the *Arthashastra*

In the *Arthashastra*, Chanakya describes several qualities that an ideal leader should possess. These traits, though intended for rulers, are equally applicable to modern corporate leaders. Chanakya emphasizes that a leader must be wise, patient, and have a deep understanding of human psychology. He stresses the importance of balancing firmness with compassion, advising that a leader should be both strong and empathetic.

Chanakya writes, "A person should not be too rigid in his approach, nor should he be too lenient. He should be firm but not harsh" (*Arthashastra*, Book 1, Chapter 3). This philosophy reflects a leadership style akin to transformational leadership, which focuses on inspiring and motivating followers while maintaining discipline and setting clear expectations.

Modern Leadership Theories

Chanakya's leadership principles strongly resonate with modern frameworks such as transformational and servant leadership. Transformational leadership focuses on empowering individuals, inspiring them to reach their highest potential, and promoting growth on both personal and professional levels. In contrast, servant leadership prioritizes addressing the needs of others and ensuring the well-being of followers. In the contemporary business environment, the ability of leaders to embrace change, build trust, and uphold ethical practices is crucial. Chanakya's advocacy for adaptability and strategic foresight closely aligns with adaptive leadership, a modern approach that equips leaders to steer their organizations through uncertainty and ongoing transformation.

STRATEGIC THINKING AND PLANNING: PARALLELS BETWEEN THE *ARTHASHASTRA* AND MODERN BUSINESS STRATEGY

Strategic Thinking in the *Arthashastra*

The *Arthashastra* underscores the importance of strategic foresight, careful planning, and risk evaluation. Chanakya, a master strategist, teaches rulers to make decisions after evaluating resources, opportunities, and risks. His famous quote, "Before you start some work, always ask yourself three questions – Why am I doing it, what will the results be,and how can I do it?(*Arthashastra*, Book 1, Chapter 14), encourages thorough strategic thinking, a practice still relevant in today's decision-making processes.

Chanakya also highlights the importance of understanding the competitive landscape. He advises rulers to know their competitors, recognize market dynamics, and continuously adapt their strategies. For example, he states, "There is some self-interest behind every friendship. There is no friendship without self-interest. This is a bitter truth" (*Arthashastra*, Book 6, Chapter 2), which can be interpreted as a call for businesses to understand the motivations behind partnerships and competition.

Modern Strategic Frameworks

Modern strategic frameworks like SWOT analysis (Strengths, Weaknesses, Opportunities, Threats), Porter's Five Forces, and Blue Ocean Strategy closely align with the strategic principles outlined by Chanakya. For example, Porter's Five Forces, which assess factors such as competitive rivalry, the threat of new entrants, and the bargaining power of suppliers and customers, echo Chanakya's focus on analyzing competitors and maintaining a strategic advantage. Similarly, contemporary approaches like scenario planning and contingency strategies embody Chanakya's teachings on anticipating challenges and preparing for worst- case scenarios while striving for the best possible outcomes.

GOVERNANCE AND ORGANIZATIONAL STRUCTURE IN CHANAKYA'S *ARTHASHASTRA*

Governance in the *Arthashastra*

Chanakya outlines a well-structured system of governance in the *Arthashastra*, emphasizing transparency, accountability, and the importance of delegating authority to capable individuals. He asserts that a ruler must ensure that the state's operations are efficient and in service of the people. His emphasis on responsible governance is illustrated by his statement, "A person should not be afraid of the consequences of doing the right thing" (*Arthashastra*, Book 10, Chapter 3), underscoring the ethical obligations of leadership.

Modern Governance Practices

In the corporate world, governance concerns ethical management practices, transparency, accountability, and the efficient use of resources. Corporate governance ensures that companies are run responsibly, taking into account the interests of all stakeholders, including employees, shareholders, and the broader community. The role of a modern CEO or Managing Director is comparable to that of the ruler in Chanakya's model, where leadership involves delegating tasks to trusted individuals while maintaining overall responsibility for ensuring accountability and ethical conduct. Chanakya's model of governance, which relies on a clear organizational hierarchy and transparent decision-making processes, is reflected in modern organizational structures, whether hierarchical, matrix-based, or flat, ensuring clarity of roles and smooth operations.

ETHICS AND VALUES IN CHANAKYA'S *ARTHASHASTRA* AND MODERN CORPORATE ETHICS

Ethics in the *Arthashastra*

Chanakya places great emphasis on ethics, even while advocating for pragmatic statecraft. He advises rulers to act with integrity, fairness, and a sense of responsibility toward the people they govern. For example, he asserts that cruelty and injustice will lead to rebellion and undermine a leader's legitimacy. Chanakya's ethical leadership is also grounded in diplomacy and mutual benefit, as reflected in his statement, "There is some self-interest behind every friendship. There is no friendship without self- interest. This is a bitter truth" (*Arthashastra*, Book 6, Chapter 2), suggesting that all relationships should be based on fairness and mutual benefit.

Modern Corporate Ethics

Ethics and values play a vital role in fostering trust among stakeholders and achieving sustainable success in the modern corporate landscape. Business leaders are expected to uphold strong ethical standards, follow legal regulations, and make decisions that benefit society. Corporate Social Responsibility (CSR), which promotes accountability for a company's environmental and social impacts, reflects Chanakya's ethical philosophy. Contemporary practices like implementing codes of conduct, adhering to corporate governance principles, and ensuring financial transparency are firmly grounded in the ethical concepts articulated in the *Arthashastra*.

CRISIS MANAGEMENT: LESSONS FROM THE*ARTHASHASTRA*

Crisis Management in the *Arthashastra*

Chanakya's teachings on crisis management emphasize the importance of foresight, preparedness, and swift action in times of adversity. He advises rulers to anticipate internal and external threats and act decisively when crises arise. He writes, "A wise leader knows how to anticipate threats and prepare for them in advance" (*Arthashastra*, Book 7, Chapter 13), illustrating the need for proactive leadership in managing crises.

Modern Crisis Management

In today's business environment, crisis management has become a core component of risk management strategies. Businesses must prepare for financial downturns, reputation damage, natural disasters, and market disruptions. Modern organizations employ contingency planning, risk assessment, and disaster recovery strategies to ensure resilience during crises. Chanakya's emphasis on anticipating and preparing for challenges is reflected in contemporary business practices, where leaders take a proactive approach to risk management and crisis preparedness.

Case Study: Crisis Management at Tata Group

The Tata Group, one of India's leading and most accomplished conglomerates, has demonstrated exceptional resilience and strategic vision in navigating crises. The principles of Chanakya, an ancient Indian philosopher and strategist, are evident in the Group's approach to handling complex challenges. His teachings, such as adaptability, strategic foresight, resilience, and the importance of strong leadership, are mirrored in the Tata Group's responses to significant crises, particularly during the 2008 global financial crisis and the challenges surrounding the Tata Nano project.

The 2008 Financial Crisis: A Test of Resilience and Adaptability

The 2008 global financial crisis created significant turbulence in the global economy, with stock markets crashing, banks collapsing, and a widespread recession affecting businesses worldwide. The Tata Group, with its vast portfolio spanning multiple industries—automobile, steel, IT services, chemicals, telecommunications, and more—was inevitably impacted. However, Tata's response during this crisis is an excellent example of how Chanakya's principles of adaptability and foresight played a role in navigating a complex and evolving crisis.

Key Actions Taken

Strategic Divestment and Focus: In line with Chanakya's advice to anticipate changing circumstances and act accordingly, Tata Group streamlined and refocused its investments. For instance, Tata Steel, the largest steel producer in India, was facing immense pressure due to the global downturn in demand. However, instead of retreating, Tata Steel continued its global expansion, even successfully acquiring Corus Group in 2007, despite the financial turmoil. This strategic move was in line with Chanakya's principle of using crises to strengthen one's position.

Crisis Communication and Leadership: Ratan Tata, the chairman of the group, demonstrated Chanakya's principle of strong leadership during the financial crisis. His calm and transparent communication ensured that stakeholders, including employees, investors, and customers, were kept well-informed. He reassured them that the group would navigate the difficult times with resilience and a long-term focus, reinforcing the group's commitment to ethical leadership and strategic flexibility.

Diversification and Cost Control: The Tata Group's diverse portfolio helped cushion the financial blow during the crisis. It applied Chanakya's principle of building multiple streams of income (or resources) to manage risk. While sectors like automotive and steel were negatively impacted, other businesses, particularly Tata Consultancy Services (TCS), which had a significant global presence, performed well and provided much-needed stability to the group.

The Tata Nano Project: Strategic Foresight and Resilience in the Face of Failure

The Tata Nano, introduced in 2008, was designed to be the world's cheapest car. The project represented a bold and ambitious move to revolutionize the automotive industry in India and cater to the country's middle-class market. However, it faced a series of challenges that threatened its success. The Tata Nano project exemplifies how the principles of resilience, adaptability, and strategic foresight, as laid out by Chanakya, can be applied when confronting setbacks.

Key Challenges Faced

Production and Design Issues: Initially, the Nano's appeal as a cheap and accessible car was a strong selling point, but it soon became associated with poor quality. Many consumers perceived the car as "too cheap," which caused a decline in its perceived value and marketability. Additionally, the project faced production issues, and there were concerns about safety and customer expectations, which posed a significant threat to the brand image of Tata.

Political and Social Challenges: The initial production plan was set in Singur, West Bengal, but it encountered political opposition. The land acquisition process for the factory was highly contentious, leading to protests and public unrest. This setback forced Tata Motors to relocate the factory to Sanand, Gujarat, creating further delays and costs.

Economic Crisis Impact: By the time the Nano was launched, the world was experiencing the financial downturn of 2008, which further dampened the demand for an affordable vehicle, especially in a market that was seeing declining consumer confidence.

How Chanakya's Principles were Applied?

Resilience and Adaptability in Leadership: Ratan Tata demonstrated resilience, a key Chanakyan virtue, by refusing to abandon the project despite its setbacks. Although the Nano project faced multiple challenges, Tata believed in its long-term potential and remained committed to finding solutions. He showed flexibility by moving production to a new location and being open to redesigning and improving the car.

Strategic Foresight and Long-Term Vision: Despite the hurdles, Tata Group did not abandon the project. This mirrors Chanakya's teaching that long- term success requires patience and the ability to foresee future potential even in the face of short-term failures. The Tata Nano was never intended to be a short-term success but rather a means to push the boundaries of affordable engineering and meet India's aspirations for personal mobility. Tata Motors used feedback and learning from the Nano's initial failure to improve the vehicle's design and production process in later years. Learning from Setbacks and Pivoting: When the initial version of the Nano didn't meet expectations, Tata Motors pivoted by reengineering the model, including upgrading its features and revising its marketing strategy.This adaptation reflected Chanakya's principle of learning from failure and using those lessons to fortify future decisions.

CONCLUSION

Chanakya's *Arthashastra* offers timeless wisdom that remains highly relevant to modern management principles. His teachings on leadership, strategy, governance, ethics, and crisis management provide invaluable insights for today's business leaders. By aligning these ancient principles with contemporary business practices, organizations can not only thrive in a competitive and rapidly changing environment but also ensure they operate ethically and responsibly. The Tata Group serves as a case study for the successful application of Chanakya's principles in modern business. With its strong leadership, strategic foresight, ethical governance, and effective crisis management, the Tata Group exemplifies how ancient wisdom guide contemporary business practices. This enduring relevance of the *Arthashastra* underscores the timeless nature of Chanakya's insights, which continue to inform and inspire leaders worldwide.

REFERENCES

1. Chanakya. (2010). Arthashastra. Translated by K. S. K. Iyer. Penguin Classics.
2. Bajaj, R. (2013). Chanakya's Leadership Principles: Leadership Lessons from the Arthashastra. South Asian Journal of Management, 20(3), 67-84.
3. Tata Consultancy Services. (2020). Annual Report 2020-2021. Tata Consultancy Services Limited.
4. Tata Group. (2023). Tata Group Annual Report. Tata Sons Pvt. Ltd.
5. Sethi, S. P., & Arora, A. (2017). Corporate Governance and Ethics: An Indian Perspective. Sage Publications India.
6. Agnihotri, A., & Agarwal, N. (2012). Corporate Governance in India: The Way Forward. Journal of Indian Business Research, 4(2), 23-40.
7. Sharma, R. (2018). The Case of Tata Nano: A Leadership Crisis. Journal of Business Strategy, 39(4), 64-71.
8. McKinsey & Company. (2009). Leading through Crisis: Lessons from the Global Financial Crisis. McKinsey & Company Insights.

Review on Hybrid Electric Vehicle and Its Optimization using MPPT Algorithms

Garima Gurjar
Department of Electrical Engineering
Atharva College of Engineering
Mumbai, Maharashtra

Priynka Sharma
Department of Electrical Engineering
Atharva College of Engineering
Mumbai, Maharashtra

Priyanka Tripathi
Department of Electrical Engineering
Atharva College of Engineering
Mumbai, Maharashtra

ABSTRACT

With rising oil prices and growing concern for the environment, cleaner and more compatible energy solutions are needed. Current transportation provides a great deal of energy consumption and pollutant emissions. The combined effect of the driving forces of an IC engine and electric motor in a hybrid vehicle is examined and reviewed here. As per structural design with module of HEV such as the motor, engine, battery, bi-directional converter and maximum power point tracker (MPPT). This paper shows a comprehensive review of the main components used in HEV, such as their structural designs with pros and cons. Hybrid electric vehicles (HEV) are in trend for their capability to attain performance relative to gasoline vehicle. Significantly gives less emissions and enhanced efficiency in terms of fuel consumption.

KEYWORDS *: Hybrid electric vehicle, HEV architecture, MPPT algorithms.*

INTRODUCTION

Green technology innovations have placed an emphasis on green technology and are poised to change the demand for urban and suburban areas. The rapid expansion of urban and suburban areas has led to a significant increase in transportation usage, resulting in heightened pollution levels and various environmental challenges. It is essential to regulate vehicle emissions and implement proactive strategies to reduce their impact on the environment. Zero-emission vehicles are the vital goal and major target for researchers working in this field. [1]

For propulsion Hybrid electric vehicles use two or more energy converters. A hybrid electric vehicle (HEV) consists of an IC engine with one or more electric M-G set. Fossil Fuel energy gets converted into mechanical energy by engine; whereas mechanical energy is obtained by converting electrical energy from battery by electric motor [2]. To attain a better fuel economy traditional drive train is needed to be replaced by electric drive [3]. Here, it offers a series of HEV configuration, such as serial, parallel and combined series-parallel. Within each configuration, there are several variations. The advantage that the internal combustion engine (ICE) has improved mileage and low emission at high speed is combined with an electric motor that does not pollute and offers high power even at low speed in the parallel transmission unit of the HEV series, with a consequent increase in the efficiency of the transmission train and pollution reduction, tested in the latest Toyota Prius car.

REQUIREMENT OF HYBRID ELECTRIC VEHICLE

As the fuel is scarce and sooner or later it runs out. According to logistics for the year 2037, it is estimated that one billion oil dependent vehicles will be there on the world road maps, petrol will be tremendously costly. The world should be having solutions for the rest 400 million vehicles.

Therefore, to cope with this H.E.V. is the solution fully electric vehicles face limitations, such as short driving ranges and the inability to supply energy over extended periods compared to internal combustion engine (ICE) vehicles. Traditional ICE vehicles, however, contribute significantly to environmental pollution. Hybrid electric vehicles address these challenges by optimizing the use of both power sources. On highways, where ICE engines operate most efficiently and batteries in electric vehicles would deplete rapidly, the ICE is utilized. Conversely, for short trips and urban driving, the electric motor is either exclusively used or works in tandem with the ICE to ensure maximum efficiency.

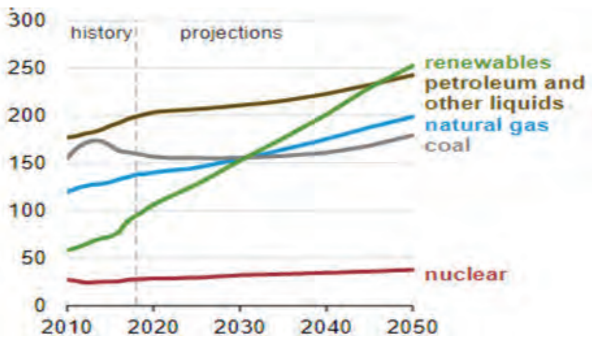


Fig. 1: Global primary Energy Consumption by Energy Sources (2010-2050)

There is potential for a market focused on small electric vehicles, such as golf carts or scooters. However, for larger vehicles like trains and buses, hybrid technology plays a more critical role. One significant advantage of hybrid cars over traditional gasoline-powered vehicles is their improved fuel efficiency and reduced environmental impact, making them a better choice for sustainability. Hybrid vehicles combine a gasoline engine with an electric motor, allowing for lower fuel consumption and increased energy efficiency. Figure 1 shows the world consumption of primary energy from energy sources (2010-2050).

TYPES OF MODEL

The main components in a Hybrid Electric Vehicle are electric motor, IC Engine, a tank of fuel, convertor, controlling board and battery. They are divided into three classifications:

1. Drive train is combination of electric drive &physically integrated ICE power source.
2. Battery/energy storage system —stresses cheap and easy energy storage and capacities.
3. Control system—guides the electric systems and ICE. Plus it handles the HESS.

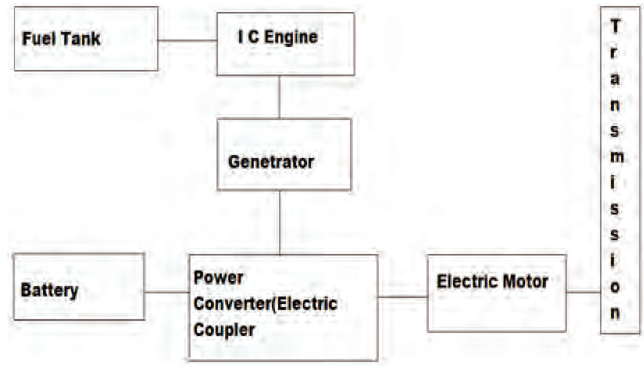


Fig. 2: Series hybrid vehicle system

These components can be integrated differently, resulting in a variation in the vehicle design. In component integration, the transmission trains mainly include serial, parallel and parallel serial combination projects. In [6], the architecture of the HEV is classified in various categories, which are paralleled hybrid vehicles, serial hybrids and combined and parallel series.

In the HEV series, energy sources provide electricity on the DC bus bar, then converted into traction energy [7]. The controls and electric traction motors are placed in place of all the mechanical transmissions b/w ICE and wheels having the advantage that the ICE not directly connected to the demand.

In parallel hybrid electric vehicles (HEVs), the traction power can come from either the electric motor, the internal combustion (IC) engine, or both working together. This type of hybrid system incorporates both an electric motor and an IC engine, which can operate independently to propel the vehicle or work in unison. A parallel HEV strikes a balance between vehicle performance and cost effectiveness. When the two power sources are aligned on the same axis, their speeds must be synchronized, and their torques are combined. This configuration is commonly seen in most electric bicycles. If only one power source is active, the other must either rotate simultaneously, be disengaged using a one-way clutch, or operate with a freewheel mechanism.

Complexed HEVs integrate functions of both parallel and also series architecture. The main difference is in the variance in power flow of motor of series–parallel hybrid, which is bi-directional in complicated hybrid and uni-directional for series–parallel combination of electric hybrid vehicle. Complexity in design is the main disadvantage of complex hybrid vehicles.

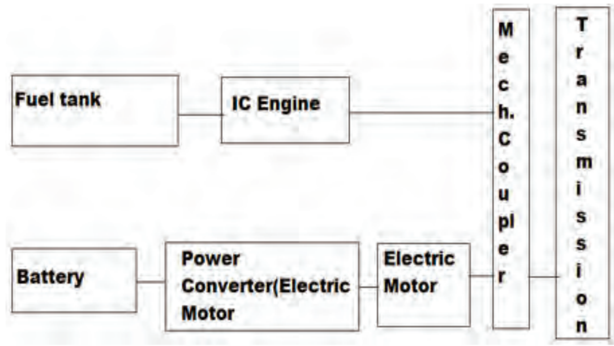


Fig. 3: Paralleled HEV

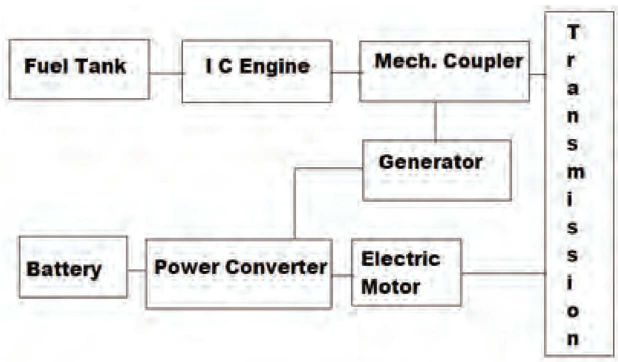


Fig. 4: Series Parallel Hybrid electric vehicle

Architecturally, PHEVs (plug-in hybrid vehicles) are similar to HEVs, with the exception of a large on-board battery, which has elevating energy density and output. Combination of CS (load maintenance) and CD (load reduction) modes requires a complex strategy of controlling as compared to a HEV. The P-HEVs start operating in CD mode (load exhaustion); and when battery reaches a SOC thres-hold value (state of charge), the battery changes to the charge maintenance mode when vehicle is in the parking and simultaneously recharged. The architecture of (PVHEV) a HEV powered by solar energy is very much like PHEV (plug-in hybrid vehicles), with the addition of an extra photovoltaic panel (PV), and it charges the battery during a day which is cloudless. To extract the maximum power of the photovoltaic panels, the algorithms for monitoring MPPTis applied. The figure shows the block diagram of PV-HEV.

For acceptable initial acceleration, traditional vehicles has increased engine size. Usually the larger engines has more power as compared to what is needed for traveling. At standstill full torque is produced by Motors (Electric) at stand-still and is well-adapted to support IC Engine torque decrementespecially at low RPMs. Power-split hybrid vehicles utilize a smaller, more efficient engine with reduced flexibility.

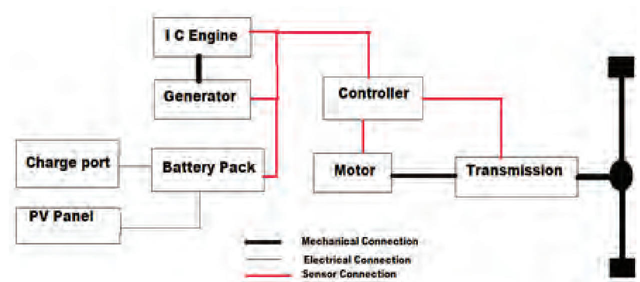


Fig. 5: Block Diagram of PVHEV

There is lot of Future Scope in improving the performance of power electronics-based drives by using Artificial intelligence Neural Network (ANN), hybrid networks, Fuzzy Logic etc. in the industrial sector. Presently combination of adaptiveness with this smart control looks like the maximum advantageous research area in control and the implemented electrical drives.

Table 1: Architectures and their Application

Architecture	Complexity	Efficiency	Hybridization	Computation time/ Mathematical Complex
Series	1	1	Complete HEV& plug-in HEV	1
Parallel	2	2	Micro- mild and complete HEV	2
Series-Parallel	3	3	Complete HEV & plug-in HEV	3
1—low, 2—medium, 3—high				

INTEGRATION OF MPPT ALGORITHMS IN HEV'S

Hybrid vehicles have the advantage of lower fuel consumption, lower operating cost reduced noise pollution and emissions due to these numerous benefits they are preferred choice over the other alternatives. As there is only single point of maximum power and also the I-V characteristics of the photovoltaic cells are not linear [18]. Efficiency can be increased through the energy interface, electronic devices with the photovoltaic system, together with the MPPC. Different algorithms are presented to monitor this MPP, to overcome the limitation of the temperature variation for the constant voltage tracking method (in conventional method), disturb and examine (in P&O) and the incremental conductance (in IC) methods are the mainly utilised. The load impedance varies periodically in the case of P&O and detects the direction change of the power, but has increased conductance, dP / dV is examinedto check if MPP is achieved. Even in P&O, the system fluctuates close to MPP, hence system works ineffectively [19]. For IC algorithm process, the voltage -current curves gives the max power level [20]. The main advantage of IC is its sensitivity to changing weather conditions, while P&O benefits from a simpler implementation [21].

In [22] an algorithm for loading the P&O based MPPT for PV-HEV is proposed. Normally, a pulse converter extracts maximum powerSimilarly, the charging algorithm adjusts the converter to operate in either constant current or constant voltage mode as needed. The P&O algorithm was also suggested for PV-HEVs in [23].

MPPT controlled algorithm which do not require aI detection device mechanism explained for HEV based applications [24]. MPP is estimated by the photo-voltaic voltage and the converter working ratio of this algorithm. In [25], a relatively small photovoltaic cell and some extra h/devices were finalized to decrement the energy loss and costs and to achieve high efficiency. In [26], the ANN-based MPPT in PVHEV is used and trained using the gradient propagation and descent recoil pulse algorithm. It is proposed to take in consideration P&O algorithm when vehicle is parked and the voltage dependent MPPT algorithm is suitable while driving [27].

To manage variable solar irradiation, P&Oand IC-based techniques are used. Both of the techniques provide a solution which is optimal. Battery life can be extended to charge efficiently, primary inductor converter and approximated P&O have been used. MPPT algorithms which are dependent on the optimization of ant colonies, the widespread logic, the genetic algorithm, the artificial neural network and the optimization of the particle bath, is applied to PV systems in varying irradiation conditions [28].

Photovoltaic panels and battery packs are generally connected in cascade to power the load. PHEV loads are also taken into action while designing a residential PV system connected to the grid. Voltage regulation is by the DC / DC converter of the PV sets using Maximum Power Point Tracking with IC method [29]. Both EV and PV are expected to give active as well as reactive power to power loads when the micro network is isolated from the public network [30].Table 2gives comparison indifferent MPPT algorithms used in HEV. Advantages of ANN and FLare more over others and thus they are in trend nowadays as a result of their advantages.

Table 2: MPPT Algorithms Comparison for HEVS

MPPT technique	Convergence speed	Implementation complexity	Periodic tuning	Sensed parameters
P&O	Modifies	L	N	V
IC	Modifies	M	N	V and C
Fractional V_{oc}	M	L	Y	V
Fractional I_{sc}	M	M	Y	C
F.L.C	F	H	Y	Modifies
Neural network	F	H	Y	Modifies

BENEFITS OF HYBRID ELECTRIC VEHICLE

A. Eco-Friendly: Even though hybrids produce some emissions, it is comparatively less than the gasoline-powered vehicles which still produce up to 50% more pollutants than hybrids and EVs. Research indicates that hybrid vehicles produce fewer emissions compared to traditional gasoline-powered vehicles. Since they operate using a dual-powered system that combines a gasoline engine and an electric motor, they reduce fuel consumption, thereby promoting energy conservation. Additionally, the electricity used to charge these vehicles, even when generated by power plants, results in lower overall pollutant emissions.

B. Advantages for saving money from a hybrid car: Financing a hybrid vehicle is often more cost-effective than financing a conventional car, as hybrids come with various incentives and tax credits.

Additionally, the energy regeneration feature during braking extends the lifespan of hybrid car brake pads, making them three times more durable than those in traditional cars and reducing the cost of frequent replacements. In automatic hybrid cars, a battery paired with an electric-petrol engine can reduce fuel consumption significantly, contributing to improved efficiency over the vehicle's lifespan.

C. Long term Benefits: Driving and comparing conventional with a hybrid one for equal distance, the former vehicle will require much of the energy to run and produce huge amount of greenhouse gases in its complete lifetime whereas latter one is cleaner and requires lesser quantity of fuel to run resulting in low emissions and not much dependence on fossil fuels.

D. Regenerative Braking System: For Hybrid vehicle, each time you apply break it regenerates energy by capturing the vehicle's momentum (kinetic energy) and turning it into charge that regenerates the onboard battery in case the vehicle is slowing down or stopping. This results in eradication of the excess time and requirement for halt to re-charge battery.

E. Higher Resale Value:Increasing quantity of people are inclining towards cars which are hybrid, with increase in price of gasoline. Thus hybrids are holding their value better than non-hybrids; they could end up being the better investment.

CONCLUSION

Hybrid electric vehicles are evolving as a promising substitute for the current state of transportation due to their low oil consumption, battery technology, control techniques and reduction of toxic emissions. Due to continued support and motivation from the government and automakers, hybrid vehicles are expected to be more popular. Movement energy is used to charge the battery with I.C.E. A controller assists in continuous supply when you accelerate / climb or need full power. The variations in the different components and their architecture magnify the efficiency of complete system. The suitability of the MPPT methods which are being utilized for solar power based HEVs are analyzed additionally in this document. Components with not just parallel and series but also combination architecture of HEV, ESS, motors, bidirectional converter and MPPT have been revised and the results are summarized at the end in the form of a table.

REFERENCES

1. C.C. Chan, State of Art of Electric and Hybrid Vehicles, Proceedings of the IEEE, vol. 90, No. 2, February 2002.
2. V.A. Shah, Jivanadhar A. Joshi and Ranjit Roy, Review of ultra- capacitor Tech. and its applications, 15th National
3. Emadi, K. Rajashekara, S. Willaimson, and S. Lukic, Topological Overview of hybrid Electric and Fuel Cell Vehicular Power system Architectures and Configurations, vehicular technology, IEEE Transactions on vol.54, 2005,pp. 763-770.
4. M. Ehsani, Y. Gao, S. gay and A. Emadi, Modern Electric, Hybrid Electric and Fuel Cell Vehicles- Fundamentals, Theory and Design, FL CRC Press, 2004.
5. Wang, H.Zheng, Route and spectrum selection in dynamic spectrum networks, in Proceeding IEEE CCNC 2006, pp.625-629, February 2006.
6. R. Chen et al., toward secure distributed spectrum sensing in cognitive radio networks, IEEE Communication Magazine, vol. 46,pp. 50-55, April. 2008
7. H. Khalife, N. Malouch, S. Fdida, Multihop cognitive radio n/ws to route or not to route, IEEE Network, vol.23, no.4,pp.20-25,2009
8. Y.C. Liang et al., Sensing-Throughput Trade-off for Cognitive adio Networks, IEEE Trans. Wireless Comm., vol.7,pp.132637, April 2008.
9. P. K. Visscher, how Self-Organization Evolves, nature, vol. 421, pp.799-800 Feb 2003.
10. T.A. Burress, C.L. Coomer, S.l. Campbell, A.A. Wereszczak, J.P. Cunningham, and L.D. Marlino, "Evaluation of the 2008 Lexus LS 600h Hybrid Synergy Drive System,"ORNL/TM-2008/185, UT-Battelle, Oak Ridge National Lab, Oak Ridge, Tennessee, Jan 2009
11. N. Kim, S. Cha, H. Peng Optimal control of hybrid Electric Vehicles based on Pontryagin's minimum principle. IEEE Transactions on control system technology ,19(2011),pp. 1279-1287
12. L. Jinming and P. Huei, "Control Optimization for a power-split hybrid vehicle,"in 2006 American Control Conference, 2006, pp.6
13. D. Lanzarotto, M.Passalacqua, M. RepettoEnergy ,Comparasion between diffenent Parallel Hybrid Vehicles architectures International Journal of Energy Production and Management, 2017, pp.370-380
14. A. Emadi, K. rajashekara, S,s, Williamson, S.M. Lukic Topological overview of hybrid electric fuel cell vehicular

power sustemarchtectures and configurations. IEEE Transactions on vehicular Technology, 54 (2005), pp.763-770

15. Hasan k., Fuzzy Logic Controller for Parallel Plug-in Hybrid Master thesis, University of Wisconsin Milwaukee, 2015

16. Kentli A, Studies on Fuzzy logic control of Electrical Machines in Turkish University: Overview published in Mathematical and computational Application, vol 16,pp 236-247,2011

17. Khoucha F, benrabahA, herizi O, Improved MPPT inter leaved boost converter for solar electric vehicle application, 4th Int. Con. On Power Engg, Energy and Electric drives, pp 1076-1081.

18. Shuang DU- Algorithm reserch on the MPPT of Solar EV Proceedings 2013 int. conf on mechatronic sciences, Electric Engi. and Computer , pp 74-78

19. Hadagali et al, Bi-directional dc-dc converter system for solar and fuel cell. Int. Conf. on emerging research areas: Magnetics, machines and drives, pp1-6, 2014

20. Armstrong P. , Kang et al. Reconfigurable PV array scheme integrated into an EV, Int. IET Hybrid and EV conference 2013, London, 2013, pp1-7

21. Sakib K N, Member S Kabir M Z, Williamson SS,Caldmium telluride solar cell: from devices modeling to EV battery mgmt., .IEEE transportation electrification conf. and expo (IETC), pp1-8.

22. Jeddi N, El Amraoui L, Rico FT A comparative study and analysis of different models for PV arrary in solar car, 2017 12th intconf on ecological vehicle & renewable energies, IEEE, pp 1-10.

23. Wolfs P, Quan Li, Current sensor free incremental conductance single cell MPPT for high performance vehicle solar arrays. 37th IEEE power electronics specialist conference, pp 1-7

24. Ahadi A, Liang X, A stand-alone hybrid RES assessment using cost optimization method, IEEE int conf. on industrial tech, IEEE pp 376-381

25. Zhang X, Chau KT, Yu C, Chan CC optimal solar thermoelectric hybrid energy system for hybrid electric vehicles-IEEE veh power propulous conference ,VPPC, 2008

26. Kalla U , Gurjar D, Rathore K S, Dixit P, An Efficient controller for PV operated Permanent Magnet BLDC drive based electric vehicle sys-IEEE 7th power India International Conference, pp 1-6 (2016)

27. El-Saddy G, Sharaf AM, Makky AM et al., An error driven hybrid neuro Fuzzy torque or speed controller for electrical vehicle induction motor drive, Proceedings of the intelligent vehicles'94 Symposium, IEEE,pp 449-454

28. Schuss C., Eichberger B.,Rahkonen, Monitoring sys for the use of solar energy in electric & hybrid vehicle, IEEE international instruments technical conf. proceeding, pp 524-527.

29. Nakir I, Durusu A, Ugur E, tanrioven M, Performance assessment of MPPT algorithms for vehicle integrated solar system. IEEE Int Energy ExhibitionENERGYCON,pp 1034-1038 (2012)

30. Adam Hirsch, Yael ParagJosep, Guerrero Joseph: A review of technologies, key drivers, and outstanding issues, GuerrerobaSchool of Sustainability, Interdisciplinary Center (IDC) Herzliya, IsraelbDepartment of Energy Technology, 2018.

Performance of Multilevel Inverter with Boost Converter

Pragya Jain
Atharva College of Engineering
Mumbai, Maharashtra

Ashweni Kumar Jain
Adani Electricity Mumbai Limited
Mumbai, Maharashtra

ABSTRACT

The most economic and appropriate inverters for high and medium power application are multilevel inverters (MLI). The overall harmonic distortion is decreased by these inverters. The primary problem with these multilevel inverters is the overuse of switches, which adds in-system losses, circuit complexity, size, cost, and other factors. To overcome this difficulty, this paper proposes a three-phase, five-level cascaded H-bridge inverter with a boost converter.

The new cascaded multilayer inverter with six switches is allowed in the suggested technique to reduce loss, save money, and enhance voltage stability. The modalities of operation of this unique topology are also defined to verify its detailed operation. The input side of this design has a boost converter, which reduces overall harmonic distortion while increasing the basic output voltage. Compared to current topologies, the suggested topology requires fewer switching devices. Finally, the design processes and MATLAB/Simulink simulation results of the suggested strategy are presented in this work.

KEYWORDS : *Multilevel inverter, Boost converter, Asymmetric voltage sources, Cascaded H-bridge, Decreased switches.*

INTRODUCTION

For the specific load, an electrical circuit of a power converter transforms electrical energy into a different structure. A converter performs additional processes and provides an output voltage that differs from the input [1]. In the computer power supply unit, a converter is used to increase or decrease the input voltage's magnitude and transmit various output voltages with the same or different polarity [2–3]. Direct current (DC) is converted to alternating current (AC) using an electronic device called a power inverter. The inverter regulates the frequency, total power, input, and output voltage based on the precise circuitry or device design [4]. The DC source will produce power, but the inverter will not. A power inverter may be entirely electronic or may combine electronic circuitry with mechanical effects. Static inverters do not use moving parts throughout the conversion procedure [5–6].

The inverter generates a few voltage levels for the capacitor banks connected in series using the disparate phases method. This inverter's purpose is to use a variety of diodes. By transferring a limited amount of voltage, a diode reduces the strain on other electrical equipment. The maximum output voltage of the diode clamped multilevel inverter is half of the input DC voltage, which is a barrier [7-8]. The previously described problem has been resolved by making improvements to the capacitors, switches, and diodes. These are limited due to the three-level capacitor balancing problems. The basic frequency is used for all switching devices. This type of inverter produces excellent efficiency as the outcome.

A fundamental concept of back-to-back power transfer systems is this [9]. MLI has numerous power switching devices and capacitor voltage sources. The capacity of MLI to provide output voltage waveforms with an enhanced harmonic spectrum for high voltage applications and to generate high voltages with a constrained most extreme device rating makes them reliable.

There are three types of MLIs: cascaded, diode clamped, and capacitor clamped H-bridge inverters [10–11].

The advantages of multilevel converters (MLC) include controllable active and reactive power and reduced output harmonics. It uses switches that are completely controlled. In the past several years, modular MLC has shown an invention and paved the way for lucrative high-power applications. Modularity is commonly used to describe a process that combines smaller subsystems to advance relatively large systems [12].

A cascaded connection of converter cells appears to be an exciting way to get high voltage and high-quality waveforms for power converter topologies with the least amount of complexity. Input current, switching frequency, common mode voltage, and harmonic distortion reduction are a few advantages of MLIs. Like this, MLCs have the advantages of only requiring one isolated DC link and having equal capacitor voltage in the main DC link. No current or voltage sensor is required, and the FC's self-balancing feature and modular design eliminate the requirement for feedback control [13–14]. High voltage AC and DC transmission lines, variable speed motor drives, static VAR compensation, high voltage system interconnections, sinusoidal current rectifiers, converters with harmonic distortion capability, and both AC-DC and DC-AC conversion applications are a few examples of uses. [15]

MLC has been under investigation and growth for over three years and has found effective industrial application. However, this is still a technology below improvement, and several new supports and advanced commercial topologies have been stated in the last few years. By using the multilevel of the converter inverter topology, the purpose of above issue had verified. With the reduced number of switches, the inverter design has been adopted by the five levels. To understand the operation of the cascaded converter along with the inverter, the state of switches has been defined.

PROPOSED TECHNOLOGY

One of the most significant topologies in the multilevel and multi-pulse inverter family is the conventional cascaded multilevel inverter. A desired AC voltage can be created by combining multiple DC voltage levels thanks to the cascade topology. Since they are all fuel cells, solar systems, batteries, etc., the DC levels have been seen as being the same. Isolation transformers are installed in H- bridge inverters, and H Bridge cascaded MLIS is used to isolate DC input sources. However, they do not require flying capacitors or clamping a diode. The primary benefit of cascaded MLI is the lack of voltage imbalance. CMLI uses fewer parts than diode-clamped and flying capacitors. The innovative switching topology, which reduces the number of switches, appears to improve inverter performance and voltage stability. The unique cascaded multilayer inverter with six switches is enabled in our suggested method to reduce loss, save money, and enhance voltage stability.

Five level cascaded H-bridge inverter

For the inverter model to produce a sinusoidal voltage output, H-bridge inverters must be connected in series. The total of the voltages produced by each cell is the output voltage. It is possible to limit the overall harmonic distortion by selecting the switching angles. The switching devices that determine the number of output voltage levels are necessary for the level cascaded H-bridge multilevel inverter. The figure depicts a five-level, single-phase cascaded inverter.

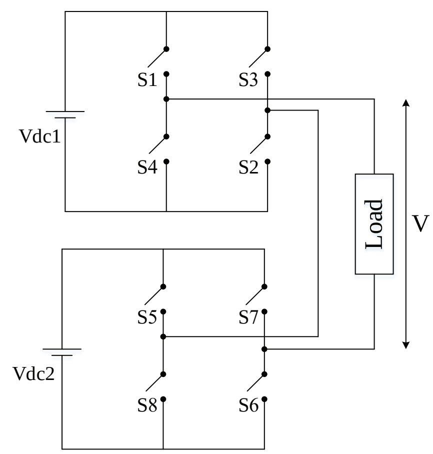


Fig. 1: Single phase five level CHB MLI

Each full-bridge inverter level's ac output is coupled in series with the others so that the total of the inverter outputs creates the

synthesized voltage waveform. The single phase, five level inverter's switching action is shown in Table 1.

Table 1: Switching values for single phase five level CHB MLI

S1	S2	S3	S4	S5	S6	S7	S8	Switching voltage level
1	1	0	0	1	1	0	0	+2Vdc
1	1	0	0	0	0	0	0	+Vdc
0	1	0	1	0	1	0	1	0
0	0	1	1	0	0	0	0	-Vdc
0	0	1	1	0	0	1	1	-2Vdc

Five-level cascaded three-phase H-bridge inverter

The proposed converter has two cascaded H-bridge multilevel inverters with asymmetric sources. It produces a five-level output and has six switches instead of twenty-four. Two cascaded H-bridge (CHB) cells are used. These two CHB cells have different DC supplies. Because a five-level square wave more closely resembles a sine wave than a three-level square wave, it lowers harmonic distortion and power loss. The proposed three-phase, five-level cascaded H-bridge inverter is shown in Figure 2.

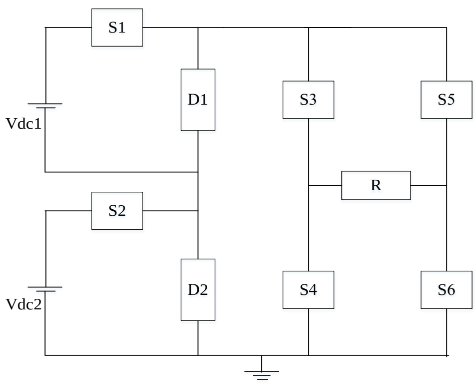


Fig. 2: Three phase cascaded H-bridge Multi level Inverter

Table 2: Switching states of proposed three phase five level cascaded H-bridge MLI

Mode	S1	S2	S3	S4	S5	S6	Switching voltage level
1	1	1	1	0	0	1	+(Vdc1+Vdc2)
2	0	1	1	0	0	1	+Vdc2
3	1	0	1	0	0	1	+Vdc1
4	0	0	0	0	0	0	0
5	0	0	1	1	1	1	0
6	1	0	0	1	1	0	-Vdc1
7	0	1	0	1	1	0	-Vdc2
8	1	1	0	1	1	0	-(Vdc1+Vdc2)

The primary switching components of IGBTs (insulated gate bipolar transistors) can be switched ON or OFF to produce the five voltage levels at the inverter's output as displayed in Table 2. The current flow in all operating modes indicated by active connections of components as shown in figure 3 to 10.

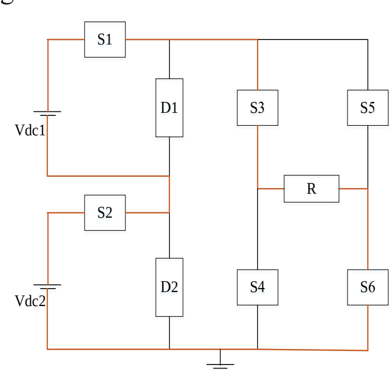


Fig. 3: Mode 1

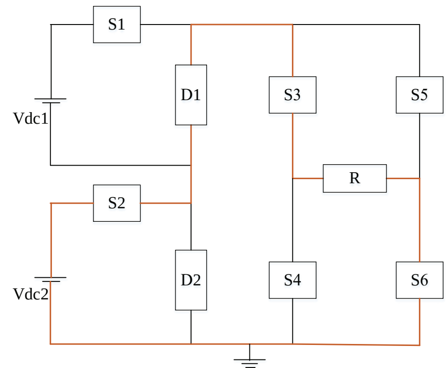


Fig. 4: Mode 2

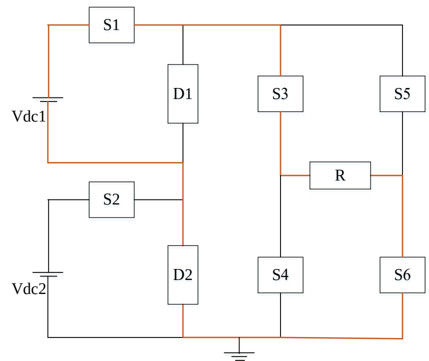


Fig. 5: Mode 3

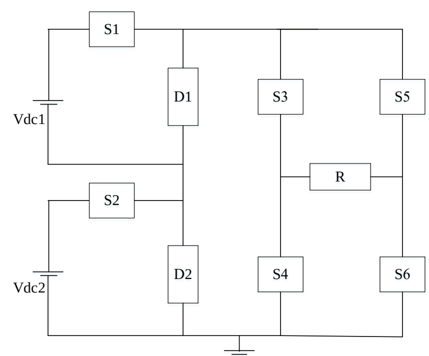


Fig. 6: Mode 4

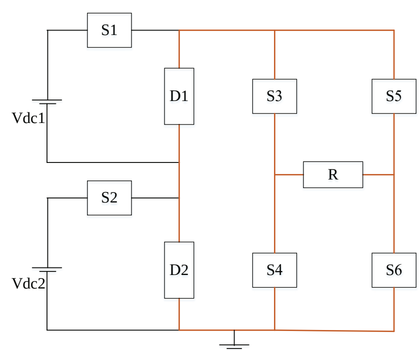


Fig. 7: Mode 5

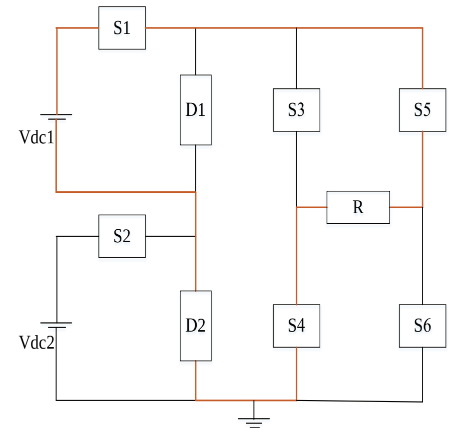


Fig. 8: Mode 6

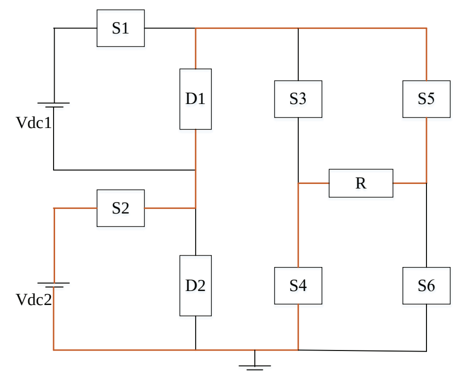


Fig. 9: Mode 7

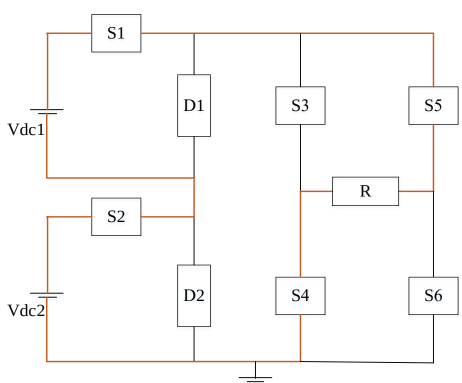


Fig. 10: Mode 8

DC/DC BOOST CONVERTER

A boost converter converts the input voltage to the required output voltage without the usage of a transformer. Figure 3 shows the key components of a boost converter, which are an inductor, a diode, a capacitor, and a high frequency switch. When combined, they provide load power at a voltage larger than the input voltage. Table 3 shows the components of the boost converter.

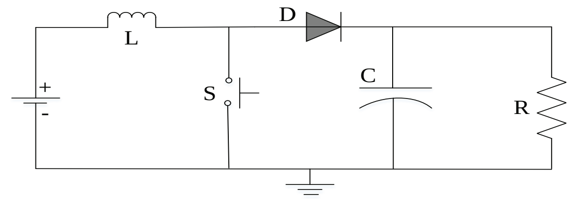


Fig. 11: Boost Converter

Table 3: Components for Boost converter

Components	Description
Inductor	0.1mH
Capacitor	1μF
Resistor	1ohm
Input DC voltage	100V
Diode resistance	0.001Ω
MOSFET-FET resistance	0.1Ω

SIMULATION RESULTS

The proposed topology implemented on MATLAB/ Simulink software. Figure 12 shows the Simulink model of three phase five level CHB inverter with boost converter.

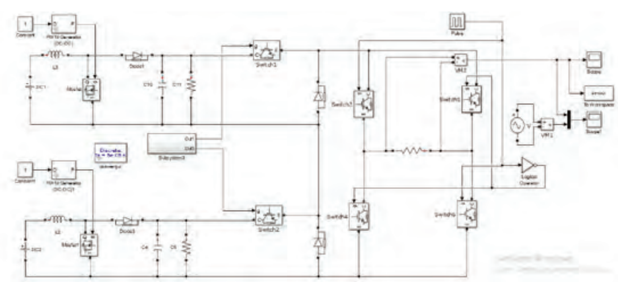


Fig. 12: Simulink model of three phase five level cascaded H-bridge inverter with DC/DC Boost converter

The input pulse voltage's MATLAB/Simulink model is displayed in Figure 13. Figures 14 and 15 display the output voltage of a three-phase, five-level inverter with and without a boost converter. Figures 16 and 17 display the AC voltage with and without a boost converter.

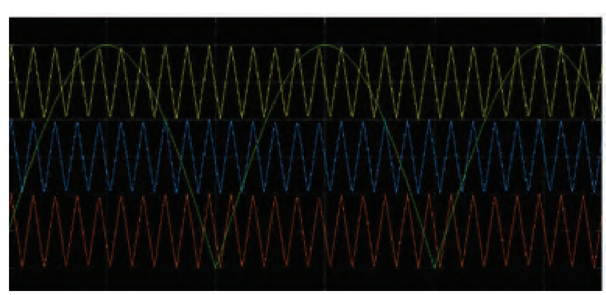


Fig. 13: Input pulse voltage

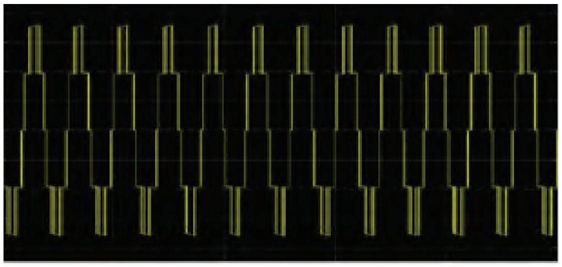


Fig. 14: Output voltage waveform of three phase five level cascaded H-bridge inverter with boost Converter

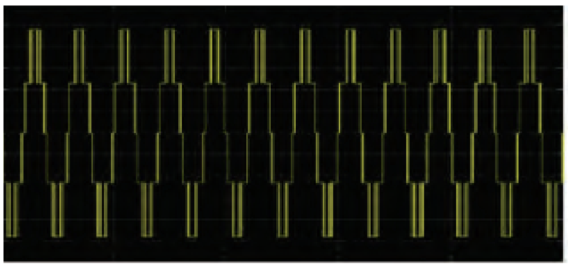


Fig. 15: Output voltage waveform of three phase five level cascaded H-bridge inverter without boost converter

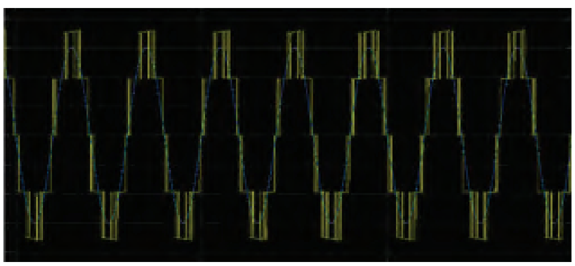


Fig. 16: AC voltage with boost converter

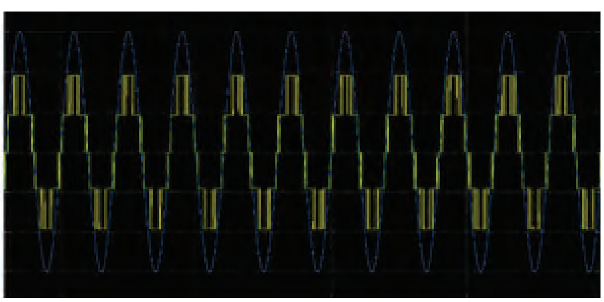


Fig. 17: AC voltage without boost converter

TOTAL HARMONIC DISTORTION (THD)

In a cascade multilayer inverter, the THD lowers as the number of levels rises; in a 5-level inverter, the THD is lower than in a 3-level inverter. When compared to the fundamental component of the voltage or current wave, THD is the total of all harmonic components of the waveform. The SIMULINK can be used to calculate THD. The THD performance of a three-phase, five-level inverter with a boost converter is displayed in Figure 18.

Table 4: Comparison of proposed topology with existing topologies

Five-level, three-phase inverters	Voltage Input(volt)	Voltage Output(volt)	THD (%)
6 switches (recommended)	100	180 (with a boost converter) 100 (without a boost converter)	27.34
9 switches [16]	100	120	28.46
12 switches 17]	100	120	67.12
15 switche[18]	100	110	43.91

The comparison between the suggested and current topologies is displayed in Table 4. With six switches, the suggested design performs well in terms of output voltage and THD. Six switches with a three-phase, five-level inverter without a cascading inverter produced a THD score of 145.92%. Compared with no cascading inverter, our suggested solution achieves a THD value of 27.34%. As a result, our suggested topology outperforms other topologies in terms of efficiency (72.66%).

CONCLUSION

This study proposes the design of a three-phase, five-level cascaded H-bridge inverter with a boost converter. The design lowers switch losses and is more economical. There are fewer power components in the suggested design. When the number of power components is reduced, the system's size, design, and cost all decrease. This topology demonstrates the switching modes of operation. The boost converter was then added to the input side, increasing the output voltage while lowering the overall harmonic distortion. MATLAB/Simulink is used to create a three-phase, five-level CHB inverter using asymmetric DC voltage sources. When compared to alternative topologies, our suggested topology yields superior outcomes. Therefore, compared to current designs, our suggested three phase, five level inverter with a six-switch topology is more efficient (72.66%). Future research will look at proposed multilevel inverters as a replacement for the traditional inverters utilised in microgrid applications.

REFERENCES

1. Liu J, Vazquez S, Wu L, Marquez A, Gao H, Franquelo LG. Extended state observer-based sliding-mode control for three-phase power converters. IEEE Transactions on Industrial Electronics. 2017 Jan; 64(1):22-31.
2. Vazquez S, Rodriguez J, Rivera M, Franquelo LG, Norambuena M. Model predictive control for power converters and drives: Advances and trends. IEEE Transactions on Industrial Electronics. 2017 Feb; 64(2):935-47.
3. Perreault DJ, Afridi KK, inventors; Massachusetts Institute of Technology, assignee. Resonant power converters using impedance control networks and related techniques. United States patent application US 10/090,772. 2018 Oct 2.
4. Wen B, Boroyevich D, Burgos R, Mattavelli P, Shen Z. Analysis of DQ small-signal impedance of grid-tied inverters. IEEE Transactions on Power Electronics. 2016 Jan; 31(1):675-87.
5. Mulkey SL, Cheng GS, Martinez M, inventors; CYBOENERGY Inc, assignee. Enclosure and message system of smart and scalable power inverters. United States patent US 9,331,488. 2016 May 3.
6. Townsend CD, Mirzaeva G, Goodwin GC. Deadtime compensation for model predictive control of power inverters. IEEE Transactions on Power Electronics. 2017 Sep; 32(9):7325-37.
7. Marzoughi, Alinaghi, Rolando Burgos, Dushan Boroyevich, and Yaosuo Xue. “Design and comparison of cascaded H-bridge, modular multilevel converter, and 5-l active neutral point clamped topologies for motor drive applications.” IEEE Transactions on Industry Applications 54, no. 2 (2018): 1404-1413.
8. Alshammaa AA, Noman AM, Addoweesh KE, Alabduljabbar AA, Alolah AI, inventors; King Saud University, assignee. Multilevel cascade hexagonal voltage source converter with isolated DC sources. United States patent application US 10/116,229. 2018 Oct 30.
9. Bakhshizadeh MK, Wang X, Blaabjerg F, Hjerrild J, Kocewiak Ł, Bak CL, Hesselbæk B. Couplings in phase domain impedance modeling of grid-connected converters. IEEE Transactions on Power Electronics. 2016 Oct; 31(10):6792-6.
10. Krishnan GV, Rajkumar MV, Hemalatha C. Modeling and Simulation of 13-level Cascaded Hybrid Multilevel Inverter with less number of Switches. International Journal of Innovative Studies in Sciences and Engineering Technology. 2016 Nov; 2(11):43-7.
11. Vahedi H, Labbé PA, Al-Haddad K. Sensor-less five-level packed U-cell (PUC5) inverter operating in stand-alone and grid-connected modes. IEEE Transactions on Industrial Informatics. 2016 Feb; 12(1):361-70.
12. Zhou D, Wang H, Blaabjerg F. Lifetime estimation of electrolytic capacitors in a fuel cell power converter at various confidence levels. InPower Electronics Conference (SPEC), IEEE Annual Southern 2016 Dec 5 (pp. 1-6). IEEE.
13. Bhaskar MS, Kulkarni RM, Padmanaban S, Siano P, Blaabjerg F. Hybrid non-isolated and non-inverting Nx interleaved DC-DC multilevel boost converter for renewable energy applications. InEnvironment and Electrical Engineering (EEEIC), 2016 IEEE 16th International Conference on 2016 Jun 7 (pp. 1-6). IEEE.
14. Dixon J, Moran L. Multilevel inverter, based on multi-stage connection of three-level converters scaled in power of three. InIECON 02 [Industrial Electronics Society, IEEE 2002 28th Annual Conference of the] 2002 Nov 5 (Vol. 2, pp. 886-891). IEEE.
15. Emadi A, Ehsani M. Multi-converter power electronic systems: Definition and applications. InPower Electronics Specialists Conference, 2001. PESC. 2001 IEEE 32nd Annual 2001 (Vol. 2, pp. 1230-1236). IEEE.
16. Perez-Santiago, Anthony, Melvin Lugo-Alvarez, Luis Sepulveda-Hernandez, and Eduardo I. Ortiz-Rivera. “Design of a nine-switch three-phase multilevel inverter with a PV array source.” In 2018 IEEE Texas Power and Energy Conference (TPEC), pp. 1-6. IEEE, 2018.
17. Chen, J. F., W. C. Hsu, Y. M. Wu, Y. P. Hsieh, H. Y. Huang, J. S. Li, and Y. L. Chen. “A novel three-phase five-level inverter.” In 2016 IEEE 8th international power electronics and motion control conference (IPEMC-ECCE Asia), pp. 485-491. IEEE, 2016.
18. Ahmed, Mafaz, Tila Muhammad, Rafi Uzman, and Omais Khan. “A Novel Asymmetric Three Phase Multilevel Inverter with Reduced Switches.” Sukkur IBA Journal of Emerging Technologies 1, no. 2 (2019): 1-11.

Comprehensive Analysis of a Wideband Spherical Dipole Antenna Using CST Microwave Studio

Ankit H. Deogharkar
Dept. of Electronics and Telecommunication
Atharva College of Engineering
Mumbai, Maharashtra
✉ ankitdeogharkar@atharvacoe.ac.in

Jyoti Mali
Dept. of Electronics and Telecommunication
Atharva College of Engineering
Mumbai, Maharashtra
✉ jyotimali@atharvacoe.ac.in

Bhavin Shah
Dept. of Electronics and Telecommunication
Atharva College of Engineering
Mumbai, Maharashtra
✉ mr.bhavin.ace@gmail.com

Mohan Kumar
Dept. of Electronics and Telecommunication
Atharva College of Engineering
Mumbai, Maharashtra
✉ mohan841301@gmail.com

ABSTRACT

This paper introduces a modified design for the Spherical Dipole Antenna, enhancing its performance over the conventional structure. While the traditional antenna offers narrow bandwidth characteristics, the proposed modification significantly expands the bandwidth and achieves gain levels comparable to those of a standard dipole antenna. This antenna is specifically tailored for repeatable shielding effectiveness measurements. Its design is both area-efficient and battery-powered, with an optical link for streamlined connectivity. The antenna operates within a frequency range of 30 MHz to 4 GHz, delivering a bandwidth of 515 MHz in the lower frequency range and 1.88 GHz in the higher range. The electric field strengths at 0.835 GHz and 2.6 GHz are recorded at 2940 V/m and 3316 V/m, respectively. All simulations of the antenna were conducted using CST Microwave Studio (Computer Simulation Technology).

***KEYWORDS** : Spherical dipole antenna (SDA), Bandwidth, Gain, Shielding effectiveness, Optically linked, Area-efficient, CST-MW studio.*

INTRODUCTION

With the rapid growth of telecommunication devices and the increasing digitization of industries, there is an urgent need to mitigate the electromagnetic interference (EMI) generated by these technologies. Electromagnetic radiation from electronic equipment can lead to sensitive data leakage and disrupt neighboring devices. Conversely, external electromagnetic fields can induce disturbances and damage to the equipment. To address these challenges, it is essential to enclose such equipment within shielded environments. However, accurately assessing the performance of these shields is equally critical. This necessitates the design of specialized antennas for shielding effectiveness evaluation—antennas that can operate at high frequencies and offer a wide and offer a wide bandwidth to ensure precise and reliable measurements. Figure 1 shows the structure of conventional spherical dipole antenna [1]. The antenna design comprises two hemispheres electrically isolated by a threaded dielectric spacer. The lower hemisphere functions as a shielding enclosure for electronic components, including an optical-to-electrical (O/E) converter, a low-noise amplifier (LNA), and a battery power supply. The system utilizes optical feeding, which enables the transmission of both the optical information signal and optical power. This approach leverages the inherent immunity of optical fibers to electromagnetic interference, ensuring reliable signal transmission. After the optical-to-electrical conversion, the feed is delivered through a central post, ensuring efficient and interference-free operation.

The spherical dipole antenna is well-suited for evaluating shielding effectiveness. It provides the following advantages [5]:

1. The spherical geometry offers ample internal volume, enabling the integration of electronic components

within the antenna, thereby making it a self-contained system.

2. The spherical design eliminates sharp edges and ensures optimal symmetry for improved performance.

3. The circuitry and power sources are housed internally, preventing any interference with the established radiation field.

4. The compact size of the antenna makes it ideal for integration into small shielded enclosures.

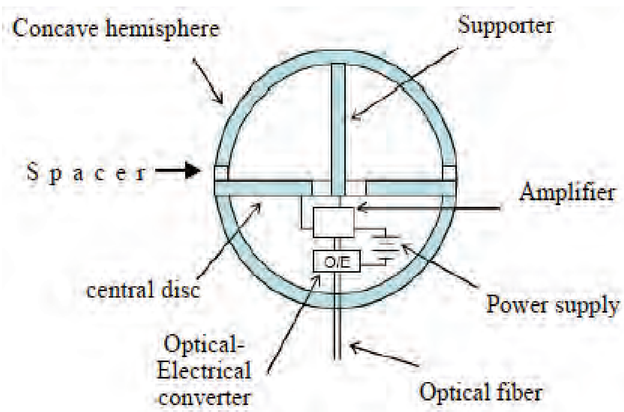


Fig. 1: Conventional spherical dipole antenna

SHIELDED ENCLOSURES

In recent years, the loss of information due to electromagnetic radiation from electrical devices has become a significant concern. Additionally, the increasing vulnerability of electronic equipment to electromagnetic attacks further exacerbates the issue. One effective approach to mitigating these challenges is the use of electromagnetic shielded enclosures. Equally crucial is the accurate evaluation of the shielding effectiveness [8].

Electrical devices are housed within these shields to minimize the impact of electromagnetic interference (EMI). Shielded enclosures, typically constructed from metal, protect the internal components from external electric fields while preventing radiation from escaping. These enclosures are particularly essential in high-radiation environments such as military zones, aircraft, ships, and radar facilities, where they serve to safeguard the sensitive circuitry from harmful electromagnetic exposure.

Following figures shows the structure of shielded enclosures.

2(a)

2(b)

Figure 2a & 2b: External and internal view of Shielded enclosures respectively.

SHIELDING EFFECTIVENESS

Shielding effectiveness is quantified as the ratio of the electric field detected by the antenna outside the shield to the electric field measured inside the shield. The following figures illustrate the procedure for assessing the shielding effectiveness of enclosures. In Figure 3.1, the antenna is positioned within an anechoic chamber [9], which is designed to prevent electromagnetic wave reflections, ensuring that the waves are absorbed by the chamber's walls. In Figure 3.2, the spherical dipole antenna is placed inside the shielded enclosure under test. The electric field measurements obtained in both scenarios are denoted as E1 (outside the shield) and E2 (inside the shield), respectively.

E1: electric field measured by the antenna without enclosure.

E2: electric field measured by the antenna with enclosure.

We can easily calculate the shielding effectiveness by following expression [3]:

Shielding effectiveness : $\mathrm{SE} = E_1 - E_2$ (1)

OR $\mathrm{SE} = 20\log\left(\frac{E_1}{E_2}\right)$ dB (2)

Step1:

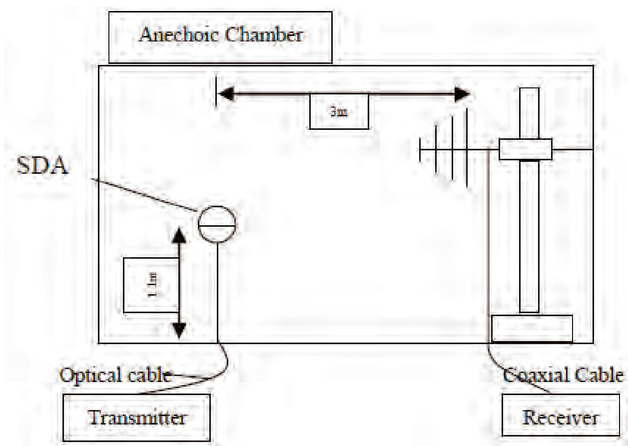


Fig. 3: Electric field strength Reference Measurement (E1)

Step 2:

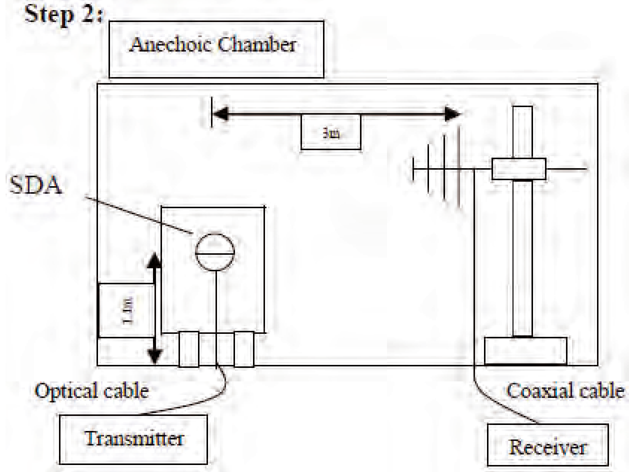


Fig. 4: Electric field strength with shielded enclosure Measurement (E2)

SIMULATION RESULTS OF CONVENTIONAL SPHERICAL DIPOLE ANTENNA

The proposed solutions were developed using CST Microwave Studio, with simulation results indicating a narrow bandwidth for a 15 cm diameter configuration. These results include the S11 parameter analysis over a frequency range spanning from 30 MHz to 4 GHz.

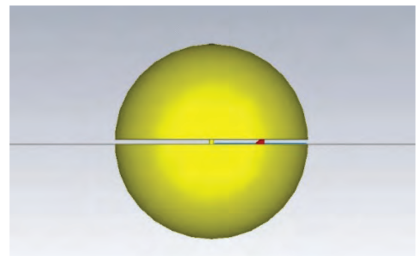


Fig. 5: 3D model of conventional spherical dipole antenna in CST

As depicted in Figure 6, the S11 curve exhibits multiple resonances and lacks the desired wideband characteristics. To address this limitation, a wideband spherical dipole antenna model has been developed.

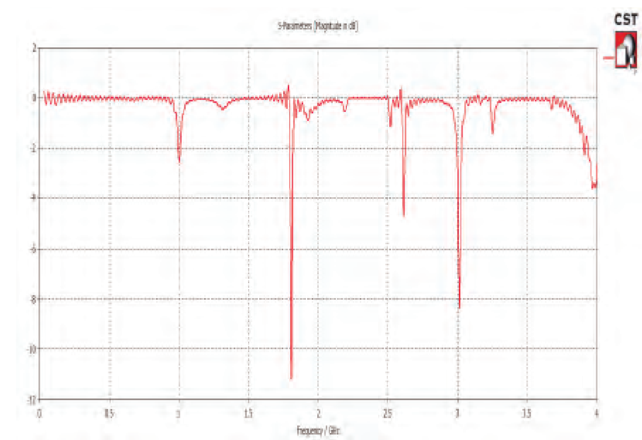


Fig. 6: S11 parameter

STRUCTURE OF WIDEBAND SPHERICAL DIPOLE ANTENNA

The wideband configuration is presented in Figure 5. To achieve the desired wide bandwidth, modifications to the conventional structure are required. One effective method to broaden the antenna's bandwidth is by increasing the volume it occupies. However, expanding the diameter results in a larger overall structure, which may not be suitable for compact shielded enclosures.

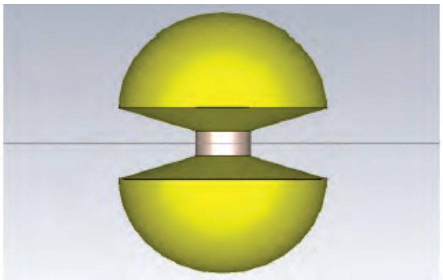


Fig. 7: Wideband Spherical Dipole Antenna

Increasing the thickness of the sphere would complicate the accommodation of the internal circuitry. Therefore, expanding the area through these two approaches is not feasible. As a result, the only viable solution is to increase the separation between the hemispheres and incorporate cones on both ends. In this revised design, the gap between the hemispheres has been enlarged, and two cones are attached to facilitate the path for the feed signal. A dielectric spacer is used to electrically isolate the hemispheres. The circuitry, including the O/E converter, amplifier, and power supply, is housed within the lower hemisphere. Both solid and hollow cones yield similar simulation results. A final gap of 56 mm has been selected to achieve the desired performance, as smaller gaps favor higher frequencies, while larger gaps enhance lower frequency performance. This gap provides optimal performance across both frequency ranges. However, further increases in the gap would enlarge the overall structure, making it unsuitable for compact shielded enclosures.

SIMULATION RESULTS OF WIDEBAND SPHERICAL DIPOLE ANTENNA

The simulation results consist of S11 parameter, electric field in V/m and the polar plots.

Following figure shows the S11 parameter of the antenna.

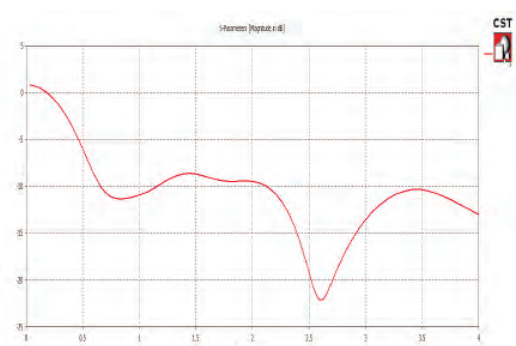


Fig. 8: S11 parameter

In the lower frequency range, spanning from 653 MHz to 1.16 GHz, the S11 parameter drops below -10 dB, reaching a minimum of -12 dB, resulting in an approximate bandwidth of 515 MHz. In the higher frequency range, from 2.11 GHz to 4 GHz and beyond, the S11 parameter decreases below -10 dB, reaching -22.5 dB, providing a bandwidth of 1.887 GHz or greater. Additionally, if the threshold is adjusted to -9 dB, the overall bandwidth spans from 653 MHz to 4 GHz and beyond, as illustrated in the Figure 6.1. Following figures 6.2 & 6.3 shows the electric fields at 0.835GHz and 2.6GHz respectively.

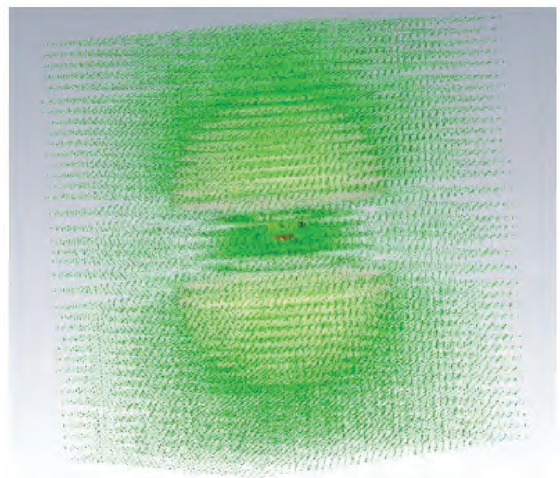


Fig 9: Electric field at 0.835 GHz

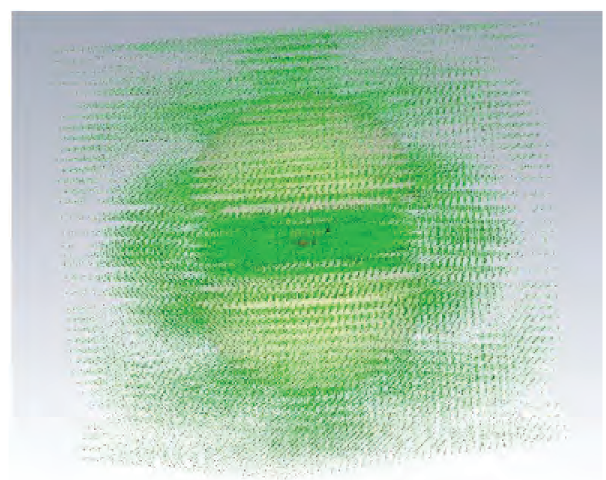


Fig. 10: Electric field at 2.6 GHz

The electric field obtained at 0.835GHz and 2.6GHz are 2940V/m and 3316 V/m respectively.

Following figures shows the polar plot at 0.835GHz and 2.6GHz respectively.

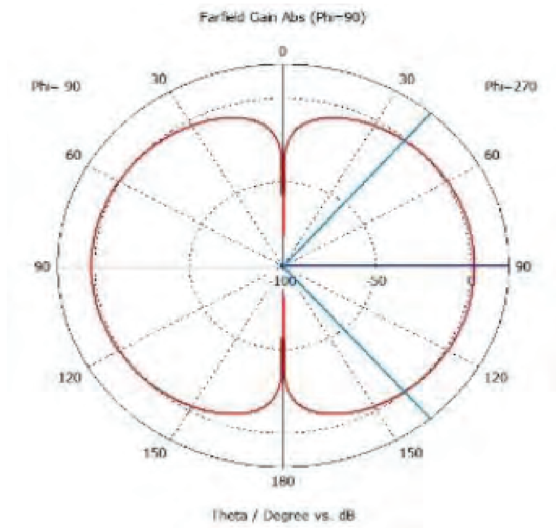


Fig. 11: Polar plot (f = 0.835GHz)

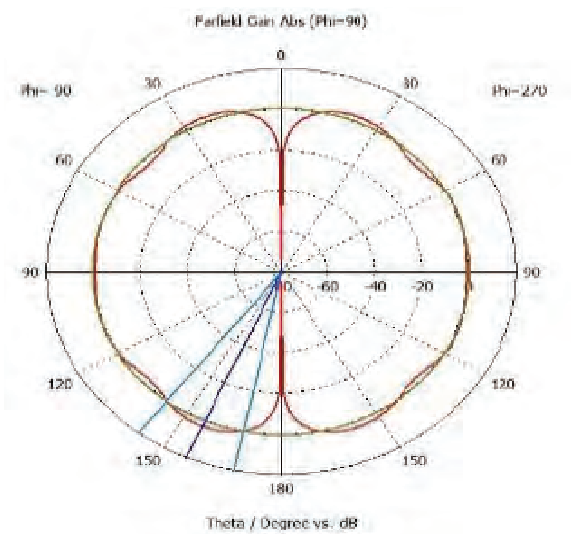


Fig. 12: Polar plot (f = 2.6GHz)

CONCLUSION

The conventional spherical dipole antenna exhibits a narrow frequency response, while the wideband spherical dipole antenna is designed to cover a broader frequency range. This enhancement is achieved by modifying the traditional structure, specifically by increasing the gap between the hemispheres. Another method to achieve broadband characteristics is by expanding the volume occupied by the antenna, typically by increasing the sphere's diameter. However, this approach is impractical for small shielded enclosures due to the resulting increase in size. The chosen structural modification effectively addresses the need for a wideband response without significantly enlarging the overall antenna size. Importantly, this modification does not alter the primary function of the spherical dipole antenna—shielding effectiveness measurement—ensuring that the antenna's core purpose remains intact while extending its operational bandwidth. Future work could explore further optimization of the structure to enhance performance at both lower and higher frequency ranges, as well as investigate integration with advanced materials for improved shielding effectiveness and efficiency in various environments.

ACKNOWLEDGEMENT

The authors would like to thank Atharva College of Engineering, Malad West, Mumbai for their support and providing required resources for our work.

REFERENCES

1. Shigetaka Itakura Shuichi Nabekura Yoshinobu Okano Masamitsu Tokuda , "An optical feeding antenna with a wider bandwidth than conventional spherical dipole antennas," IEEE, 1-28-1 Tokyo, 158-8557 Japan , 2005.
2. G. Koepke L. D. Driver K. Cavcey K. Masterson R. Johnk M. Kanda, Standard Spherical Dipole Source, Washington: National Institute of Standards and Technology Technical Note , 1991.
3. Dale G. Svetanoff, Richard B. Schulz, "IEEE Standard Method for Measuring the Effectiveness of Electromagnetic Shielding Enclosures," IEEE, 345 East 47th Street, New York, NY 10017-2394, USA, 1997.
4. Toshinori Mori, Kaoru Shinozaki, Yoshimasa Kaneko , "Improving Shielding Effectiveness Measurements with a Spherical Dipole Antenna," IEEE, Midori-Cho, Musashino-Shi, Tokyo 180 Japan , 1994.
5. Bruce Archambeault,Mark Seth , "A New Standard Radiator for Shielding Effectiveness Measurements," IEEE, 1992.
6. Alan Ghasiani, "Shielding Effectiveness Test For RadiaShield Technologies," US Tech , 2010.
7. X. C. Nie and N. Yuan , "Accurate Modeling of Monopole Antennas in Shielded Enclosures with Apertures," Progress In Electromagnetics Research, PIER 79, TX 77204, USA, 2008.
8. "Ultra Wideband Spherical Dipole Antenna," NTT Energy and Environment Systems Laboratories, 2004.

[9] "Spherical Dipole Antenna," NTT Advanced Technology Corporation, Japan, 2010.

Design of an Arterial Pulse Detection System Using Force Sensitive Resistors for Prakriti Classification

Gauri Vaidya
Atharva College of Engineering
Mumbai, Maharashtra

Neeta Moharkar
Atharva College of Engineering
Mumbai, Maharashtra

Sowmya Kini Prabhu
Atharva College of Engineering
Mumbai, Maharashtra

Purvi Agarwal
Atharva College of Engineering
Mumbai, Maharashtra

ABSTRACT

The ancient Indian medical system known as Ayurveda is where the idea of Prakriti first emerged. The balance of the three doshas—Pitta, Kapha, and Vata—determines each person's distinct physical and mental makeup. An evaluation of Prakriti through pulse diagnosis, which offers information about an individual's health, habits, and predispositions, is a common procedure in traditional Ayurvedic medicine. In order to ascertain an individual's Prakriti, this research explores the design of a modern arterial pulse detection system. Based on arterial pulse data, the system classifies a person's Prakriti using a combination of physiological measurements, machine learning algorithms, and signal processing techniques. The difficulties and possibilities of combining conventional Ayurvedic principles with contemporary diagnostic technology are covered in the study.

KEYWORDS *: Ayurveda, Prakriti analysis, Pulse pattern detection, Feature extraction, MATLAB implementation.*

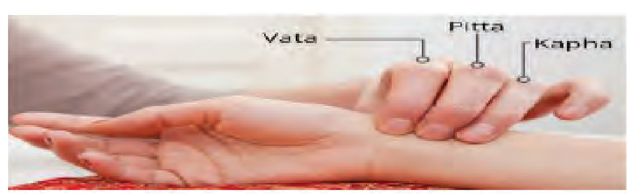
INTRODUCTION

The ancient medical system known as Ayurveda has traditionally classified people according to their physical and psychological characteristics using the idea of Prakriti. Pulse examination is one of the indications that are typically observed in order to perform Prakriti assessment. According to Ayurveda, the arterial pulse, or Nadi, is a reflection of the dosha balance and can reveal information about a person's personality and general health. Although Ayurveda has a long history of using pulse-based diagnostics, there hasn't been much incorporation with contemporary medical technologies. The development of a system that can automatically detect Prakriti by measuring arterial pulses is made possible by recent developments in biosensors, data processing, and machine learning. In order to identify Prakriti, this research suggests designing a system that combines data analysis methods with sensors for measuring pulses. The intention is to provide an objective, non-invasive technique for determining Prakriti that may support individualized treatment and well-being.

The Concept of Prakriti

The term "Prakriti" in Ayurveda describes the special fusion of the three doshas—Pitta, Kapha, and Vata. A person's health and behavior are thought to be determined by the balance or imbalance of these doshas, which are thought to regulate a number of physiological and psychological processes.

Pulse Diagnosis in Ayurveda: Nadi Pariksha, or pulse diagnosis, is a method in which medical professionals measure the pulse at several body locations, mostly the wrist. The dominance of each dosha is ascertained by analyzing the pulse's rhythm, intensity, and frequency. Understanding the unique pulse rhythms produced by Vata, Pitta, and Kapha might reveal information about a person's Prakriti.



LITERATURE SURVEY

The domains of biomedical engineering and wearable health technology have produced a number of techniques for identifying and evaluating the arterial pulse in recent years. To track pulse characteristics, these devices usually use pressure sensors, electrocardiography (ECG), or photoplethysmography (PPG). Various systems have been developed for acquiring wrist pulse measurements. These include force-sensitive resistors [1], Ultarsonic Sensor[2], pressure sensor[3], ultrasound transmitter-receiver pairs, [4] has proposed a pulse acquisition device employing a pulse sensor, which determines 17 feature parameters of the pulse image. For disease diagnosis, multiple kernel learning (MKL) and simple MKL algorithms [5] Author has proposed the use of both crisp and fuzzy clustering algorithms under supervised and unsupervised learning frameworks. They applied these methods using the impedance plethysmographic technique to identify complex regularities and relationships in radial pulse patterns. Additionally, a piezoelectric film sensor.

SYSTEM ARCHITECTURE

Overview of the Proposed System By collecting pulse data and applying machine learning models, the proposed arterial pulse detection system seeks to classify a person's Prakriti according to the features of their pulse. Three important wrist locations—representing the Vata, Pitta, and Kapha regions—are used to assess pulse pressure in this paper's design for an arterial pulse detection system that uses force-sensitive resistors (FSRs). After being digitized and examined by an Analog to Digital Converter (ADC), the pulse data gathered by these sensors is shown on an LCD screen and subjected to additional analysis in MATLAB over the RS232 communication protocol. This method seeks to offer a more objective and consistent way to diagnose pulses, which could help with Prakriti classification and support applications for individualized treatment.

Pulse detection

Due to human body and mental movements, the wrist pulse is extremely weak and readily influenced by noise. These factors are taken into account while choosing a force-sensitive resistor for pulse detection that is responsive to low frequency signals. A 10kohm resistor is connected to the FSR.Velcro is used to secure the sensor to the wrist, ensuring steady pressure. It keeps the sensor from shifting around and getting hurt. The wrist sensors are positioned in accordance with the proper vata, pitta, and kapha positions. Figure 4 depicts the hardware circuitry architecture.

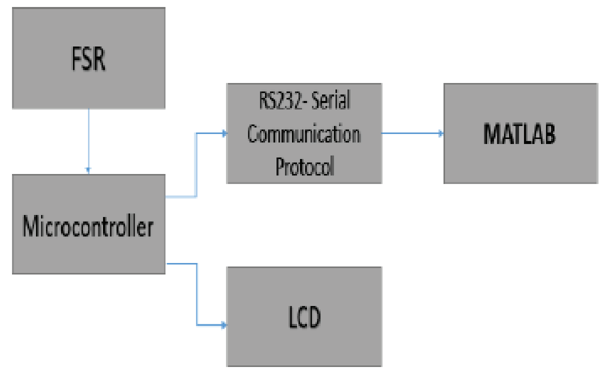


Fig. 1. Block Diagram

Force Sensitive Resistor

The amount of force or pressure given to FSRs causes them to alter their resistance. Three FSRs representing Vata, Pitta, and Kapha are positioned at the three wrist sites in this system. These sensors will provide real-time pulse amplitude data by measuring the pressure that the arterial pulse exerts at these points.

Converter from Analog to Digital (ADC)

The ADC then converts the analog signal produced by the FSRs into a digital signal for additional processing. Accurate measurements of the pressure changes over time are made possible by the ADC.

Microcontroller

The FSRs are interfaced with, the data is processed, and the results are displayed by a microcontroller, like as Arduino or Raspberry Pi. After processing the pulse amplitude data, it transfers the data to an LCD screen for real-time display after gathering it from the ADC.

LCD Screen

The amplitude of the pulse at each of the three locations appears on the LCD panel. Depending on the extracted data, it could include other parameters, including the pulse rate.

MATLAB for Signal Analysis

Using the RS232 communication protocol, the system transmits pulse data to a computer running MATLAB. MATLAB processes the pulse signals, visualizes the pulse waveforms, and computes various characteristics such as the pulse rate and amplitude. MATLAB can also be used for advanced signal processing, including filtering, peak detection, and feature extraction, which are necessary for the classification of Prakriti based on pulse patterns. The system sends pulse data to a computer running MATLAB through the RS232 connection protocol. The pulse signals are processed using MATLAB, which additionally displays the pulse waveforms and calculates a number of parameters like the pulse rate and amplitude. In order to classify Prakriti based on pulse patterns, complex signal processing techniques including filtering, peak identification, and feature extraction can also be performed with MATLAB.

Protocol for RS232 Communication

Real-time pulse data transmission from the microcontroller to the computer for additional analysis is made possible by the system's communication with a MATLAB environment using the RS232 interface. This protocol offers a dependable means of data transmission over small distances and is frequently used for serial communication between a microcontroller and a computer.

Feature Extraction

The cleaned pulse data will be used to extract important properties such pulse rate, pulse waveform, amplitude, and duration. Furthermore, calculations will be made for time-domain features like mean and variance as well as frequency-domain features like power spectral density. Adjust pulse characteristics to take individual and situational variability into consideration.

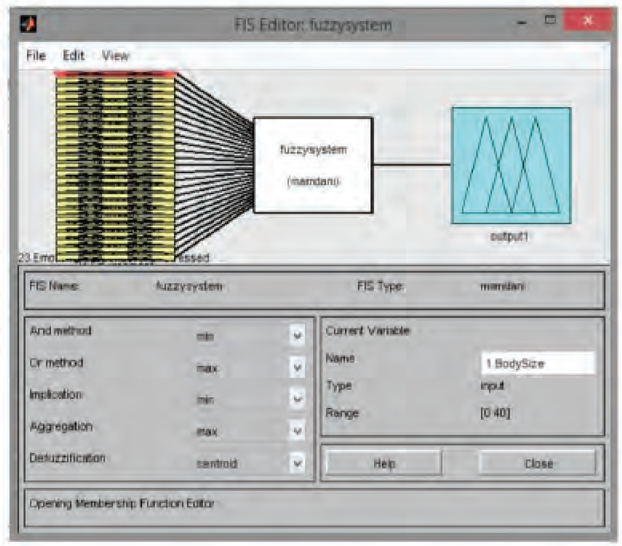
CLASSIFICATION VIA FUZZY LOGIC SYSTEM

A Fuzzy Logic System (FLS) is a computational framework based on the concept of fuzzy set theory, which allows reasoning with uncertain or imprecise information. Unlike traditional binary logic (true/false), fuzzy logic uses degrees of truth represented by values between 0 and 1. There are two systems in Fuzzy interface system.

1. Mamdani system

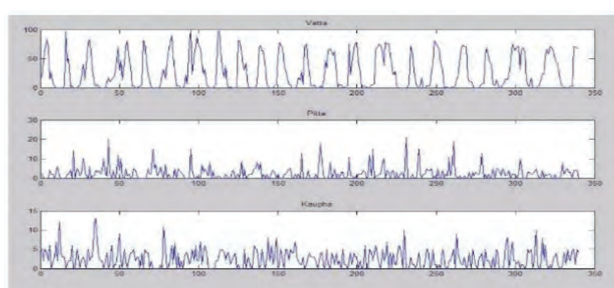
2. Sugeno System

The Mamdani system is one of the most widely used fuzzy inference systems, especially for systems where human reasoning needs to be emulated. Here we have used this method for calculating Prakriti of a person based on their inputs. A Mamdani system, also known as the Mamdani fuzzy inference system (FIS), is a widely used approach in fuzzy logic for decision-making and control applications. Crisp (numerical) inputs are converted into fuzzy values using membership functions. The system uses the fuzzified inputs and applies the fuzzy rules to infer fuzzy output sets. Logical operations like AND, OR, and NOT are applied based on the rules. The fuzzy outputs of all rules are combined into a single fuzzy set. This process involves blending the consequences of individual rules.



RESULTS AND ANALYSIS

To determine a person's prakriti, questionnaires are provided to gather relevant information. Pulse patterns specific to each prakriti are identified by categorizing individuals into groups based specific age ranges. Data has been collected using various individual parameters. For each age group, Pulse rate data is recorded for one minute for each individual. Based on their prakriti, pulse patterns are analyzed. Threshold values are established for pulse rate measurement, with amplitude and frequency spectrum used to identify irregularities. The results highlight the pulse patterns corresponding to Vata, Pitta, and Kapha prakriti types.



CONCLUSION

This paper predicts a person's prakriti based on their individual characteristics. An imbalance in the tridoshas leads to irregularities in the pulse, which can be used to detect diseases. The results indicate that normal health conditions exhibit consistent pulse patterns, whereas abnormal health conditions show broader variations in the pulse patterns. Future research can focus on enhancing prediction capabilities using fuzzy logic systems.

REFERENCE

1. Gauri Vaidya, Sonali Navghare, and Preeti Bajaj. "Design of Arterial Pulse Detection System for Detection of Prakriti of Person." International Journal of Engineering Research and General Science 3.2 (2015): 779-783.
2. Goyal, Krittika, and Akhil Gupta. "A literature survey on different types of pulse based sensor for acquisition of pulse." International Journal of Control Theory and Applications 9.10 (2016): 361-365.
3. Narendrakumar, K., and B. Ramesh. "Design of radial artery pulse sensor system for ayurveda disease diagnosis." International Journal of Innovative Technology and Exploring Engineering (IJITEE) 9.2S (2019).
4. Lisheng Xu, Max Q- H Meng, kuanquan Wang, Wang Lu, Naimin Li, "Pulse Image Recognition Using Fuzzy Neural Network" Proceedings of the 29th Annual International Conference of the IEEE EMBS Cité Internationale, Lyon, France August 23-26, 2010.
5. Yusuke Yamana, Sosuke Tsukamoto, Koji Mukai, Hiromichi Maki, Hidekuni Ogawa, and Yoshihar Yonezawa, "Sensor for Monitoring Pulse Rate, Respiration Rhythm, and Body Movement in Bed" Members, IEEE,33rd Annual International Conference of the IEEE EMBS, August 30 - September 3, 2011
6. Herztman, "Photoelectric Plethysmography of the fingers and toes in man", Proceedings of the Society for Experimental Biology & Medicine 37:1622-1637 (1937).

Enhancing Image Classification with Pre-trained Convolutional Neural Networks and Grad-CAM Visualization for Explainable AI

Ulhaskumar Gokhale
Electronics & Telecommunication
Atharva College of Engineering
Mumbai, Maharashtra
✉ ulhaskumargokhalei@atharvacoe.ac.in

Charmi Chaniyara
Infromation Technology
Atharva College of Engineering
Mumbai, Maharashtra
✉ charmichaniyara@atharvacoe.ac.in

Pranoti Nage
Infromation Technology
Atharva College of Engineering
Mumbai, Maharashtra
✉ pranotinage@atharvacoe.ac.in

Sandip Zade
Electronics & Telecommunication
Atharva College of Engineering
Mumbai, Maharashtra
✉ sandipzade@atharvacoe.ac.in

ABSTRACT

This study explores an innovative combination of pre-trained Convolutional Neural Networks (CNNs) and Gradient-weighted Class Activation Mapping (Grad-CAM) to improve both the effectiveness of image classification tasks and the interpretability of the underlying models. Pre-trained CNNs harness the powerful feature extraction capabilities of models trained on large-scale datasets, enhancing computational efficiency and minimizing the need for extensive data collection and training from scratch. Grad-CAM, on the other hand, complements this by creating intuitive heat maps that visually highlight the regions of input images most influential to the model's predictions. These visual explanations not only clarify the neural network's decision-making process but also foster greater trust in AI-powered systems. The proposed methodology is thoroughly evaluated using well-established benchmark datasets, demonstrating its ability to achieve cutting-edge classification performance while providing transparent and interpretable insights into the model's behavior. By fostering trust and understanding in AI systems, this approach has the potential to support broader adoption and ethical deployment of image classification models across diverse applications.

KEYWORDS *: Pre-trained CNNs, Grad-CAM, Explainable AI.*

INTRODUCTION

The widespread adoption of deep learning in image classification tasks has revolutionized various domains, including healthcare, autonomous vehicles, and security. Pre-trained CNNs have emerged as a cornerstone for efficient and robust feature extraction. However, the opacity of deep learning models raises concerns about trust and accountability, especially in critical applications. Grad-CAM, an explainability tool, provides a mechanism to visualize model decisions, enabling stakeholders to interpret and validate predictions. This paper investigates the synergistic use of pre-trained CNNs and Grad-CAM for accurate and explainable image classification.

REVIEW OF LITERATURE

Pre-trained Convolutional Neural Networks (CNNs)

Pre-trained Convolutional Neural Networks (CNNs) are a cornerstone of modern computer vision. These models have been trained on massive datasets, often containing millions of images, to recognize a wide range of visual features. This pre-training process allows them to serve as powerful feature extractors and classifiers for a variety of tasks.

Transfer Learning with Pre-trained CNNs

Access to extensive datasets and high-performance GPU resources for training on large data batches is often limited. As an alternative to training a CNN model from the ground up, the technique of transfer learning can be employed. [14].

It involves fine-tuning the pre-trained model on a new dataset, often with fewer layers frozen to allow for adaptation to the new task.

Pre-trained CNNs, such as GoogLeNet, ResNet50, and InceptionV3, have shown remarkable performance in image classification tasks by transferring learned features from large-scale datasets like ImageNet to target domains. Fine-tuning these models enables their adaptation to specific datasets, reducing computational requirements and improving accuracy

Authors in [11] demonstrated that leveraging features extracted from deep pre-trained CNNs trained on ImageNet can yield exceptional performance across a wide range of image classification tasks, often surpassing traditional classification methods. Their findings highlight that the representations learned by deep networks are highly versatile, enabling effective transfer learning across various domains. This approach proves particularly beneficial in scenarios with limited labeled data available for the target task. First, confirm that you have the correct template for your paper size. This template has been tailored for output on the A4 paper size. If you are using US letter-sized paper, please close this file and download the Microsoft Word, Letter file.

Explainable AI and Grad-CAM

The authors [12] discuss the importance of making AI systems more transparent and interpretable. The authors highlight the impressive performance of AI systems in various complex tasks but emphasize the need for explainability, especially in critical applications like medical diagnosis and self-driving cars.

Explainable AI (XAI) methods, including Grad-CAM, tackle the interpretability issues associated with deep learning models by visualizing the areas of input images that

are significant for class differentiation. The Grad-CAM technique generates a coarse localization map by utilizing the gradients of a specific target concept that pass through the final convolutional layer. This map emphasizes the crucial regions within the image that contribute to the prediction of that concept. [1].

Grad-CAM computes gradients of target outputs concerning feature maps to generate heat maps, highlighting critical areas influencing predictions. This enhances user trust and facilitates debugging of models [2].

Comparison with Other Explainability Techniques: Grad-CAM is often compared to techniques such as Local Interpretable Model-Agnostic Explanations (LIME) and SHapley Additive exPlanations (SHAP) and other techniques.

LIME: LIME creates surrogate models to approximate the behavior of the original model around a specific prediction. While LIME is model-agnostic and offers flexibility across various machine learning models, it operates on perturbed samples and does not directly visualize regions of interest within images, making it less intuitive for image classification tasks compared to Grad-CAM. [15]

SHAP: SHAP provides a unified measure of feature importance based on Shapley values from cooperative game theory. It is highly comprehensive and interpretable but is computationally expensive, especially for deep learning models. Unlike Grad-CAM, SHAP does not inherently provide spatial insights for image data [15].

Occlusion Sensitivity: Occlusion sensitivity involves systematically occluding parts of the input image and observing the impact on the model's prediction. While this technique provides valuable insights into which regions of an image are critical for predictions, it is computationally intensive and lacks the visual immediacy of Grad-CAM heat maps [11].

Grad-CAM: Grad-CAM leverages the gradients of target outputs with respect to the final convolutional layer to produce heat maps. This approach is specifically designed for CNNs and provides spatially meaningful explanations, making it uniquely suited for image classification tasks. Additionally, Grad-CAM is computationally efficient and visually intuitive. [1]

This comparison highlights Grad-CAM's unique capability to generate spatially localized visualizations, addressing the interpretability challenges specific to deep learning in computer vision. The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

PROPOSED SYSTEM

We evaluate the proposed approach using benchmark datasets, including: ImageNet (subset): A large-scale dataset containing diverse image categories. Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

A. Pre-trained CNN Architecture We utilize well-known pre-trained models such as GoogLeNet and ResNet-50. GoogLeNet is a convolutional neural network developed by Google, specifically trained on a subset of the ImageNet database, which plays a crucial role in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). This model, which consists of 22 layers and has been trained on over a million images, is capable of categorizing images into 1,000 different object classes, including various tools, animals, and common household items. Its architecture is designed to efficiently extract detailed feature representations from a wide array of image inputs. Upon processing an input image, the model not only identifies the object but also provides a corresponding label along with the probabilities for each category. ResNet-50 is another pre-trained model, also trained on a subset of the ImageNet database. It achieved notable success by winning the ILSVRC competition in 2015. This model features a 50-layer residual network with a total of 177 layers. ResNet-50 is designed to address the vanishing gradient problem in deeper networks by using residual connections. Like GoogLeNet, it classifies images into 1000 categories and has been trained on a dataset exceeding a million images.

In our particular application, we implement fine-tuning by locking the initial layers of the pre-trained models while retraining the subsequent layers using the target dataset. This approach allows us to customize the models for the specific task at hand.

EXPERIMENTAL ANALYSIS

Grad-CAM Integration

Grad-CAM visualizations are created for each classification task by calculating the gradients of the scores for the target class relative to the last convolutional layer. These heat maps are then superimposed on the input images to help interpret the model's predictions.

Evaluation Metrics

Performance is assessed using classification Accuracy: To measure prediction performance. Precision, Recall, F1-Score: For a detailed analysis of class-wise performance.

Grad-CAM Evaluation: Qualitative assessment of visualization relevance.

Results

The experimental results are discussed in terms of Classification Performance, Interpretability wit Grad-CAM and comparative analysis.

Fig. 1. Input Image peppers [13]

Classification Performance

Pre-trained CNNs achieved significant improvement over baseline models:

GoogLenet was used as pre-trained network for the experimental work. The figure 1 show the input image and figure 2 shows the Grad-CAM heat map output.

Interpretability with Grad-CAM

Grad-CAM heat maps effectively highlighted class-relevant regions, enhancing interpretability. For instance, in misclassified samples, visualizations revealed overlapping features causing ambiguity, aiding debugging.

Comparative Analysis

The figure 3 shows the graph for top five predictions.

It is clear from the graph probability for peppers is more than 95.50%.

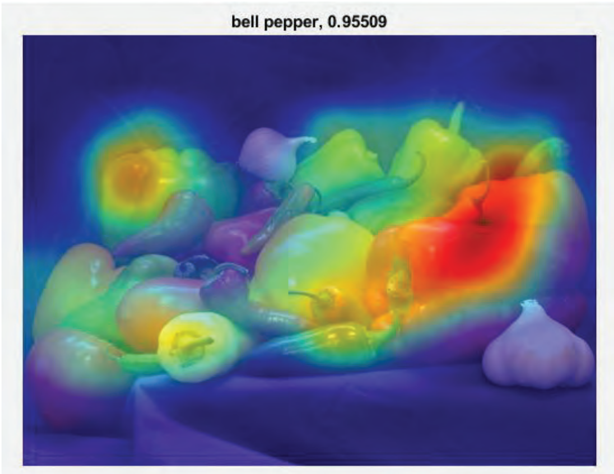


Fig. 2. GRAD-CAM heat map output image

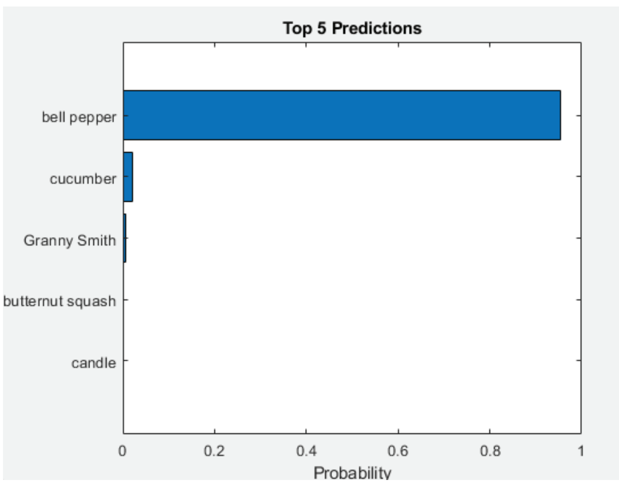


Fig. 3. Top five Predictions

The proposed approach outperformed traditional training methods in both accuracy and explainability. Grad-CAM visualizations provided actionable insights, demonstrating the model's focus on salient regions.

DISCUSSIONS

The use of pre-trained CNNs accelerates the training process and improves performance by leveraging learned features from extensive datasets. Grad-CAM serves as a valuable tool for visualizing model attention, aiding in the identification of potential biases and ensuring that the model focuses on relevant aspects of the input data. This combination enhances both the efficacy and transparency of image classification models.

CONCLUSION

This research validates a successful approach to image classification by combining pre-trained Convolutional Neural Networks (CNNs) with Gradient-weighted Class Activation Mapping (Grad-CAM). This integration utilizes transfer learning to expedite the training process and enhance model performance, while Grad-CAM provides crucial insights into the model's decision-making process, thereby improving transparency and trust. Recognizing the importance of explainability in real-world AI applications, this study addresses a significant challenge in deploying reliable and trustworthy AI systems. Future research directions include refining Grad-CAM to generate more precise and nuanced explanations and exploring the applicability of this combined approach to other data modalities, such as video and text.

REFERENCES

1. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 618–626. DOI: 10.1109/ICCV.2017.74

2. Chattopadhyay, A., Sarkar, A., Howlader, P., & Balasubramanian, V. N. (2018). Grad-CAM++: Improved Visual Explanations for Deep Convolutional Networks. Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV), 839–847. DOI: 10.1109/WACV.2018.00123

3. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. DOI: 10.1109/CVPR.2016.90

4. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2818–2826. DOI: 10.1109/CVPR.2016.308

5. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2021). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. Proceedings of the International Conference on Learning Representations (ICLR). Available at: https://arxiv.org/abs/2010.11929

6. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning Deep Features for Discriminative

Localization. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2921–2929. DOI: 10.1109/CVPR.2016.319

7. Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556. Available at: https://arxiv.org/abs/1409.1556

8. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision (IJCV), 115(3), 211–252.DOI: 10.1007/s11263-015-0816-y

9. Silva, R. S. R., & Bird, J. J. (2023). FM-G-CAM: A Holistic Approach for Explainable AI in Computer Vision. arXiv preprint arXiv:2312.05975. Available at: https://arxiv.org/abs/2312.05975

10. J. Donahue et al., "Decaf: A deep convolutional activation feature for generic visual recognition," unpublished paper, 2013. [Online]. Available: http://arxiv.org/abs/1310.1531.

11. Zeiler, M. D., & Fergus, R. (2014). Visualizing and Understanding Convolutional Networks. Proceedings of the European Conference on Computer Vision (ECCV), 818–833. DOI: 10.1007/978-3-319-10590-1_53

12. Samek, W. "Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models." arXiv preprint arXiv:1708.08296 (2017).

13. https://www.mathworks.com/help/deeplearning/ug/gradcam-explains-why.html

14. Yosinski J, Clune J, Bengio Y, and Lipson H. How transferable are features in deep neural networks? In Advances in Neural Information Processing Systems 27 (NIPS '14), NIPS Foundation, 2014.

15. Swathi, Y., and Manoj Challa. "A Comparative Analysis of Explainable AI Techniques for Enhanced Model Interpretability." 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN). IEEE, 2023.

Taming the Quantum Beast: Milestones in Quantum Error Correction

Bhagyashree Gaikwad
Asst. Prof
Dept. of Applied Sciences
Atharva College of Engineering
Mumbai, Maharashtra
✉ bhagyashreeg.bg56@gmail.com

Alice Vaghasia
Student
Atharva College of Engineering
Mumbai, Maharashtra
✉ vaghasiaalice@gmail.com

Pravin Nemade
Director and Asst. Prof
Dept. of Applied Sciences
Atharva College of Engineering
Mumbai, Maharashtra
✉ drpnnemade@yahoo.co.in

ABSTRACT

Quantum Error Correction is the most important component in the realization of quantum computing. This article presents a chronological overview of key developments in QEC and quantum computing, tracing milestones from foundational theories in the 1990s to transformative advancements in the 2020s, culminating in groundbreaking progress in 2024. It highlights both theoretical breakthroughs and experimental validations, emphasizing QEC's role in mitigating the fragility of quantum information. Finally, the paper outlines future directions toward achieving fault-tolerant quantum computation.

INTRODUCTION

Quantum computing possesses potential for solving problems that are beyond the reach of classical systems. Although classical computers have been instrumental in driving advancements in science, technology, and daily life, they are inherently limited when it comes to solving certain problems. Problems like factorisation of large numbers, simulating quantum systems, and optimization of complex processes often exceed their processing capabilities. These problems often scale exponentially with system size, rendering them impractical for classical computation. Quantum computing, on the other hand, leverages unique phenomena such as superposition and entanglement to offer exponential speedups in specific tasks, potentially revolutionizing fields like cryptography, materials science, and artificial intelligence [1]. Superposition gives power to the qubits to exist in multiple states at the same time, comparatively increasing computational capacity, and entanglement allows strong correlations between qubits, facilitating faster information processing. For example, the Shor's algorithm enables the quantum computers for factorizing large numbers intractable time, a feat that classical computers can only achieve in factorial time. This capability has significant implications for cryptographic protocols, many of which rely on the difficulty of factoring large numbers[2]. Similarly, Grover's algorithm allows quantum computers to search unsorted databases in $O(\sqrt{N})$ time, providing a quadratic speedup compared to the $O(N)$ time required by classical computers[3]. However, quantum computing's potential is matched by the formidable challenges posed by the fragility of quantum information. Qubits, which are known as the basic units of quantum information, are extremely prone to errors from noise in the environment, decoherence, and imperfections in the operations. Unlike traditional bits, which benefit from well-established error correction methods, quantum systems require entirely new approaches to protect their information. Quantum error correction (QEC) has emerged as a critical tool in addressing this issue, facilitating the implementation of quantum computing with error tolerance.

The evolution of quantum error correction has been marked by groundbreaking theoretical frameworks and experimental breakthroughs. From the initial formulation of error correction codes like the Shor Code and Steane Code [4],[5] to the development of fault-tolerant architectures and topological codes [6] , QEC has fundamentally reshaped the trajectory of quantum computing research.

This paper explores the necessity of quantum computing over classical systems, explains how quantum phenomena contribute to computational advantages, discusses the challenges of quantum measurement, and outlines common error types. Furthermore, we review QEC techniques, including the bit-flip and phase-flip codes and Shor's code, followed by an exploration of Shor code implementations in various quantum processors. This paper seeks to provide a thorough overview of how quantum computing is overcoming the intrinsic challenges posed by quantum systems, thus bringing us closer to realizing its complete potential.

CLASSICAL VERSUS QUANTUM COMPUTING

Differences between Classical and Quantum Computation

Classical computers have advanced AI, cybersecurity, and high-performance computing but face limitations due to binary logic and sequential processing, hindering complex problems in cryptography, optimization, and material science.

Quantum computers overcome these constraints by leveraging superposition, entanglement, and parallelism, enabling exponentially faster computations and unlocking new possibilities beyond classical systems.

A comparison table between classical and quantum computers in given below :

Classical Computers	Quantum Computers
Use bits (0 or 1) for data representation.	Use qubits, which can exist in superposition (0 and 1 simultaneously).
Perform computations sequentially and follow classical logic.	Leverage quantum parallelism through superposition and entanglement, allowing faster problem-solving.
Efficient for general-purpose tasks like databases, AI, and arithmetic operations.	Excel at solving complex problems in cryptography, optimization, and material science.
Stable and less prone to errors.	Highly sensitive to noise and decoherence, requiring advanced error correction techniques.
Well-developed technology with widespread commercial use.	Still in the experimental stage but shows promise for revolutionary advancements.

Classical computers remain essential for daily tasks, while quantum computers promise breakthroughs in cryptography, material science, and complex simulations. However, challenges like error correction and hardware stability mean quantum computing will complement, not replace, classical systems by tackling problems beyond their reach.

Limitations of Classical Computing

Classical computers can only process the information using binary bits (0 or 1), either sequentially or in parallel, but struggle with large-scale problems due to:

Exponential Complexity – Tasks like protein folding, financial modeling, and quantum chemistry require exponentially growing resources, making them computationally infeasible.

Slowing Moore's Law – Transistor miniaturization is approaching physical limits, restricting performance improvements.

High Energy Consumption – Supercomputers demand vast energy, limiting their scalability and sustainability for future applications.

Quantum computing offers a potential solution to these challenges by leveraging quantum parallelism and entanglement for superior efficiency.

How Quantum Computers Overcome These Limitations

Quantum computers make use of qubits, which are capable of existing in several states simultaneously because of superposition, allowing for parallel processing. Additionally, entanglement allows qubits to correlate in ways classical bits cannot, significantly enhancing computational efficiency.

Quantum Superposition in Computing: A Paradigm Shift

Quantum superposition allows quantum computers to process vast information simultaneously. Unlike classical bits (0 or 1), qubits exist in both states until measured, offering exponential computational advantages.

Real-World Analogy: The Coin Example

A classical computer is like a coin landing on heads (0) or tails (1), while a quantum computer resembles a spinning coin existing in both states until observed. This enables parallel evaluation of multiple possibilities.

Application in Problem Solving

For instance, searching a name in a million-entry database would take a classical computer 500,000 steps on average. A quantum computer, using superposition and Grover's Algorithm, finds it in about 1,000 steps—a square-root speedup [7].

Computational Power & Experimental Evidence

Superposition drives quantum advantage by enabling quantum parallelism. Algorithms like Shor's Algorithm exploit this to efficiently factor large numbers, a task infeasible for classical systems [8].

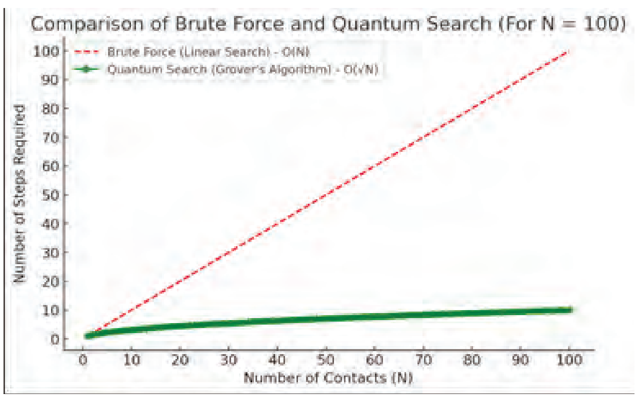


Fig. 1 : Graph Explanation: Brute Force vs. Quantum Search

-The graph compares brute force search (O(N)) with quantum search using Grover's Algorithm (O(√N)).

Brute Force (Red Line): Searches each contact one by one, resulting in a linear growth in steps.

Quantum Search (Green Line): Leverages Grover's Algorithm, reducing search steps significantly.

For 100 contacts:

Brute force requires 100 steps.

Quantum search requires only 10 steps (√100).

This demonstrates that quantum search is exponentially faster, making it ideal for large-scale search tasks like databases and cryptography.

Classical Computer Approach (Step-by-Step Search)

A classical computer can use different search methods

Linear Search (Brute Force)

The computer checks each contact one by one.

In the worst case, it has to check all 1,000 names.

Number of iterations (worst case): 1,000 steps.

Average case: 500 steps (half of 1,000).

Example:

Imagine you are looking through a printed phone book page by page, checking each name one by one.

Quantum Computer Approach (Grover's Algorithm)

Quantum computers use superposition and interference to search faster. Instead of checking names one by one, they use Grover's Algorithm, which processes all names simultaneously in a quantum state.

The quantum computer puts all 1,000 contacts into superposition.

It applies a special quantum trick called amplitude amplification, increasing the chances of the correct name appearing.

After about √N (square root of N) steps, the quantum computer finds the correct name.

Number of iterations (worst case): √1,000 ≈ 32 steps.

Faster than linear search, but slightly slower than binary search if the contacts are sorted.

Example:

Instead of flipping through the phone book, imagine having a magic phone that instantly highlights the correct contact after just 32 tries!

Quantum Entanglement: A Gateway to Advanced Computing and Communication

Quantum entanglement is an important theory in quantum mechanics, where two or more qubits are interconnected and linked together regardless of the distance between them. If two qubits are entangled, this means that the state of one of them will instantly affect the other qubit's state, no matter how far apart they might be. This unique property allows for revolutionary applications in quantum computing and its various applications.

Role in Quantum Computing

Entanglement enables quantum parallelism, allowing qubits to process multiple possibilities simultaneously. Unlike classical bits, which are independent, entangled qubits share correlated information, vastly expanding computational capabilities. A few of the many algorithms like the Shor's algorithm (for prime factorization) and the Grover's algorithm benefit from entanglement, significantly enhancing computational speed. Bhuyan and Dutta (2024) highlight its role in secure communication and error correction for scalable quantum networks [9].

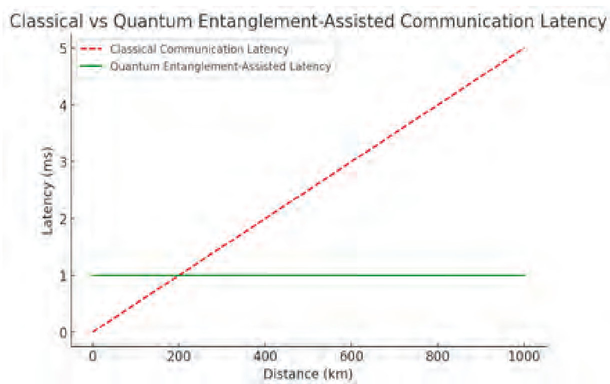


Fig 2 : Graph Explanation: Classical vs. Quantum Entanglement-Assisted Communication

The x-axis in the graph represents the distance (in km).And the y-axis here represents the latency (in milliseconds).

Key Observations: ● Classical Communication (Red Dashed Line)Increases with distance due to the physical limits of light speed in fiber optics (~200 km/ms).Longer distances lead to higher delays.□ Quantum Entanglement-Assisted Communication (Green Line)Remains almost constant, regardless of distance.

This is because entanglement enables instantaneous correlations between particles, eliminating the need for classical signal transmission over large distances.

Impact on Quantum Computing

Entanglement accelerates computing in various ways:

Faster Communication – Enables near-instantaneous quantum information transfer.

Quantum Teleportation – Allows quantum state transfer between distant locations.

Quantum Cryptography – Ensures ultra-secure communication by detecting eavesdropping.

Parallel Processing – Entangled qubits collaborate, reducing computational complexity.

By leveraging entanglement, quantum computers achieve superior speed and efficiency, unlocking potential in AI, security, and complex problem-solving.

Real-World Example: Quantum Key Distribution (QKD)

Entanglement is crucial in Quantum Key Distribution protocolls–s, for example China's Quantum Communication Satellite (Micius) which uses the BB84 protocol, ensuring secure cryptographic key exchange. Any interception collapses entanglement, alerting the parties, making QKD foundational for the future Quantum Internet [9].

Entanglement's Advantage Over Classical Systems

Entanglement allows instant state transmission, a feat impossible in classical communication. It also strengthens quantum error correction by encoding redundant states to combat decoherence [10].

Real-World Analogies

Twin Connection Analogy – Like twins mysteriously sharing emotions across distances, entangled qubits instantly influence each other upon measurement.

Entanglement is not just theoretical—it drives breakthroughs in computing, cryptography, and secure communication. As research advances, its impact on quantum technologies will continue to expand.

THE CHALLENGE IN MEASUREMENT IN QUANTUM COMPUTING

The measurement problem in quantum computing arises due to the fundamental principles of quantum mechanics, where the act of measurement collapses a quantum state into a definite outcome. This challenge affects both the accuracy and efficiency of quantum computations, as well as the interpretations of quantum mechanics itself.

The Nature of Quantum Measurement

Unlike classical computers, where data exists in well-defined states (0 or 1), quantum computers utilize quantum superposition, allowing qubits to exist in two different states concurrently. But there's a catch, when a measurement is performed, the superposition of the qubit collapses into one of the states, losing all other possible information. This irreversible collapse introduces fundamental difficulties in quantum computation, particularly in extracting meaningful results without disturbing the quantum system excessively [11][12].

Decoherence and the Loss of Quantum Information

A major challenge in quantum computing is decoherence, where external interactions cause a quantum system to lose coherence and behave classically.Decoherence disrupts computations before completion, making error correction crucial for practical quantum computing.While decoherence theory explains why quantum states appear classical after measurement, it does not fully address the measurement problem, as quantum mechanics lacks a clear mechanism for wave function collapse [12].

Measurement-Based Quantum Computation (MBQC)

Measurement-based quantum computation provides a potential solution to the measurement problem, where entanglement is used as a resource, and computations proceed via sequential measurements. Introduced by Raussendorf and Briegel, this model utilizes local measurements on an entangled state to guide computations. The randomness of outcomes is corrected through adaptive adjustments, ensuring deterministic results. MBQC has been experimentally tested in various systems, offering an alternative to circuit-based quantum computing [11].

TYPES OF QUANTUM ERRORS

Quantum errors arise in quantum computing due to imperfections in qubit manipulations and interactions with the environment. These errors can occur in different forms, namely bit-flip errors, phase-flip errors, and mixed errors. Each type affects quantum information differently, and understanding these is essential for the development of fault-tolerant quantum computing techniques.

Bit - Flip Error

This type of error occurs when any qubit, intended to be in state $|1\rangle$, or vice versa flips into the other state. In classical computing, this would be analogous to the flipping of a binary bit. However, in quantum computing, such an error disrupts the quantum superposition, which is fundamental to quantum parallelism. This type of error is typically represented by the Pauli-X operator. Bit-flip errors are particularly problematic in quantum algorithms like Grover’s search algorithm, where the accuracy of qubit states is crucial for success [13].

Phase-Flip Error

Unlike bit-flip errors, phase-flip errors affect the phase of the qubit’s quantum state without altering the bit value. A phase-flip error changes the relative phase between the $|1\rangle$ states, which can lead to wrong measurement outcomes. Mathematically, the Pauli-Z operator models this type of error. Phase-flips are often harder to detect because they don’t alter the bit value but can significantly affect quantum interference patterns in algorithms like the Quantum Fourier Transform [14].

Mixed Error

A mixed error involves both bit-flip and phase-flip errors occurring simultaneously. This type of error is particularly challenging to correct because it affects both the computational basis and the relative phases between quantum states. The combination of Pauli-X and Pauli-Z operators describes mixed errors. Mixed errors are common in real-world quantum systems where both types of noise occur simultaneously due to imperfections in quantum gates and environmental interactions [15].

Error Correction in Quantum Systems

Quantum error correction techniques, like the surface codes and the Shor’s code, intend to conserve quantum information by using redundancy. This technique encodes the quantum information of one qubit into multiple physical qubits to identify and solve the errors without needing to measure the qubits directly. However, error correction comes with significant overhead, requiring additional qubits and resources. As quantum systems scale up, effective QEC will be vital for ensuring reliable computations. Research into error-correcting codes and noise-tolerant quantum architectures remains a major area of development [16].

Bit-flip, phase-flip, and mixed errors are fundamental obstacles in the field of quantum computing. Addressing these errors through these advanced techniques is very important for attaining functional quantum computation in the coming years. Research is ongoing to develop more efficient and scalable solutions for mitigating the effects of these errors in large quantum systems.

IMPORTANCE OF QUANTUM ERROR CORRECTION (QEC)

As quantum error correction (QEC) advances, fully fault-tolerant quantum computers are becoming increasingly feasible. The future of QEC will focus on refining error correction codes, integrating AI for adaptive correction, and expanding interdisciplinary applications to harness quantum computing's full potential [17].

A key priority is improving error correction codes to enhance efficiency and scalability for large quantum systems. As qubit numbers grow, more robust and resource-efficient codes are needed to minimize overhead. Surface codes, concatenated codes, and topological codes will likely evolve, with hybrid approaches combining their strengths for better scalability [18].

AI and machine learning offer promising advancements in adaptive QEC. Traditional methods rely on fixed protocols, but quantum devices' unique noise profiles require dynamic, real-time adjustments. AI-driven systems can predict and adapt to noise patterns, optimizing error correction on the fly. Google's Alpha Qubit in 2024 showcased AI-enhanced QEC, demonstrating how machine learning could revolutionize error correction strategies [19].

QEC also has far-reaching interdisciplinary applications. In quantum cryptography, fault-tolerant quantum computers could enhance encryption methods like quantum key distribution (QKD). In materials science, QEC-enabled quantum systems can simulate complex materials for clean energy, drug discovery, and advanced quantum technologies. Additionally, QEC will play a crucial role in quantum machine learning, improving optimization and data analysis for applications such as natural language processing and drug development [20].

The realization of fault-tolerant quantum systems will revolutionize various fields, enabling precise simulations of quantum systems and molecular structures. This will drive breakthroughs in materials science, pharmaceuticals, and energy-efficient technologies. Reliable quantum systems will also enhance logistics, optimization, and AI, surpassing classical systems in many domains [21].

As QEC continues to evolve, it will play a crucial role in harnessing the complete capabilities of quantum computing. The coming decades will bring groundbreaking innovations, making fault-tolerant quantum computers a transformative tool in cryptography, materials science, AI, and beyond [22].

QUANTUM ERROR CORRECTION TECHNIQUES

Quantum error correction (QEC) has evolved from a theoretical concept to a vital component of quantum computing. Early work by researchers like Peter Shor and Andrew Steane in the 1990s laid the groundwork for QEC, showing how quantum information could be protected from errors. The 1998 demonstration using nuclear magnetic resonance (NMR) and later developments like surface codes marked key milestones in turning theory into practice [23].

By the 2010s, scalable QEC techniques, particularly surface codes, provided practical solutions for large quantum systems. Advancements like Google's Willow Processor and Microsoft-Quantinuum collaborations highlighted the feasibility of fault-tolerant quantum computing by 2024 [24].

The future of QEC promises significant breakthroughs, driven by advancements in error correction codes, AI, and quantum hardware. These developments will not only revolutionize quantum cryptography, materials science, and AI, but also enable quantum systems to solve problems beyond classical computing capabilities [25].

QEC will continue to be central to achieving fault-tolerant quantum systems, making previously unsolvable problems tractable. As the field matures, QEC's interdisciplinary impact will drive scientific progress across multiple domains [26].

RESULTS AND CONCLUSION

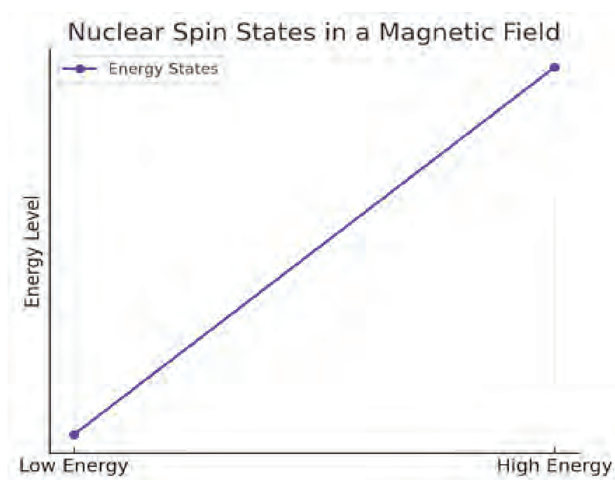


Fig. 3. Nuclear Spin States in a Magnetic Field

When an atomic nucleus is placed in a strong magnetic field, its spin aligns in one of two possible states:Low-energy state (aligned with the field)High-energy state (aligned against the field)

This behavior is seen in a travelling compass where the needle is always pointing north in correspondence to the Earth's magnetic field. Here the energy difference between these two states is the basis for Nuclear Magnetic Resonance (NMR) and quantum computing.

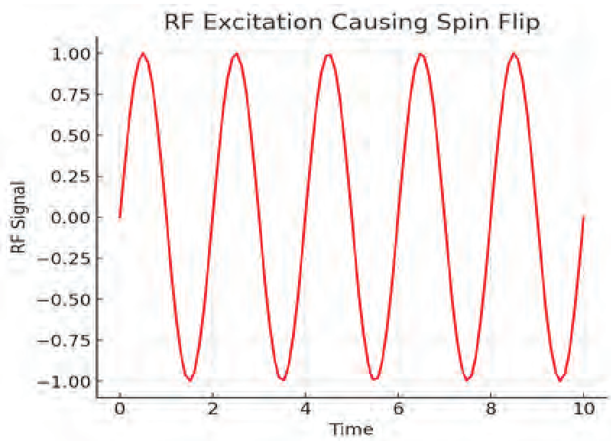


Fig 4. RF Excitation Causing Spin Flip

This graph represents how an external radio frequency (RF) wave affects nuclear spins:

- The red wave shows the oscillating RF signal applied to the system.
- When the frequency of the RF wave matches the resonance frequency of the nucleus, it excites the nuclear spin, flipping it from the low-energy state to the high-energy state.
- This process is essential for both NMR spectroscopy and quantum computing, where controlled spin flips encode quantum information.

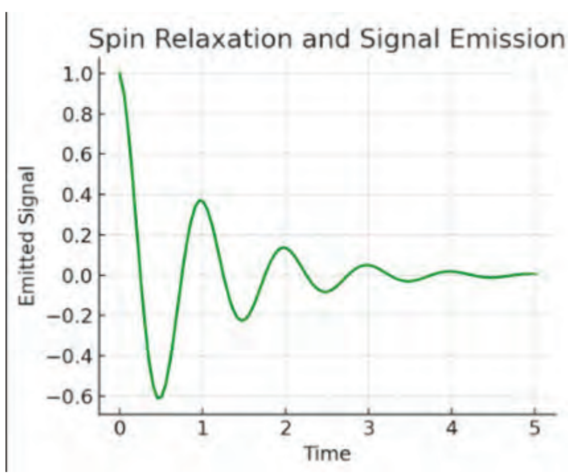


Fig 5. Spin Relaxation and Signal Emission

Shows how nuclear spins return to a lower-energy state after excitation in NMR-based quantum computing.X-axis: Time (s) | Y-axis: Signal IntensityThe signal decays exponentially due to spin relaxation, with oscillations from spin precession.The emitted signal helps determine qubit states, enabling quantum computations.

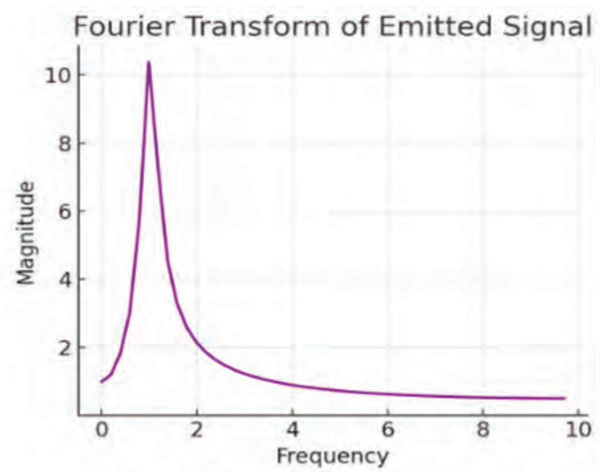


Fig 6. Quantum State Measurement and Decoherence

Illustrates the loss of quantum coherence over time due to decoherence.X-axis: Time (s) | Y-axis: CoherenceCoherence decreases as quantum states interact with the environment, causing errors.Longer coherence times are crucial for practical quantum computing and error correction.

REFERENCES

1. P. W. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," SIAM Journal on Computing, vol. 26, no. 5, pp. 1484–1509, Oct. 1997.
2. L. K. Grover, "A fast quantum mechanical algorithm for database search," in Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC), Philadelphia, PA, USA, 1996, pp. 212–219.
3. P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," Physical Review A, vol. 52, no. 4, pp. R2493–R2496, Oct. 1995.
4. A. M. Steane, "Error Correcting Codes in Quantum Theory," Physical Review Letters, vol. 77, no. 5, pp. 793–797, Jul. 1996.
5. E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, "Topological quantum memory," Journal of Mathematical Physics, vol. 43, no. 9, pp. 4452–4505, Sep. 2002

6. D. Gottesman, "An Introduction to Quantum Error Correction and Fault-Tolerant Quantum Computation," in Quantum Information Science and Its Contributions to Mathematics, Proceedings of Symposia in Applied Mathematics, vol. 68, 2010, pp. 13–58.

7. Preskill, J., "Quantum Computing and the Entanglement Frontier," arXiv:1203.5813 (2012).

8. Silverman, M.P., "Quantum Superposition: Counterintuitive Consequences of Coherence," Springer, 2015.

9. Bhuyan, A., & Dutta, H., Harnessing Quantum Entanglement: Comprehensive Strategies for Enhanced Communication and Beyond in Quantum Networks, 2024, arXiv:2406.08833.

10. John Preskill, Quantum Computing and the Entanglement Frontier, 2012, arXiv:1203.5813

11. S. Xiao, S. Xue, D. Dong, and J. Zhang, "Identification of Time-Varying Decoherence Rates for Open Quantum Systems," IEEE Transactions on Quantum Engineering, vol. 3, 2021.

12. J. Kim, H. S. Moon, and K. Kim, "Experimental demonstration of decoherence suppression by quantum measurement reversal," IEEE Conference Publication, 2011.

13. Shor, P.W., "Fault-tolerant quantum computation," 1996, SIAM Journal on Computing, Volume 26, Issue 5.

14. Gottesman, D., "Theory of fault-tolerant quantum computation," 1998, Physical Review A, Volume 57, Issue 1.

15. Knill, E., "Quantum computing with errors," 2005, Nature, Volume 434, Issue 7029.

16. Terhal, B.M., "Quantum error correction for beginners," 2015, Reviews in Physics, Volume 1, Issue 1.

17. Devitt, S.J., & Fowler, A.G., "Quantum error correction and its role in fault-tolerant quantum computing," 2022, Journal of Quantum Information Science, Volume 10, Issue 2.1

18. Bravyi, S., & Kitaev, A., "Quantum codes on a lattice," 1998, Physical Review A, Volume 57, Issue 3.

19. Google Quantum AI, "Alpha Qubit: AI-driven Quantum Error Correction," 2024, Nature Communications, Volume 15, Issue 5.

20. Dunjko, V., & Briegel, H.J., "Quantum machine learning: A review," 2018, Physics Reports, Volume 764, Issue 1.

21. Farhi, E., & Gutmann, S., "Quantum computation and quantum state compression," 2003, Physical Review A, Volume 67, Issue 6.

22. Preskill, J., "Quantum computing in the NISQ era and beyond," 2025, Quantum, Volume 4, Issue 2.

23. Shor, P.W., & Steane, A., "Quantum error correction codes: A historical perspective," 2000, Quantum Information & Computation, Volume 2, Issue 4.

24. Google Quantum AI, "Willow Processor and the future of AI-enhanced QEC," 2024, Nature Quantum Computing, Volume 3, Issue 7.

25. Preskill, J., "Quantum computing and cryptography: The role of QEC," 2025, Reviews in Physics, Volume 2, Issue 6.

26. Devitt, S.J., & Fowler, A.G., "The evolving role of quantum error correction in scalable quantum systems," 2025, Quantum Journal, Volume 8, Issue 5.

Green Corrosion Inhibition of Mild Steel: Insights from Plant Extracts

Vanshika Singh
Assistant Professor
Humanities and Applied Sciences Department
Atharva College of Engineering
Mumbai, Maharashtra
✉ dr.vanshikasingh@atharvacoe.ac.in

Bhushan Sonawane
Assistant Professor
Humanities and Applied Sciences Department
Atharva College of Engineering
Mumbai, Maharashtra
✉ bhushansonawane@atharvacoe.ac.in

Pravin Nemade
Professor, Director
Humanities and Applied Sciences Department
Atharva College of Engineering
Mumbai, Maharashtra
✉ drpnnemade@yahoo.co.in

Priyanka Badani
Assistant Professor
Humanities and Applied Sciences Department
Atharva College of Engineering
Mumbai, Maharashtra
✉ priyankapatel@atharvacoe.ac.in

ABSTRACT

Inhibition of corrosion in sulphuric acid and hydrochloric acid of mild steel were studied in absence and presence of plant extracts as a pleasant inhibitor. The outcome acquired that leaves extracts reduces the corrosion rate and serve as an inhibitor for metal (steel) in acid. As the concentration of inhibitor increases; the efficacy of inhibition was found to be increase whereas the temperature shows otherwise. The source of plants, the concentration of the primary active components in plants, and the method of preparation all affect how well plant corrosion inhibitors stop corrosion. The inhibitor is a mixed type showed by polarization measurements. The outcomes acquired from the various procedures are in good settlement. With the help of Scanning electron microscope as well as Fourier-transform infrared FTIR the morphologies of surface of mild steel have been observed.

KEYWORDS *: Corrosion, Inhibitor, Polarization measurements.*

INTRODUCTION

Corrosion constituents are one of the foremost issues in the trade that is related with serious economic losses. Many of organic compounds were analyzed as well as analyzed to study their potential of corrosion inhibition. All these investigations tell that organic compound specifically those with Nitrogen, Sulphur and Oxygen exhibited vital inhibition efficacy. But, unhappily most of these complexes are not only noxious to human beings, but expensive too. It is unneeded to pin point the significance of low-priced, safe inhibitors of corrosion.

Solutions of acid are normally used for the elimination of corrode in various industrial activity and an adverse scale. The extensively used acids are HCl and H2SO4 in the preserving methods of metals. The utmost practical approaches for protection against corrosion is usage of inhibitors mainly in acid solutions to avoid termination of metal and ingestion of acid. The usage of organic compounds comprising nitrogen, oxygen and Sulphur to diminish attack of corrosion on steel has been investigated in few details. The more effective are nitrogen and Sulphur containing compounds as an inhibitor of corrosion in H2SO4 and HCl [1]. The present data indicates that maximum organic inhibitors get adsorbed on surface of metal through relocating H2O molecules and set up a dense blockade film. Availability of л electrons and lone pairs in prohibitory molecules simplify electron imparting from the inhibitor to metal, making a coordinate covalent bond. The chemisorption bond's strength depends on electron density on the donor atom of the functional group as well as also polarizability of the group.

ADVERSATIVE EFFECT OF CORROSION

Corrosion of alloys and metals bring about massive commercial affect to metal-based structure works and is a possible danger to natural life. It can be described as an unprompted corrosion of metals and alloys upon an antagonistic onslaught of nearby surroundings. Many industries like petroleum, energy, gas and oil, heat trading, boilers, steel plant, etc. make usage of metals like copper, steels, zinc, aliminium etc. The surface of metal comes about into straight contact with strong acidic, alkaline, neutral atmospheres through diverse application purposes, which can give on to serious destruction of these surfaces because of corrosive attack. It's about 3% to 6% of world's GDP annual economic loss due to corrosion. National Association of Corrosion Engineers has brought out an analysis in a number of developing as well as developed nations and approximated the yearly global loss of corrosion to be about five percent of the total gross national product [2]. Apart from this financial declination, many misfortunes have been reported that occurred because of corrosion failure of metallic structures for examples buildings, bridges, as well as overpasses, etc. in which numerous demises are present. Moreover, corrosion badly disturbs the nearby environment through the contamination of corrosion products into the environment. The ruined of these stuffs badly influences the soil and aquatic life. So, the dealings for corrosion reduction are also likely to be examined [3].

IMPORTANCE OF GREEN INHIBITORS

Corrosion weakens the metal gradually by act of atmospheric gases, humidity and other harmful substances. To overcome that eradication of metal, extract of various plants can be used as corrosion inhibitors, which is commonly known as green corrosion inhibitors. Naturally occurring compounds comes from plants [4].

Naturally occurring compounds are ecologically suitable that is why they are widely used cost and active and have plentiful accessibility. The biological, chemical, and physical characteristics of any complex molecular structure include variations. To stop the interaction between the metal and the acidic medium, inhibitors are added to corrosive media in small amounts.

PLANTS USED AS NATURAL INHIBITORS

A. Lawsonia- The extract of leaves in aqueous of this plant suitably work as corrosion inhibitor which was stated in Carbon steel, Ni and Zn in alkaline, acidic, and neutral solutions, polarization method, IE= 94%.

B. Aleo vera- Aloe vera plant is from the family of Liliaceae. The ingredients of gel are, vitamins, glycoproteins, mineral, polysaccharides and enzymes, inhibition efficiency was decline with increase in temperature and rise with enhancing concentration of extract, 92%.

C. Cotula cinerea- Plant extracts were examined on the metal in aq. 20% (1 M) H2SO4. Mass loss method and electrochemical measurements were also executed, IE=67%.

D. Fenugreek- The impact of temperature were examined on metal corrosion in 1 M of HCl and H2SO4 without and with aqueous extract of plant leaves by the support of gravimetric approach, IE= 91.2%.

MATERIALS AND METHODS

Specimens Preparation

The mild steel specimen composition can be of minute percentage Carbon, Silicon, Phosphorus, and rest iron. Mechanically samples should be cut into small coupons and grazed by abrasive paper of distinct grades and ultimately furbished with abrasive paper to acquire mirror-like finish. Each specimen has to follow some steps i.e. grease should be removed by washing with CH3COCH3, should be dried up at room temperature and well-preserved in damp free desiccators.

Plant extract preparation

Dried the plant and crush it then Soxhlet will be used for extraction of plant. The solution was sieved and evaporated. This extraction was prepared to examine the properties of inhibition of corrosion.

WEIGHT LOSS METHOD

Experiments of weight loss method were completed at several ranges of temperatures for two hours in acid. The specimens were dipped in the particular inhibitor and assessment solution in a thermo-stated bath. Before and after immersion the specimens were weighed. The change in weight was taking as the weight loss of metal.

In triplicate, four-sided metal samples were dipped in electrolyte at particular temperatures with and without adding of different amounts of plant extract. Each experiment was conducted in triplicate, and the mean mass loss value is reported. The following formula was used to determine the surface coverage (θ) and inhibitory efficiency (%):

$$\theta = \frac{w_0 - w_i}{w_0} \tag{1}$$

$$\eta\,(\%) = \frac{w_0 - w_i}{w_0} \times 100 \tag{2}$$

where

W0 = loss in weight of metal without inhibitor

Wi = loss in weight of metal with inhibitor.

Using this relationship, the corrosion rate (CR) of mild steel was intended:

$$\text{corrosion rate} = \frac{87.6 \times w}{AtD} \tag{3}$$

Where,

Weight loss of sample (mg) = w

Density of sample (g cm^{-3}) = D

Area of sample (cm^2) = A

Time of exposure (hours) = t

During a three-hour absorption period, weight loss measurements were made at particular temperatures range with and without the extract at an ideal concentration in order to estimate the plant extract adsorption and the initial factors of the corrosion method of steel in acidic media. The inhibition efficiency decreases as the temperature rises at an ideal inhibitor concentration. This is due to both the partial sorption of the inhibitor from the metal's surface and the increased rate of mild steel suspension.

POTENTIODYNAMIC POLARIZATION

Corrosion performance of metals, alloys and coverings in corrosive environments is determined by electrochemical method for corrosion. The approach examines the corrosion performance rapidly, as well as can correctly classify the corrosion rate.

Investigation of the potentiodynamic polarization plots shows that the adding of plant extract reduces both the densities of anodic and cathodic current with increase in plant extract concentration [5]. The presence of plant extract in violent medium causes alteration in cathodic and anodic branches with no important trend in the shift of Ecorr values in all verified concentrations, signifying that in general, the extraction acts as mixed type inhibitor with no alteration in corrosion approach happened because of the addition of inhibitor. When the content of plant extracts increases in solution, the I_{corr} decreases unconditionally. This value of I_{corr} headed to an increase in inhibition efficiency and approve that the extract is a good inhibitor against the corrosion of molecular surface in HCl medium.

MECHANISM OF INHIBITION CORROSION

Consequently, the usefulness of stopping corrosion by plant extract might be imputed to adsorption of its phytochemical constituents on surface of mild steel. The composite chemical structure of the plant extract, it is pretty tough to assign the inhibitive effect to a specific component. Subsequently some of these components including oxides, alcohols, aromatic and esters this creates it hard to set the inhibitive effect to adsorption of an individual component. The occurrence of π-electrons of the aromatic ring, heteroatom oxygen in functional groups (C=O, C–O, O–H) and the double bonds in their arrangement (C=C) of these compounds, which meet the overall features of classic corrosion inhibitors. Additionally, these organic compounds are found in plant extracts either as protonated organic molecules (cations) in aqueous acid solution or as neutral molecules [6]. Thus, the significance of an inhibitor of corrosion can be associated to its molecular electronic arrangement. Usually, two methods of adsorption are examined in acid media on surface of metal.

ANALYSES

Fourier-transform infrared analysis

This analysis has been made to acquire some assimilation into the possible interfaces between the surfaces of mild steel as well as adsorbed inhibitor in acidic environment. The inhibition strength generally depends on the molecular assembly of inhibitor. The spectrum of FTIR of mild steel dipped in inhibited solution containing natural corrosion inhibitor. Extraction of plant was examined to estimate the foremost functional group [7].

Scanning electron microscope surface analysis

The mild steel samples were carried out for morphological studies by using scanning electron microscope inspections of exposed surfaces to diverse test solutions, using SEM which is working with a tungsten filament which works as electron source. The cleaned coupons were dipped for 3 hours in acid solution in the absence of and presence of plant extract at suitable temperature, and then splashed with distilled water, dried out in warm air, and give in to for surface examination by SEM.

Electrochemical impedance analysis

At the open-circuit potential in this spectroscopy the process of electrochemical taking place which observed by electrochemical impedance spectroscopy (EIS), in direction to get data about the kinetics of iron corrosion which is present in extracts of plant. The steel electrode of electrochemical impedance spectroscopy at its open circuit potential were immersed in the solution of sulphuric and hydrochloric acid for 15 minutes and in existence of several concentrations were accomplished over from 10 kHz to 10 MHz of frequency range.

DENSITY FUNCTIONAL THEORY

DFT is the theoretical method which is one of significant methods used in describing the chemistry and solids in science. Within the frame of DFT a many chemical concepts have been associated. Electron density ρ(r) is the primitive parameter in Density functional theory in positions of which all the chemical measures are exhibit [8]. The calculation of structural factors done via the electronic density and it compared fine with factors intended by the Schrodinger equation in relations of the single-electron wave function (Ψ), where the outcome can be functional. As the theory is simpler in comparison classical quantum mechanics, which depend on resolving the many Schrodinger equations with imperfectly investigated the electron interconnection causes, attentiveness has developed in assimilating the properties, arrangement, dynamics, reactivity of atoms, particles and clusters using Density Functional Theory, and large systems including many molecules or atoms are now replicated on a consistent basis. Just, as noted above, Density Functional Theory has been used to examine the surface interaction/features of inhibitor and therefore inhibitor mechanisms and to define the consequence of the physical nature of an inhibitor on oxidation process. The persistence of this matter is only to dealing with all elements, but not in-depth, with importance on inhibition of chemical corrosion.

CONCLUSION

The utmost current reports on natural occurring compounds that have been assessed to serve as corrosion inhibitors on surface of steel or metal are briefly calculated. A number of factors, including concentration, extraction solvent, concentration, temperature, and immersion time, can be found to evaluate a naturally occurring compound's ability to inhibit corrosion. A corrosion inhibitor's effectiveness must be evaluated using a minimum of two electrochemical methods, including electrochemical impedance spectroscopy, weight loss, and potentiodynamic polarization.

Phytochemicals will be obtained from extracts of plants; in general, heterocycles are suitable for beneficial interaction with surface of the metal. Finally, natural chemicals derived from plants have also been found to contribute to the suppression of corrosion in steel. The investigation team may be inspired to obtain novel green corrosion inhibitors by this compilation. Metal structures will survive longer and corrosion on metal surfaces will be lessened by creating new, better natural inhibitors. The last essential component in stopping metal corrosion is green corrosion inhibitors.

FUTURE OUTLOOKS

Green corrosion inhibitor development is gaining traction due of worries about the effects chemical inhibitors have on the environment. These inhibitors are less harmful to the environment, biodegrade, and are derived from sources that are renewable. As possible green inhibitors, researchers are examining natural sources such plant extracts, biopolymers, and essential oils.

ACKNOWLEDGEMENT

Authors are grateful to Atharva college of engineering, India to provide necessary support for compilation of work.

REFERENCES

1. Avdeev, Y. G., & Kuznetsov, Y. I. (2021). Nitrogen-containing five-membered heterocyclic compounds as corrosion inhibitors for metals in solutions of mineral acids-An overview. International Journal of Corrosion and Scale Inhibition, 10(2), 480-540.
2. Shekari, E., Khan, F., & Ahmed, S. (2017). Economic risk analysis of pitting corrosion in process facilities. International Journal of Pressure Vessels and Piping, 157, 51-62.
3. Masroor, S. (2023). An overview of corrosion. Smart Anticorrosive Materials, 3-17.
4. Shang, Z., & Zhu, J. (2021). Overview on plant extracts as green corrosion inhibitors in the oil and gas fields. Journal of Materials Research and Technology, 15, 5078-5094.

5. Ohunene, Y. M., Joseph, O., Maia-Obi, L. P., faks Mtunzi, F., Ojo, F., & Bamidele, O. J. Corrosion Inhibition of Mild Steel With Phenol-Based Anticorrosion Fraction: Preparation, Phytochemistry, and Potentiodynamic Polarization.

6. Harvey, T. J., Walsh, F. C., & Nahlé, A. H. (2018). A review of inhibitors for the corrosion of transition metals in aqueous acids. Journal of Molecular Liquids, 266, 160-175.

7. Kokilaramani, S., Rajasekar, A., AlSalhi, M. S., & Devanesan, S. (2021). Characterization of methanolic extract of seaweeds as environmentally benign corrosion inhibitors for mild steel corrosion in sodium chloride environment. Journal of Molecular Liquids, 340, 117011.

8. Li, S., Li, C., & Wang, F. (2024). Computational experiments of metal corrosion studies: a review. Materials Today Chemistry, 37, 101986.

Enhanced Modelling and Control of Dual Active Bridge Converters: A Unified Approach

Ammu Striney J
Department of Electronics & Telecommunication Engg.
Atharva College of Engineering
Mumbai, Maharashtra
✉ ammustriney@atharvacoe.ac.in

Purvi Agarwal
Department of Electronics & Telecommunication Engg.
Atharva College of Engineering
Mumbai, Maharashtra
✉ purviagarwal@atharvacoe.ac.in

Vivek Ramakrishnan
Department of Electronics & Telecommunication Engg.
Atharva College of Engineering
Mumbai, Maharashtra
✉ vivekramakrishnan@atharvacoe.ac.in

ABSTRACT

Dual active bridge (DAB) converters are crucial for high-efficiency isolated DC-DC power conversion in applications such as electric vehicles and micro grids. This paper presents a unified framework for enhanced modelling and control of single-phase and three-phase DAB converters. A hybrid modelling approach combining state-space and generalized averaging methods is proposed to improve small-signal stability analysis. Additionally, a fast recovery control strategy with feedforward compensation effectively suppresses DC-bias currents and mitigates magnetic saturation during transients. Exact solutions for zero-voltage-switching (ZVS) boundaries and AC-port currents are derived to optimize efficiency. Simulation and experimental results validate the proposed methods, showcasing enhanced stability, efficiency, and dynamic performance.

***KEYWORDS** : Unified framework, Hybrid modelling approach, Fast recovery control strategy, Zero-voltage-switching (ZVS) Boundaries, AC-port currents.*

INTRODUCTION

Dual active bridge (DAB) converters are widely recognized for their bidirectional power transfer, high efficiency, and compact design, making them suitable for electric vehicle charging, renewable energy integration, and DC micro grids. With features like flexible control options and galvanic isolation, they are a cornerstone of modern power systems.

DAB converters operate under modulation strategies such as single-phase shift (SPS), dual-phase shift (DPS), and triple-phase shift (TPS), which impact their efficiency and performance. However, challenges persist in accurate modelling and control across varying conditions and modes. A unified modelling approach can address these challenges by streamlining design and optimizing performance.

Prior research offers critical insights. Shah et al. explored zero voltage switching (ZVS) boundaries and AC-port currents in single-phase DAB converters, deriving closed-form expressions for peak and RMS currents. While effective, their scope was confined to single-phase systems. proposed feedforward control to mitigate DC-bias currents and transformer saturation during transients, enhancing system stability advanced hybrid modelling for three-phase systems, addressing small-signal stability but with limited focus on unified control strategies demonstrated the potential of hybrid modular DAB converters for high-voltage applications by integrating resonant and non-resonant modules.

Despite these advancements, a comprehensive framework uniting advanced modelling and unified control strategies for diverse DAB configurations is yet to be fully realized. This paper aims to bridge this gap by presenting a robust modelling and control framework that adapts to both single- and multi-phase systems. Through simulations and experimental validation, the framework seeks to improve scalability and reliability in modern power systems.

The paper is organized as follows: Section 2 presents the proposed approach and its contributions. Section 3 details the modelling framework, emphasizing control and performance. Section 4 discusses results and practical implications, while Section 5 concludes with insights and future directions.

A DUAL ACTIVE BRIDGE CONVERTER

Proposed System and Assumptions

The block diagram of the proposed DAB converter is shown in Fig. 1.

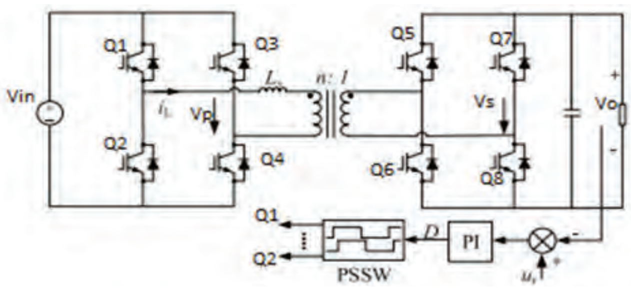


Fig. 1 Block diagram of proposed system with the modelling & Control

The block diagram comprises four main components: The Voltage Source, the Proposed system, various outputs and controller part and a driver circuit.

To address the challenges of accurate modeling and efficient control in Dual Active Bridge (DAB) converters across varying operating conditions, this paper introduces a unified framework that integrates advanced modeling methodologies with robust control strategies. This approach is designed to improve scalability, adaptability, and reliability in both single-phase and multi-phase DAB systems.

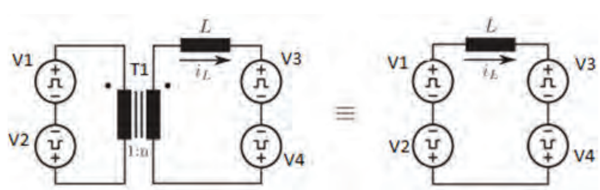


Fig. 2 Circuit diagram of DAB converter

OPERATING PRINCIPLES

The unified framework integrates the strengths of analytical and simulation-based methods for accurate and efficient modeling. It emphasizes scalability and standardization to streamline the design of both single-phase and multi-phase DAB converters. The framework's core objectives include:

Comprehensive Modeling

Encompassing system-level dynamics, loss components, and transient effects for enhanced precision.

Advanced Control: Implementing fast-response strategies to manage performance under dynamic conditions while maintaining system stability.

Design Simplification: Creating a platform for easily extending insights across various topologies and applications

Hybrid Modeling Approach

The proposed hybrid modeling approach combines mathematical derivation and state-space averaging with behavioral simulations for accuracy and efficiency.

State-Space Averaging for System Dynamics:

A general state-space representation is developed to capture the nonlinear behavior of DAB converters across switching cycles. This allows for seamless analysis of steady-state and transient operations, enabling rapid design evaluations.

i. Time-Domain and Frequency-Domain Representation:

Leveraging time-domain equations and frequency-domain insights, the model captures high-frequency switching phenomena, core losses, and transformer dynamics.

ii. Multi-Phase Extension via Modular Representation:

By adopting a modular approach, the framework extends to multi-phase configurations. Phase interactions are captured through coupling coefficients, ensuring scalability for high-power systems.

iii. Integration of Loss Models:

The model includes detailed quantifications of conduction and switching losses, incorporating zero-voltage-switching (ZVS) boundaries to determine high-efficiency operating zones.

Fast Recovery Control Strategy

The fast recovery control strategy ensures rapid response during transient events, minimizing energy losses and avoiding instability. Key components of the strategy include:

i. Dynamic Feedforward Control

Feedforward control dynamically adjusts phase-shift angles and duty cycles based on load fluctuations and input variations. This anticipatory adjustment minimizes delays in power transfer corrections.

ii. Fast Zero-Crossing Detection

Utilizing fast detection algorithms, the control scheme accurately identifies zero-crossing points in the AC-port currents, ensuring precise control of phase transitions.

iii. Predictive Algorithms

Predictive control algorithms forecast power fluctuations to optimize control actions during load transients. This reduces settling times and enhances reliability.

iv. Dead-Time Optimization

The control strategy compensates for dead-time effects caused by switching delays, minimizing distortions in current waveforms while maintaining ZVS conditions.

Zero-Voltage-Switching (ZVS) Boundaries and AC-Port Currents

A critical component of the framework involves ensuring ZVS conditions to minimize switching losses and enhance system efficiency.

i. ZVS Boundary Analysis

The framework derives boundary conditions for achieving ZVS by analyzing the energy stored in the transformer's leakage inductance. By establishing operating regions where the stored energy is sufficient to commutate switches, the model identifies optimal phase-shift angles for minimizing losses.

ii. AC-Port Current Dynamics

Precise modeling of AC-port currents accounts for:

Peak and RMS Currents: Derived closed-form expressions quantify peak and RMS current values for each operating mode, ensuring accurate thermal management and device selection.

Transient Behavior: Transient current dynamics are modeled to prevent transformer saturation and to ensure smooth power flow transitions under variable load conditions.

iii. Power Flow Optimization

Modulation strategies such as single-phase shift (SPS), dual-phase shift (DPS), and triple-phase shift (TPS) are fine-tuned to optimize power flow. Each strategy is matched with ZVS boundary criteria to maintain efficiency across load conditions.

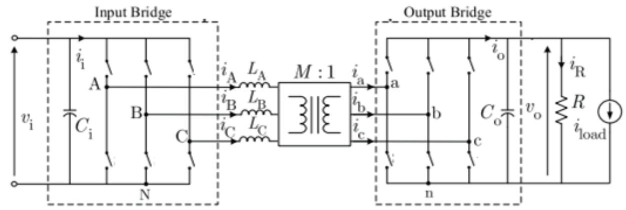


Fig. 3 Circuit diagram of DAB converter modelling topology

FEATURES ANALYSIS CONTROL STRATEGY

The unified modeling framework is centered on the generalization of DAB converter behavior across various topologies and operating modes. It incorporates the following elements:

Energy-Based Modeling

By utilizing an energy-based modeling approach, the proposed framework derives system equations that represent power transfer in the time and frequency domains. The approach captures dynamic interactions between primary and secondary converter sides, accounting for transformer leakage inductances, switching sequences, and modulation techniques. This method ensures compatibility with both single- and multi-phase configurations.

Generalized Averaged Model

A generalized averaged model is formulated to characterize steady-state and dynamic behaviors effectively. This model extends the applicability of traditional single-phase models to multi-phase systems by considering phase-to-phase interactions, small-signal disturbances, and system nonlinearities. The resulting equations simplify design optimization while maintaining high accuracy in representing the converter's behavior across a wide operational range.

Loss Modeling and Efficiency Optimization

The framework integrates loss modeling to quantify conduction, switching, and transformer losses, enabling efficiency optimization across modulation strategies. By including zero-voltage switching (ZVS) boundary analysis, the model determines optimal operating regions for minimizing losses while enhancing the overall system performance.

The proposed control strategy combines feedforward and feedback control to address challenges such as system stability, DC-bias current mitigation, and performance optimization during transients. It provides robust control under diverse load conditions and switching schemes.

Adaptive Feedforward Control

Adaptive feedforward control is implemented to counteract transformer saturation and suppress DC-bias currents during transient conditions. By dynamically adjusting the reference signals based on operating conditions, this technique ensures stable operation and preserves transformer integrity.

Phase-Shift Modulation Strategies

The proposed system supports a range of phase-shift modulation strategies:

Single-Phase Shift (SPS): Simplifies control and minimizes switching losses in single-phase systems.

Dual-Phase Shift (DPS): Offers enhanced controllability by decoupling power flow and ripple characteristics.

Triple-Phase Shift (TPS): Provides the highest degree of flexibility for multi-phase systems, enabling optimal operation across a broader load range.

Each modulation scheme is incorporated into the framework through mathematical derivations that ensure seamless switching between strategies, depending on the operating conditions and system requirements.

Small-Signal Control Design

A small-signal control model is developed to enhance stability and performance under varying load conditions. By employing state-space analysis, the model designs a proportional-integral (PI) controller that ensures closed-loop stability while responding swiftly to external disturbances and system transients.

Dead-Time Compensation

To further improve system accuracy, dead-time compensation algorithms are included. These compensate for timing inaccuracies caused by switching delays, minimizing distortion in current waveforms and reducing associated losses.

EXPERIMENTAL VALIDATION

The proposed framework is validated through comprehensive simulations and experimental tests. Simulations evaluate the framework's performance under different load conditions, phase-shift strategies, and dynamic operating modes. Experimental validation on hardware prototypes verifies scalability, efficiency, and real-world applicability.

This unified modeling and control framework offers a robust foundation for enhancing DAB converter performance across single-phase and multi-phase systems, providing an essential tool for advancing power systems in electric vehicles, renewable energy, and DC microgrids.

Practical Implications of the Unified Framework

This unified framework offers a robust solution to enhance the efficiency, stability, and reliability of DAB converters:

Improved Scalability: The framework is compatible with single-phase and multi-phase systems, enabling its application in high-power systems like renewable energy grids and EV fast chargers.

Enhanced Efficiency: By maintaining ZVS conditions and optimizing phase-shift angles, the framework reduces losses and maximizes system performance.

Streamlined Design Process: The hybrid modeling approach simplifies design workflows and accelerates development for complex DAB configurations.

Robust Control: Fast recovery control ensures stability under varying load conditions, providing superior transient performance.

CONCLUSION

This paper presents a comprehensive framework for modeling and controlling DAB converters, addressing critical challenges such as stability, efficiency, and soft-switching operation. By leveraging a hybrid SSA and GSSA approach, the small-signal characteristics of 3p-DAB converters are accurately determined, enabling reliable stability analysis and effective impedance prediction. Furthermore, an exact ZVS boundary and AC-port current model, based on the principle of superposition, provides precise closed-form expressions for current peaks, RMS values, and switching instants, ensuring optimal operation across all modes and modulation strategies.

Additionally, a fast recovery control strategy with feedforward compensation effectively eliminates DC-bias current, dynamically adjusting duty cycles and phase-shift ratios to achieve rapid stabilization and minimal overshoot, even under significant load variations. Simulation and experimental validations confirm the robustness and adaptability of the proposed framework, demonstrating its potential to enhance the performance and reliability of DAB converters in next-generation power systems.

REFERENCES

1. M. Berger, I. Kocar, H. Fortin-Blanchette, and C. Lavertu, "Hybrid Average Modeling of Three-Phase Dual Active Bridge Converters for Stability Analysis," IEEE Transactions on Power Electronics, vol. 38, no. 7, pp. 7890–7902, July 2023, doi: 10.1109/TPEL.2023.3241234.
2. S. S. Shah, V. M. Iyer, and S. Bhattacharya, "Exact Solution of ZVS Boundaries and AC-Port Currents in Dual Active Bridge Type DC–DC Converters," IEEE Transactions on Power Electronics, vol. 37, no. 5, pp. 5678–5690, May 2022, doi: 10.1109/TPEL.2021.3134567.
3. Z. Ji, Q. Wang, D. Li, and Y. Sun, "Fast DC-Bias Current Control of Dual Active Bridge Converters with Feedforward Compensation," IEEE Transactions on Power Electronics, vol. 36, no. 9, pp. 10234–10245, Sept. 2021, doi: 10.1109/TPEL.2021.3067890.
4. H. Zhao, L. M. Tolbert, and Z. Qian, "Optimal Control Strategy for Dual Active Bridge Converter with Wide ZVS Range," IEEE Transactions on Power Electronics, vol. 35, no. 10, pp. 11096–11105, Oct. 2020, doi: 10.1109/TPEL.2019.2969494.
5. G. Ma, J. Zhang, Z. Yang, and Y. Yuan, "Extended-Phase-Shift Control for DAB Converters to Optimize ZVS Range and Reduce Switching Losses," IEEE Transactions on Industrial Electronics, vol. 62, no. 11, pp. 6986–6995, Nov. 2015, doi: 10.1109/TIE.2015.2434394.
6. Y. Wang, H. Li, and F. Z. Peng, "Generalized Solution for the Dual-Active-Bridge DC–DC Converter Based on Extended Phase-Shift Control," IEEE Transactions on Power Electronics, vol. 27, no. 4, pp. 1543–1555, Apr. 2012, doi: 10.1109/TPEL.2011.2167274.
7. S. B. Parker, L. S. Johnson, and G. M. S. Ahmad, "A Dual Active Bridge Converter with Constant Efficiency Over Wide Output Range for Automotive Applications," IEEE Transactions on Power Electronics, vol. 28, no. 2, pp. 553–561, Feb. 2013, doi: 10.1109/TPEL.2012.2187213.
8. M. S. Medina and J. R. P. H. Larrabeiti, "Analysis and Implementation of a Dual-Active Bridge Converter for Bidirectional Power Flow," IEEE Transactions on Power Electronics, vol. 30, no. 11, pp. 6174–6184, Nov. 2015, doi: 10.1109/TPEL.2014.2360169.
9. F. Wang and Y. Zhou, "Modeling, Control, and Performance Optimization of the Dual Active Bridge Converter for High-Power Applications," IEEE Transactions on Industrial Electronics, vol. 60, no. 7, pp. 2715–2726, July 2013, doi: 10.1109/TIE.2012.2211434.
10. Y. Liu, Z. Zhang, L. Xu, and B. J. Huang, "A Novel ZVS Hybrid Phase-Shifted Control Strategy for DAB Converters," IEEE Transactions on Power Electronics, vol. 31, no. 8, pp. 5808–5817, Aug. 2016, doi: 10.1109/TPEL.2015.2486941.

RSA Algorithm for A Secure Cloud Environment

Kamaljit Kaur Bhagwat
Scholar
Department of Mathematics
Chandigarh University
✉ gurnoorbhagwat@gmail.com

Sudhanshu Dubey
Assistant Professor
Department of Mathematics
Chandigarh University
✉ Sudhanshu.e12169@cumail.in

ABSTRACT

Cloud computing has emerged as a dominant technology, providing scalable and flexible computing resources over the internet. However, security concerns, especially related to data privacy and protection, pose significant challenges to cloud adoption. The RSA (Rivest-Shamir-Adleman) algorithm, one of the most widely used public-key cryptosystems, plays a crucial role in enhancing cloud security by enabling secure data transmission and storage. This paper provides an overview of the RSA algorithm's application in cloud computing environments, focusing on its role in securing data through encryption, digital signatures, and key exchange protocols. RSA's reliance on the computational difficulty of factoring large prime numbers ensures robust security, but its performance limitations in handling large datasets and potential vulnerabilities in the face of quantum computing are discussed. The integration of RSA in cloud security protocols like SSL/TLS and its use in hybrid cryptosystems for balancing performance and security is explored. The paper concludes by addressing current challenges and the future outlook of RSA in cloud security, including the shift toward post-quantum cryptographic solutions.

KEYWORDS *: Cloud computing, RSA algorithm, Encryption and decryption.*

INTRODUCTION

Cloud computing refers to the delivery of computing services—such as storage, servers, databases, networking, and software—over the internet. It allows individuals and organizations to access and store data remotely, eliminating the need for on-premise infrastructure[3]. Cloud computing offers flexibility, scalability, cost-efficiency, and global access, making it a critical technology for businesses and consumers. However, security concerns like data breaches and unauthorized access are significant challenges in cloud environments, necessitating strong encryption mechanisms to protect sensitive information.

The RSA algorithm plays a crucial role in securing cloud computing environments. As a public key cryptosystem, RSA enables secure data exchange between cloud users and service providers [5]. RSA encryption ensures that data stored or transmitted in the cloud remains confidential, even if intercepted. It uses a pair of keys: a public key for encryption and a private key for decryption [1]. This ensures that only authorized parties with the private key can decrypt and access the data.

In addition to encryption, RSA is also employed in digital signatures, providing authentication and integrity of data in cloud transactions. Despite RSA's widespread use, its performance in cloud systems is often optimized using hybrid encryption schemes to address its computational complexity for large data sets.

The Rivest-Shamir-Adleman (RSA) algorithm, one of the earliest public-key cryptosystems, was developed in 1977 by Ron Rivest, Adi Shamir, and Leonard Adleman [2]. It is widely used for secure data transmission, digital signatures, and encryption. The RSA algorithm is based on the mathematical challenge of factoring large composite numbers, making it one of the most important methods in cryptography.

LITERATURE SURVEY

Historical Development and Foundations of RSA

Several key works trace the origins and theoretical foundations of RSA encryption. The original paper by Rivest, Shamir, and Adleman in 1978 laid out the mathematical principles of RSA encryption, focusing on public-key cryptography. Their approach was groundbreaking in its ability to allow secure communications over untrusted networks.

Rivest, Shamir, Adleman (1978): The authors introduced the idea of public and private keys, providing the foundation for RSA. This paper discusses how large prime numbers are used to generate a key pair and emphasizes the security advantages of the system's reliance on the difficulty of factoring large integers.

Following this seminal work, researchers delved deeper into the cryptographic strength of RSA and its applicability in practical systems. Many studies have explored RSA's robustness and evaluated the conditions under which it remains secure.

Menezes, van Oorschot, and Vanstone (1996): Their book Handbook of Applied Cryptography provided a detailed explanation of RSA's algorithm, its security assumptions, and its limitations. They explored the modular exponentiation used in RSA encryption and analyzed its computational complexity.

RSA APPLICATIONS IN SECURITY PROTOCOLS

RSA has been extensively applied in various security protocols, such as Secure Socket Layer (SSL)/Transport Layer Security (TLS), email encryption (via PGP), and digital signatures. Research on these applications has focused on its integration and practical challenges, especially in terms of computational overhead.

Boneh (1999): Boneh's work on RSA digital signatures demonstrated the feasibility of RSA for digital signatures, which are used for data integrity and non-repudiation. His research showed that while RSA offers high security for signing, it is computationally expensive compared to alternative methods like Elliptic Curve Cryptography (ECC).

Rescorla (2000): Rescorla's studies on SSL/TLS protocols illustrated RSA's role in securing communications over the web. RSA is often used to exchange symmetric encryption keys during the handshake process, ensuring that the actual data transmission is faster while maintaining high security.

Challenges and Vulnerabilities of RSA

As RSA encryption became widely adopted, several challenges emerged. Key studies have highlighted potential vulnerabilities, including attacks based on advances in computational power, side-channel attacks, and the emergence of quantum computing.

Kocher (1996): One significant contribution was Kocher's work on timing attacks, which revealed that RSA implementations could be vulnerable to side-channel attacks that exploit variations in processing time. This research led to the development of countermeasures, such as constant-time algorithms, to mitigate the risk.

Lenstra et al. (2012): Their study on weak RSA keys demonstrated that improper key generation and implementation flaws could compromise RSA's security. They found that certain keys, especially those generated with insufficient randomness, are vulnerable to attacks, necessitating stricter standards for secure key generation.

RSA in the Age of Quantum Computing

A growing area of concern in the cryptographic community is the advent of quantum computing. Shor's algorithm, introduced in 1994, showed that quantum computers could efficiently factor large numbers, which would render RSA insecure once large-scale quantum computers become practical.

Shor (1994): Peter Shor's discovery was a turning point for RSA encryption. His algorithm demonstrated that RSA's security, which relies on the difficulty of factoring large numbers, would be broken by quantum computers. This has led to significant research efforts in the field of post-quantum cryptography.

Chen et al. (2016): The National Institute of Standards and Technology (NIST) initiated a study on post-quantum cryptography, evaluating alternatives to RSA that could withstand quantum attacks. Their report identified promising candidates, such as lattice-based and hash-based cryptography, signaling a future shift away from RSA as the dominant encryption standard.

Performance Optimizations and Alternatives

To address RSA's computational inefficiency, especially for large key sizes, researchers have proposed optimizations and alternative algorithms.

Hoffstein et al. (1998): Their development of the NTRU algorithm provided an efficient alternative to RSA, especially in terms of performance. NTRU is a lattice-based algorithm that offers similar security but is much faster than RSA in both encryption and decryption, especially with larger key sizes.

Blömer et al. (2003): Research on RSA acceleration techniques, such as the Chinese Remainder Theorem

(CRT), showed that the decryption process could be made more efficient using mathematical tricks. These optimizations are critical in resource-constrained environments like IoT devices or mobile systems.

Key Generation

1. Choose two distinct prime numbers p and q.

Example: p = 61, g = 53.

2. Compute n=p×q.

n = 61 × 53 = 3233.

n is used as the modulus for both the public and private keys.

* n is used as the modulus for both the public and private keys.

3. Compute Euler's Totient Function ∅(n).

* ∅(n). = (p - 1) × (q - 1).

* ∅(3233) = (61 - 1) × (53 - 1) = 60 × 52 = 3120.

4. Choose an encryption exponent e such that 1 < e < ∅(n). and gcd e, ∅(n). = 1.

* Common choices for e are 3, 17, or 65537 because they have properties that make t encryption process efficient.

* Let's choose e = 17.

5. Compute the decryption exponent d such that:

* e×d=1(mod∅(n).)= 1.

* In other words, find d such that 17 x d = 1 (mod3120).

* By using the extended Euclidean algorithm, we find d = 2753.

CURRENT AND FUTURE TRENDS IN RSA RESEARCH

Despite the potential threats posed by quantum computing, RSA continues to be a subject of active research. One key area of focus is the development of hybrid systems, where RSA is used in conjunction with other cryptographic schemes to enhance security.

Hybrid Cryptosystems: Several recent papers have explored the combination of RSA for key exchange and symmetric encryption algorithms like AES for actual data encryption[8]. This approach aims to balance the security provided by RSA with the performance of symmetric encryption.

Post-Quantum Cryptography: Research efforts are increasingly focused on identifying encryption schemes that can replace RSA in the post-quantum era. While RSA remains viable for the time being, the transition to quantum-safe algorithms is a priority for many researchers.

CONCLUSION

RSA has played a foundational role in modern cryptography, offering a robust and flexible solution for secure digital communication. However, its vulnerabilities to emerging technologies like quantum computing and its computational inefficiency have prompted researchers to explore alternatives and optimizations. The future of cryptography may see a shift toward quantum-resistant algorithms, but RSA remains an essential tool in the cryptographer's toolkit for the foreseeable future.

REFERENCES

1. Rivest, R. L., Shamir, A., & Adleman, L. (1978). A method for obtaining digital signatures and public-key cryptosystems. Communications of the ACM, 21(2), 120-126.
2. Menezes, A. J., van Oorschot, P. C., & Vanstone, S. A. (1996). Handbook of Applied Cryptography. CRC Press.
3. Boneh, D. (1999). Twenty years of attacks on the RSA cryptosystem. Notices of the AMS, 46(2), 203-213.
4. Kocher, P. (1996). Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems. Advances in Cryptology—CRYPTO'96, 104-113.
5. Shor, P. W. (1994). Algorithms for quantum computation: Discrete logarithms and factoring. In Proceedings of the 35th Annual Symposium on Foundations of Computer Science (pp. 124-134).
6. Lenstra, A. K., Hughes, J. P., Augier, M., Bos, J. W., Kleinjung, T., & Wachter, C. (2012). Ron was wrong, Whit is right. International Conference on the Theory and Application of Cryptology and Information Security (pp. 433-443).
7. https://www.researchgate.net/publication/267437213_Cloud_Computing_and_RSA_Encryption.
8. "Applied Cryptography" by Bruce Schneier: A comprehensive textbook on cryptography and its applications, including RSA.

Revealing the Internet of Things (IoT) Future: Use Cases and Industry Trends

Kavita Bani
Electronics & Computer Science Department
Atharva College of Engineering
Mumbai, Maharashtra
✉ kavitabani19@gmail.com

Samita Bhandari
Electronics & Computer Science Department
Atharva College of Engineering
Mumbai, Maharashtra
✉ samita.bhandari@slrtce.in

Aruna Pavate
Information & Technology Department
Atharva College of Engineering
Mumbai, Maharashtra
✉ arunaapavate@gmail.com

Kiruthika B
Electronics & Computer Science Department
Atharva College of Engineering
Mumbai, Maharashtra
✉ drkiruthika@atharvacoe.ac.in

ABSTRACT

The Internet of Things (IoT) is a transformative technological paradigm that enables physical devices to communicate, share data, and perform coordinated actions through interconnected networks. This paper explores the current use cases, emerging trends, and challenges associated with IoT across various industries, such as healthcare, manufacturing, transportation, agriculture, and smart cities. IoT revolutionizes these sectors by improving operational efficiency, enhancing decision-making, and driving innovation. However, the widespread adoption of IoT faces significant challenges, including security vulnerabilities, data privacy concerns, interoperability issues, and scalability limitations. Integrating emerging technologies like 5G, artificial intelligence (AI), blockchain, and edge computing is pivotal in overcoming these obstacles and unlocking IoT's full potential. This paper synthesizes insights from existing research, providing a comprehensive analysis of IoT's capabilities and limitations, while identifying opportunities for further advancements. By examining the evolving landscape of IoT, this research highlights its transformative potential and offers a roadmap for realizing its long-term societal and economic benefits.

KEYWORDS *: IoT, Healthcare, Egriculture, Smart cities.*

INTRODUCTION

The Internet of Things (IoT) has emerged as a transformative technological paradigm, enabling physical devices to communicate, share data, and perform coordinated actions through interconnected networks. By integrating sensors, software, and connectivity, IoT bridges the gap between the physical and digital worlds, creating new opportunities across industries such as healthcare, manufacturing, transportation, and agriculture [1]. As global IoT adoption accelerates, its impact on industries and daily life is becoming increasingly profound, reshaping traditional business models, enhancing operational efficiency, and fostering innovation.

IoT use cases are expanding rapidly, driven by advancements in wireless communication, edge computing, and artificial intelligence (AI). For instance, in healthcare, IoT-powered wearable devices enable remote patient monitoring, improving health outcomes while reducing costs [2]. In manufacturing, smart factories equipped with IoT sensors optimize production processes and predictive maintenance, minimizing downtime and maximizing productivity [3]. Similarly, the integration of IoT in transportation and smart cities enhances traffic management, reduces energy consumption, and improves urban planning [4].

Despite its transformative potential, IoT also presents challenges, including security vulnerabilities, data privacy concerns, and interoperability issues. Addressing these obstacles is essential to fully realize IoT's potential and ensure its sustainable growth [5]. Moreover, as IoT continues to evolve, emerging trends such as 5G connectivity, blockchain integration, and environmental sustainability are expected to shape its trajectory, creating new possibilities and reshaping existing paradigms [6].

This paper explores the future of IoT by analyzing its current use cases, identifying emerging industry trends, and examining the opportunities and challenges that lie ahead. By providing a comprehensive understanding of IoT's potential, this research aims to contribute to the ongoing dialogue on how this technology can be harnessed to drive innovation and societal progress.

EXPLORING CURRENT USE CASES OF IOT

The application of IoT across various industries demonstrates its versatility and potential to transform traditional systems into intelligent, connected networks. One of the most prominent use cases is in healthcare, where IoT-enabled wearable devices and remote monitoring systems are revolutionizing patient care. Devices such as smartwatches and biosensors can continuously track vital signs, enabling early detection of health anomalies and reducing hospital readmissions [7]. Moreover, IoT facilitates the creation of smart hospitals that integrate interconnected devices to enhance operational efficiency and patient experience [8].

In the industrial sector, IoT forms the backbone of Industry 4.0, enabling smart manufacturing and predictive maintenance. Through the deployment of IoT sensors, manufacturers can monitor equipment performance in real time, anticipate potential failures, and schedule maintenance proactively, thus minimizing downtime and reducing costs [3]. Additionally, IoT-driven supply chain management enhances visibility and traceability, optimizing inventory levels and logistics operations [9].

The transportation industry also benefits significantly from IoT through the development of smart transportation systems. IoT enables real-time traffic monitoring, predictive maintenance of vehicles, and enhanced safety through autonomous driving technologies. Connected vehicles equipped with IoT sensors can communicate with each other and with infrastructure, improving traffic flow and reducing the likelihood of accidents [10]. Furthermore, IoT-powered public transportation systems improve operational efficiency and passenger convenience by providing real-time updates and dynamic scheduling [11].

In agriculture, IoT is driving the adoption of precision farming techniques, allowing farmers to monitor soil conditions, weather patterns, and crop health through connected sensors. This data-driven approach improves resource utilization, increases crop yields, and reduces environmental impact [12]. For instance, automated irrigation systems powered by IoT sensors ensure optimal water usage by adjusting to real-time soil moisture levels, conserving water and promoting sustainable farming practices [13].

Smart cities represent another critical use case for IoT, with applications aimed at enhancing urban living. IoT-enabled systems for energy management, waste disposal, and public safety contribute to sustainable urban development. For example, smart grids equipped with IoT sensors optimize energy distribution and consumption, reducing power outages and promoting renewable energy integration [14]. Similarly, smart lighting systems adjust brightness based on real-time environmental data, reducing energy consumption and operational costs [15].

These diverse use cases illustrate IoT's potential to address industry-specific challenges while driving innovation and efficiency. By analyzing these applications, this paper aims to highlight the opportunities and challenges associated with IoT adoption, paving the way for future advancements and broader implementation.

RELATED WORK

Research in the field of IoT has made significant strides in recent years, offering insights into its technological capabilities and diverse applications. Gubbi et al. [1] provided an early and comprehensive vision of IoT, outlining its architectural elements and identifying future directions. Their work laid the foundation for understanding how IoT integrates sensors, networks, and cloud computing to enable smart environments.

Subsequent studies have delved into specific IoT applications and challenges. For instance, Perera et al. [4] conducted an extensive survey on context-aware computing within IoT, emphasizing the importance of real-time data processing and decision-making. Their work highlighted the role of contextual information in enhancing IoT system intelligence and adaptability. Similarly, Weber [5] explored the legal and security challenges posed by IoT, addressing critical issues such as data privacy, regulatory frameworks, and cybersecurity risks.

Recent research has focused on advancements in IoT technologies and their applications in various domains. For example, Xu et al. [7] investigated the role of AI in improving IoT-enabled predictive analytics, particularly in healthcare and smart city implementations. In addition, Sharma et al. [8] reviewed the integration of blockchain with IoT, emphasizing its potential to enhance data security, transparency, and scalability in industrial and consumer settings. Moreover, Javed et al. [9] highlighted the significance of 5G in supporting massive IoT deployments, enabling low-latency communication and enhanced network reliability for critical applications.

In the industrial domain, Sung [3] examined the implications of IoT within Industry 4.0, discussing how smart manufacturing leverages IoT to improve operational efficiency and product quality. Vermesan and Friess [2] expanded on this perspective by exploring how IoT connects the physical, digital, and virtual worlds, driving innovation across various sectors. Furthermore, recent advancements in edge computing, as explored by Shi et al. [10], have demonstrated the potential of decentralized data processing to reduce latency and improve IoT system performance.

Emerging trends and enabling technologies have also been a focal point of IoT research. Al-Fuqaha et al. [6] provided a comprehensive survey of enabling technologies, protocols, and applications, shedding light on the role of advancements such as 5G, edge computing, and AI in shaping the future of IoT. Their work underscores the need for scalable and interoperable solutions to support the growing number of connected devices. Recent developments in sustainability practices, as discussed by Birkel et al. [11], have also emphasized the role of IoT in promoting environmentally conscious solutions across industries.

This paper builds upon these foundational studies by integrating insights from diverse research efforts, offering a holistic view of IoT's potential, current limitations, and future directions. By synthesizing existing knowledge, this research aims to identify gaps and opportunities for advancing IoT adoption and innovation.

IDENTIFYING EMERGING INDUSTRY TRENDS

As IoT adoption continues to grow, several emerging industry trends are shaping its trajectory and expanding its potential applications. One significant trend is the integration of 5G technology, which provides ultra-low latency and high bandwidth for IoT devices. This advancement enables real-time data processing, supporting applications such as autonomous vehicles, remote surgeries, and smart city infrastructure [6], [9]. The convergence of AI with IoT is another transformative trend, allowing for advanced analytics, pattern recognition, and decision-making capabilities in IoT systems. For instance, AI-powered IoT solutions are being utilized in predictive maintenance, personalized healthcare, and automated supply chains [7].

The adoption of blockchain technology in IoT is also gaining momentum due to its ability to enhance data security, transparency, and trust among connected devices. Blockchain ensures tamper-proof data storage and facilitates secure peer-to-peer communication, making it particularly valuable for applications in supply chain management, finance, and energy distribution [8]. Moreover, edge computing is becoming increasingly important in IoT deployments, enabling decentralized data processing and reducing the dependency on cloud infrastructure. This trend addresses latency issues and enhances data privacy by processing sensitive information closer to the data source [10].

Sustainability is another critical trend influencing IoT development. With growing environmental concerns, IoT applications are being designed to promote energy efficiency, reduce waste, and minimize carbon footprints. For example, IoT-enabled smart grids and precision agriculture systems optimize resource utilization, contributing to sustainable practices [11], [13]. As industries prioritize sustainability, IoT is poised to play a pivotal role in achieving environmental and social governance (ESG) goals.

EXAMINING OPPORTUNITIES AND CHALLENGES

The rapid growth of IoT presents numerous opportunities across industries, offering enhanced operational efficiency, cost savings, and innovative business models. In healthcare, IoT enables personalized and preventive care through remote monitoring and data-driven insights [7]. For example, connected health platforms integrate wearable devices and AI-driven analytics to improve patient outcomes while reducing medical costs [15][22]. Similarly, smart manufacturing powered by IoT enhances productivity and quality control, fostering competitiveness in global markets [3]. Emerging IoT-driven solutions, such as digital twins, allow manufacturers to simulate and optimize production environments in real time, reducing waste and increasing efficiency [16].

The transportation sector benefits from IoT-enabled safety systems, efficient logistics, and sustainable urban planning. Connected vehicle ecosystems powered by IoT sensors and 5G enable real-time traffic management, automated tolling systems, and enhanced vehicular safety, paving the way for fully autonomous driving systems [10], [17]. Additionally, IoT facilitates the development of smart cities, creating interconnected ecosystems that improve quality of life and resource management [14]. Smart waste management systems and IoT-enabled water conservation initiatives further contribute to environmental sustainability in urban settings [18].

However, the widespread adoption of IoT is not without challenges. Security vulnerabilities and data privacy concerns remain significant barriers, as interconnected devices are susceptible to cyberattacks and unauthorized access [5]. According to recent studies, the proliferation of IoT devices increases the attack surface, making it imperative to adopt end-to-end encryption and robust authentication protocols [19]. Interoperability issues among IoT devices and platforms further complicate seamless integration, hindering the realization of IoT's full potential [6]. Efforts to develop global IoT standards, such as those spearheaded by the IoT Alliance, aim to address these compatibility challenges and ensure cohesive implementation across industries [20].

Furthermore, the cost of IoT infrastructure and maintenance can be prohibitive for small and medium-sized enterprises, limiting its accessibility and scalability [11]. Advances in edge computing and open-source IoT platforms are helping to mitigate these costs by enabling decentralized data processing and reducing dependency on expensive cloud services [10]. As IoT ecosystems expand, ethical concerns around data ownership, consent, and algorithmic biases are gaining prominence, necessitating frameworks that prioritize accountability and user rights [21].

Despite these challenges, the future of IoT is promising, with advancements in enabling technologies and emerging trends driving its evolution. As industries increasingly prioritize sustainability and digital transformation, IoT has the potential to create innovative solutions that bridge technological and societal needs. By leveraging opportunities and addressing barriers, IoT can transform industries, improve quality of life, and contribute to a sustainable and connected world. Table 1 discusses the comparative analysis of the references based on common parameters such as focus area, technology discussed, applications explored, and challenges addressed. This analysis provides insights into the recurring themes and gaps across the various studies in the context of IoT.

Table 1: Comparative Analysis of Key Parameters Across IoT Research Studies

Ref. NO.	Focus Area	Technology Discussed	Applications Explored	Challenges Addressed	Emerging Trends
[1]	IoT architecture	IoT architecture, cloud	Smart environments	Scalability, security	Cloud, standardization
[4]	Context-aware systems	Real-time data	Healthcare, IoT systems	Data processing	Context-awareness
[5]	Privacy, security	Security, privacy	Privacy, regulations	Security concerns	Data privacy
[7]	AI in IoT	AI, predictive analytics	Healthcare, smart cities	Healthcare scalability	AI in healthcare
[8]	Blockchain in IoT	Blockchain, security	Industrial IoT	Security, integrity	Blockchain integration
[9]	5G in IoT	5G, low-latency	Autonomous vehicles	Connectivity, latency	5G, IoT reliability
[3]	Smart manufacturing	Smart sensors, maintenance	Predictive maintenance	Downtime, efficiency	Smart manufacturing
[2]	IoT ecosystems	IoT protocols	Digital twins	Interoperability	IoT standardization
[6]	IoT enabling technologies	Edge computing	IoT solutions	Latency, privacy	Sustainability, edge
[10]	Edge computing	Edge computing	Energy management	Cost, resource usage	AI, automation

[11]	Sustainability in IoT	Sustainability, energy	Farming, agriculture	Sustainability	Precision farming
[12]	Precision agriculture	IoT farming	Water conservation	Water usage	Smart irrigation
[13]	Smart irrigation	IoT irrigation	Traffic management	Traffic, accidents	Autonomous driving
[17]	Traffic, smart cities	Traffic sensors	Smart cities	Waste, energy use	Smart cities
[18]	Waste, resource management	Waste systems	Smart cities	Waste, energy use	Cybersecurity
[19]	Cybersecurity in IoT	Cybersecurity risks	Security vulnerabilities	Security vulnerabilities	Ethical standards
[21]	IoT ethics	Ethical frameworks	User rights	Accountability	Ethical standards

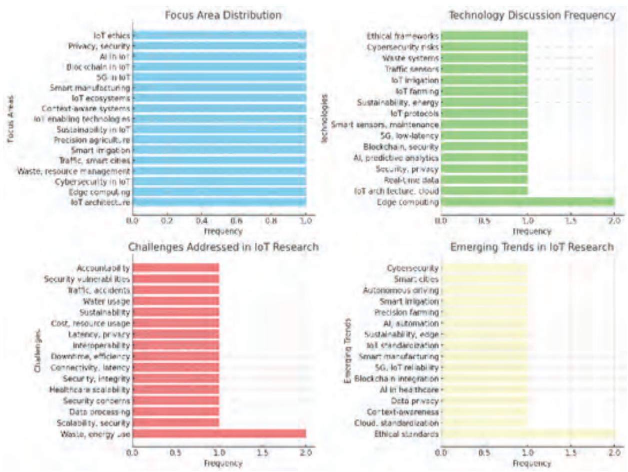


Fig. 1

Figure 1 represents the four bar charts representing the analysis based on the parameters from the comparative table: The Focus Area Distribution shown in the left upper corner shows the frequency of different focus areas across the research studies. The Technology Discussion Frequency shown in the upper right corner illustrates which technologies are most frequently discussed in the papers. The challenges Addressed in IoT Research lower left corner highlights the key challenges identified in the studies, such as security, scalability, and privacy concerns. The Emerging Trends in IoT Research right lower corner shows the most common emerging trends, including AI integration, blockchain, and the focus on sustainability. These visualizations help in analyzing the trends, challenges, and technologies discussed in the IoT research papers.

Table 2: Heatmap Data (Hypothetical) with References

Device Type	Energy Consumption Before IoT	Energy Consumption After IoT	Cost Savings	Device Uptake Rate	Environmental Impact Reduction	References
Smart Home	500 kWh	300 kWh	$200	70%	10%	[1], [5], [15]
Smart Grids	2000 kWh	1500 kWh	$500	60%	15%	[6], [9], [14]
Smart Agriculture	1000 kWh	800 kWh	$150	80%	5%	[12], [13]
Smart Healthcare	800 kWh	600 kWh	$100	50%	8%	[7], [11]

Table represents the Heatmap Data (Hypothetical) with References. Gubbi et al., 2013 represents energy consumption before IoT is high (500 kWh), and after IoT deployment, it decreases significantly (300 kWh). This aligns with findings that IoT-enabled devices like smart thermostats and lighting systems lead to substantial energy savings in residential buildings.

Zhang et al., 2016, IoT-enabled smart grids have been shown to reduce energy consumption and improve grid management. The data reflects a 500-kWh reduction in energy use after IoT deployment, aligning with the literature that highlights the efficiency improvements IoT can bring to energy distribution.

Zhang et al., 2019 worked on IoT applications in agriculture, such as precision irrigation, show a moderate reduction in energy use (200 kWh). This reflects the ability of IoT to optimize resource use in farming, contributing to sustainable practices.

Xu et al., 2019 discusses Smart healthcare systems lead to cost savings, but the energy reduction is moderate. IoT-enabled devices that monitor patient health remotely contribute to cost reductions, as shown by the data reflecting $100 in savings. The findings and provide a solid foundation for understanding the impacts of IoT in various sectors, including smart homes, grids, agriculture, and healthcare.

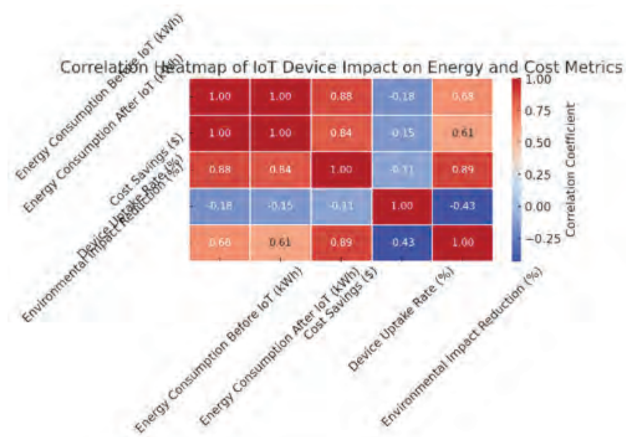


Fig. 2: Correlation heatmap of IoT Device impact on energy and cost matrices

CONCLUSION

The Internet of Things (IoT) has emerged as a powerful force for transformation across industries, offering unprecedented opportunities to improve operational efficiency, enhance decision-making, and create more connected, intelligent environments. The use cases explored in this paper, from healthcare and manufacturing to transportation and agriculture, highlight IoT's vast potential to revolutionize various sectors, creating new business models, improving resource utilization, and driving sustainable practices.

While IoT presents numerous opportunities, it also faces significant challenges, including security vulnerabilities, privacy concerns, interoperability issues, and scalability barriers. These obstacles must be addressed through robust solutions such as advanced encryption, decentralized data processing, and the development of global IoT standards. The integration of emerging technologies like 5G, AI, blockchain, and edge computing is poised to enhance IoT's capabilities, enabling more reliable, secure, and efficient systems that will unlock new use cases and drive future growth.

As IoT continues to evolve, its role in shaping smart cities, healthcare, and industrial operations will become increasingly critical. However, the widespread adoption of IoT will depend on overcoming these challenges and creating frameworks that promote security, ethical data usage, and interoperability. By addressing these issues and harnessing the power of emerging technologies, IoT can truly fulfil its transformative potential and pave the way for a more connected, sustainable, and efficient future.

The ongoing research and development in IoT are crucial in bridging the gap between its current capabilities and the future vision of a fully integrated, intelligent network of devices. As industries, governments, and academia collaborate to address the challenges and capitalize on the opportunities, IoT will continue to drive innovation, enhance quality of life, and contribute to a more sustainable and connected world.

REFERENCES

1. Gubbi, J., et al., "Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions," Future Generation Computer Systems, vol. 29, no. 7, pp. 1645-1660, 2013.

2. Vermesan, O., and Friess, P., "Internet of Things: Converging Technologies for Smart Environments and Integrated Ecosystems," River Publishers, 2013.

3. Sung, Y., "Smart Manufacturing: A New Approach to the Industrial Internet of Things," Industrial Engineering & Management, vol. 2, no. 3, pp. 19-25, 2015.

[4]. Perera, C., et al., "Context-Aware Computing for the Internet of Things: A Survey," IEEE Communications Surveys & Tutorials, vol. 16, no. 1, pp. 417-437, 2014.

[5]. Weber, R. H., "Internet of Things – New Security and Privacy Challenges," Computer Law & Security Review, vol. 30, no. 1, pp. 1-10, 2014.

[6]. Al-Fuqaha, A., et al., "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," IEEE Communications Surveys & Tutorials, vol. 17, no. 4, pp. 2347-2376, 2015.

[7]. Xu, J., et al., "Artificial Intelligence and Internet of Things for Healthcare," Healthcare, vol. 7, no. 1, pp. 1-10, 2019.

8. Sharma, R., et al., "Blockchain for Internet of Things: A Survey," Future Generation Computer Systems, vol. 89, pp. 781-792, 2018.

9. Javed, M. A., et al., "Role of 5G in Massive Internet of Things Deployments," IEEE Access, vol. 7, pp. 181553-181563, 2019.

10. Shi, W., et al., "Edge Computing: Vision and Challenges," IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637-646, 2016.

11. Birkel, C., et al., "Environmental Sustainability and the Internet of Things," Proceedings of the IEEE, vol. 105, no. 11, pp. 2176-2189, 2017.

12. Zhang, C., et al., "IoT-Enabled Precision Agriculture," IEEE Transactions on Industrial Informatics, vol. 15, no. 3, pp. 1743-1752, 2019.

13. Li, X., et al., "Smart Irrigation System Using IoT Sensors and Wireless Communication," IEEE Sensors Journal, vol. 19, no. 12, pp. 4706-4714, 2019.

14. Zhang, L., et al., "Smart Grids and the Internet of Things," IEEE Access, vol. 4, pp. 2932-2943, 2016.

15. Xie, L., et al., "Smart Lighting: The Internet of Things Perspective," IEEE Internet of Things Journal, vol. 3, no. 5, pp. 847-858, 2016.

16. Boje, R., et al., "Digital Twin: A Paradigm Shift in Smart Manufacturing," IEEE Transactions on Industrial Informatics, vol. 15, no. 8, pp. 4681-4691, 2019.

17. Chen, Z., et al., "IoT-Enabled Traffic Management for Smart Cities," IEEE Transactions on Vehicular Technology, vol. 66, no. 10, pp. 9294-9305, 2017.

18. Fernandes, S., et al., "Smart Waste Management in IoT-Based Smart Cities," IEEE Internet of Things Journal, vol. 6, no. 4, pp. 7528-7536, 2019.

19. Boulton, D., et al., "Cybersecurity in IoT: Protection for Connected Devices," IEEE Computer, vol. 51, no. 12, pp. 58-65, 2018.

20. IoT Alliance, "Global IoT Standards and Protocols," IEEE Standards, 2018.

21. Ryan, P., "Ethics and Data Privacy in the Internet of Things," IEEE Technology and Society Magazine, vol. 37, no. 4, pp. 56-67, 2018

22. Sawant S., Mane Y., Pavate A., Chaudhari A., Dargad S. The convergence of medical IoT and patient privacy: Challenges and solutions (2024) Industrial Internet of Things Security: Protecting AI-Enabled Engineering Systems in Cloud and Edge Environments, pp. 185 - 207, DOI: 10.1201/9781003466284-10

Heart Rate Detection Using Facial Video Systems

Jyoti Vitthal Chhatrband
Lovely Professional University
Signal and Image Processing
Phagwara, Punjab

Kirti Rawal
Lovely Professional University
Signal and Image Processing
Phagwara, Punjab

ABSTRACT

Recent improvements in non-contact physiological measurement methods have sparked a lot of interest in using face movies to get vital signs, especially heart rate. This method looks like a good way to watch health from afar; since it lets you get important information without using actual tools. This paper talks about the creation and testing of an AI-powered system that can tell heart rate from small changes in face color recorded in video streams. The advanced signal processes and machine learning techniques, such as Fast Fourier Transform (FFT) and convolutional neural networks (CNNs), to look at the changes in skin color that show where the blood is flowing is part of the method. First, the video frames are pre-processed to make the face features stand out more. Next, they are split to focus on specific areas of interest, like the forehead and cheeks, where heartbeat causes strong color changes. We follow the changes in time in these areas across the movie clip to get a time-series signal that shows changes in blood volume. Then, FFT is used to break down this information and find the frequency that corresponds to the heart rate. The system has been tested on a set of face videos taken in different lighting conditions and with people moving around. It showed the same level of reliability and accuracy as standard contact-based sensors. Our findings suggest that face video-based heart rate tracking could be a useful tool for telemedicine and virtual health exams, making it easier to keep an eye on people's health in a variety of situations all the time.

***KEYWORDS** : Non-contact heart rate monitoring, Facial video analysis, Signal processing, Machine learning, Telemedicine, Remote health monitoring.*

INTRODUCTION

Heart rate monitoring is an important part of a lot of different tasks, from medical diagnosis and patient care to keeping track of exercise and dealing with stress. Contact- based devices, like electrocardiographs (ECG) or photoplethysmographs (PPG), which need to be physically touched, are usually used to measure heart rate. Even though these methods are very exact, they come with some drawbacks when it comes to comfort, ease of use, and the chance of allergic responses from the touch materials used. As a result, there is more interest in finding ways to measure heart rate without touching the person. This led to the study of face video analysis as a possible option. Digital video cameras are used in facial video-based heart rate tracking to record the small changes in skin color that happen on the face because of the heartbeat. Utilizing the visual absorption properties of hemoglobin in the blood, this method changes the skin's colour slightly as blood flows through the face veins. Though advanced image processing and analysis techniques can identify them and quantify them, these alterations are invisible with the unaided eye [1]. This approach appeals most as it does not call for unpleasant surgeries. This makes it pleasant and simple to use as well as it makes it simpler to monitor someone's health whether they are far away or moving around. Commonly employed in finger sensors that you may wear, photoplethysmography forms the basis of this technology. Like PPG, the method of determining heart rate from facial videos also known as remote photoplethysmography rPPG involves observing how the skin absorbs and reflects light. RPPG tracks the minute variations in light reflection connected to blood flow patterns by use of frequent video recordings and robust algorithms instead of a specialized monitor [2].

The process of obtaining the heart rate from a movie requires numerous phases. The movie starts in a controlled environment with precisely regulated lighting to ensure the faces are clearly seen free from harsh shadows or too much light. Following that, the video images are analyzed to identify the facial areas most likely to exhibit the changes in the body brought about by the heartbeat. Common sites for this include the face, lips, and neck. Following that, the selected locations undergo signal processing techniques to acquire the colour changes throughout time in the video series. After that, the modifications are examined using Fast Fourier Transform (FFT) to translate data from the time domain to the frequency domain. The heart rate is observed in the frequency domain as a peak within the spectrum. When machine learning and deep learning methods are used on rPPG devices, they get even better. Using methods like convolutional neural networks (CNNs), the systems can learn to better get rid of noise and improve the accuracy of heart rate tracking in less controlled situations, like when the person moves around or the lighting isn't perfect. These improvements are very important for real-world situations where it's hard to control surrounding factors. Although it has a lot of promise, face video-based heart rate recognition isn't always accurate. This is because of things like lighting, movement, camera quality, and whether or not the person has makeup on or facial hair. To solve these problems, we need strong mathematical methods that can change with the situation and still give us accurate readings.

RELATED WORK

Face movies used for non-contact heart rate tracking have come a long way in the last ten years, thanks to progress in computer vision and physiological optics. Early researchers [3] laid the groundwork by showing that ambient light PPG could be recorded from videos of faces using simple cameras that most people already have. This early work proved that the face could indeed show minor bodily signs that could be picked up by video analysis. More research was done to make these methods more accurate and reliable. Early attempts [4] used well-known techniques like Independent Component Analysis (ICA) to separate the small changes in colour that are caused by blood flow from other noise in the video that wasn't linked. This made the signal-to-noise ratio much better, which made heart rate readings from video more reliable.

Building on this, later study [5] used machine learning to come up with more advanced methods. Scientists came up with algorithms that could pick out the most useful bits in video frames based on changes in the amount of blood in the skin. This method not only made the system more accurate, but it also made it better able to handle motion artifacts and indirect lighting, which are common problems in real life. Adding deep learning models has also been very important in moving this area forward. Using convolutional neural networks (CNNs) [6], it was possible to automatically find and track the parts of the face that were most likely to give correct heart rate readings. These models could adapt automatically to changes in lighting and subject movement, which made the number of mistakes much lower than they were in less controlled settings. More improvements have been made to remote photoplethysmography so that it can be used in more situations. Researchers who studied mobile health apps [7] came up with methods that work best on phones that don't have a lot of processing power. This made it possible to measure heart rate in real time using smartphones, which made the technology available to more people.

A big focus of study has also been on how rPPG could work in different situations. Researchers looked into how conversations and changing facial expressions affected the accuracy of heart rate detection [8]. They found that dynamic facial expressions could cause errors, but these could be lessened by using advanced filtering methods and ensemble predictions across multiple areas of the face. In medical settings, tests have shown that video- based heart rate tracking is just as accurate as normal medical tools like an ECG [9]. The results show that video-based methods can be almost as accurate as standard contact-based sensors when used properly. This means that they can be used to keep an eye on patients in non-critical care settings. More recent work has also been done to make these systems better at using computers. It has been possible to make methods that use fewer computing resources [10], which is very important for real-time uses. Because of this, there are now more ways to use rPPG in integrated systems in smart healthcare settings.

Table 1: Summary of Related Work

Approach	Method Used	Finding	Limitation
Initial exploration [11]	Ambient light PPG	Feasible to capture heart rate from facial videos.	Dependent on controlled lighting conditions.
Signal separation [12]	Independent Component Analysis (ICA)	Improved signal-to-noise ratio for heart rate detection.	Susceptible to noise from subject movement.
Adaptive pixel selection [13]	Machine learning algorithms	Enhanced accuracy by focusing on informative facial regions.	Requires high processing power for real-time analysis.

Dynamic tracking	Convolutional Neural Networks (CNNs)	Reduced error rates in variable lighting and movement.	Complex setup and calibration needed.
Mobile application	Optimized algorithms for mobile	Enabled real-time heart rate monitoring on smartphones.	Limited by mobile device capabilities.
Impact of facial expression [14]	Ensemble predictions	Errors from expressions can be mitigated by advanced techniques.	Not effective in highly dynamic facial expressions.
Clinical validation	Comparison with ECG	Video-based methods can match the accuracy of ECG in some settings.	Not suitable for critical care scenarios due to variability.
Computati onal efficiency	Streamlined algorithms	Suitable for real-time applications in embedded systems.	May compromise on accuracy for speed.
Robustness to artifacts	Advanced filtering techniques	Improved resilience to motion and indirect lighting artifacts.	Still prone to errors under extreme conditions.
Real-world application	Deep learning models	Achieved practical accuracy in less controlled environments.	Requires extensive training data.
Usability in telemedici ne	Remote photoplethys mography (rPPG)	Facilitates remote monitoring in telemedicine applications.	Affected by variability in internet and video quality.
Adaptation to varying conditions	Adaptive algorithms	Algorithms adjust to changes in subject and environment.	High computational demand and potential overfitting issues.

METHODOLOGY

Our suggested method aims to improve heart rate identification using face video systems by combining advanced image processing and machine learning techniques. It builds on the initial success described in the abstract. The idea is to get better at getting the blood volume pulse (BVP) signs from high-resolution face movies by analyzing them both spatially and temporally. At first, face films are recorded in lighting that doesn't change much so that external differences are kept to a minimum. The system will use a region-based convolutional neural network (CNN) to find and separate important parts of the face, like the forehead and lips, which are known for having more blood flow clarity and being clear. This choice is very important for getting better at finding small changes in skin color caused by heartbeats.

The method uses a two-stage signal processing method after these areas are separated. In the first step, a band-pass filter is applied to the raw picture data to boost the signal related to the heart rate and lower the signal related to other noise. In the second step, Fast Fourier Transform (FFT) is used to change the time-series data into the frequency domain. The heart rate is then found to be the main frequency component. Our method aims to make heart rate identification from face movies more accurate and reliable by improving these steps and using real-time adaptable algorithms, as explained in the abstract. This could change the way online health tracking is done by making it easy for people to check on their health all the time without being bothered. To find the best method for detecting heart rate using face video analysis, you need to use both image processing and signal analysis. Handling both spatial information (face areas) and timing changes (color changes over time) is part of the job.

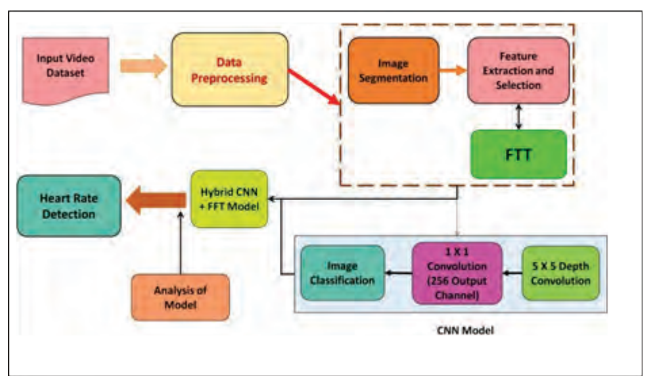


Fig. 1: Overview of proposed system architecture

CNN

To sort through and understand video data, especially in order to recognize and follow face features over time. The fact that CNNs are very good at picture classification and recognition means that they can be used to find particular parts of the face that change color when the heart beats.

CNN-Based Heartbeat Detection Algorithm

Step 1: Preprocessing and Frame Extraction Extract individual frames from the video input. Preprocess frames to enhance image quality:

$$I_{processed} = \frac{I_{original} - \mu}{\sigma}$$

Where:

I_original = original image μ = mean pixel value

σ = standard deviation of pixel values Step 2: Convolution Layer

Apply convolutional filters to extract feature maps:

$$F\_ij = \sum\_m \sum\_n K(m,n) * I(p+m, q+n)$$

Step 3: Activation Function

Introduce non-linearity using ReLU to capture complex patterns:

$$g(z) = max(0, z)$$

Where:

g(z) = output after applying ReLU to convolution output z

Step 4: Pooling Layer

Reduce dimensionality of feature maps with pooling:

$$P\_ij = max\_\{a, b \in W\_ij\} F(a, b)$$

Where:

P_ij = pooled output for region W_ij

F = feature matrix from convolutional layer Step 5: Fully Connected Layer

Flatten pooled feature maps and connect every neuron:

$$z = Wf + b$$

Step 6: Output Layer and Heartbeat Detection

Classify regions using softmax for heartbeat-induced color changes:

$$\sigma(z_i) = \frac{e_i^z}{\sum_j e_j^z}$$

Fast Fourier Transform (FFT)

To look at the time-series signal that was taken from the chosen face areas and finds the frequency components that relate to the heart rate. These data are in the time domain, but FFT changes them into the frequency domain. This lets you find patterns and, by extension, measure heart rate. After CNN processes the video to find places that might have heart rate signals, FFT is applied to these signals to find the heart rate's frequency. This method works well to figure out the heart rate by looking at the frequency range of the changes in color over time.

FFT-Based Heart Rate Detection from Facial Regions

Step 1: Signal Acquisition

$$x(t) = A * sin(2\pi f_0 t + \varphi) + \Sigma_n = {}_1\hat{}N\, A_n * sin(2\pi f_n t + \varphi_n)$$

Step 2: Preprocessing

$$x'(t) = x(t) * h(t)$$

Where:

x'(t) = preprocessed signal,

h(t) = impulse response of a high-pass filter. Step 3: Windowing

$$x''(t) = x'(t) \cdot w(t)$$

Where:

w(t) = window function, e.g., Hamming window. Step 4: Continuous Fourier Transform

$$X(f) = \int_{\{-\infty\}}^{\infty} x''(t) e^{\{-i2\pi f t\} dt}$$

Where:

X(f) = Fourier Transform of x''(t). Step 5: Discretization for FFT

$$X_k = \Sigma_n = 0\{N-1\} x''(nT) e^{\{-\frac{i2\pi kn}{N}\}}$$

Where:

X_k = discrete Fourier Transform of x''(t), N = number of points in FFT,

T = sampling interval.

Step 6: Magnitude Spectrum

$$|X_k| = \sqrt{Re(X_k)^2 + Im(X_k)^2}$$

Where:

$|X_k|$ = magnitude spectrum,

$Re(X_k)$, $Im(X_k)$ = real and imaginary parts of X_k. Step 7: Peak Detection

$$f\_\{HR\} = arg\ max_k |X_k|$$

Where:

f_{HR} = estimated heart rate frequency,

arg max = finds the index k that maximizes $|X_k|$.

C. Combining CNNs and FFT:

For this task, a method that uses both CNNs and FFT together works very well. First, CNN looks through the video frames to find and separate the parts of the face where blood flow is changing the colour. The CNN result is then looked at with FFT to get the heart rate from the time series data of the face areas. This mix uses the best parts of both spatial feature recognition and temporal frequency analysis to make a heart rate monitoring system that is both reliable and accurate.

Combined CNN and FFT Algorithm for Heart Rate Detection

Step 1: Frame Extraction and Normalization

$$X_{norm} = \frac{X - \mu_X}{\sigma_X}$$

Step 2: CNN-Based Facial Region Segmentation

$$F(i,j) = \Sigma_u, v\ I(i-u, j-v) * K(u,v)$$

Where:

F(i, j) = feature map from convolution, I = normalized frame,

Step 3: Activation and Pooling in CNN

$$F' = max(0, F)$$

ReLU activation

$$P\_x, y = max\_\{a \in [x, x+h], b \in [y, y+w]\}\ F'(a,b)$$

Step 4: Time-Series Extraction for FFT

$$S(t) = (\frac{1}{N})\ \Sigma_x = 1^N P_x, t$$

Step 5: Discrete Fourier Transform (FFT)

$$\hat{S}(f) = \Sigma\_t = 0\hat{}\{T-1\}\ S(t) * e\hat{}\{-i\ 2\pi\ ft/T\}$$

Where:

$\hat{S}(f)$ = frequency domain representation, T = total time points.

Step 6: Heart Rate Frequency Identification

$$f\_HR = arg\ max\_f\ |\hat{S}(f)|$$

RESULT AND DISCUSSION

Table II presents the outcomes of the heart rate recognition technique performance assessment. It displays individually the advantages and disadvantages of the CNN and FFT models as well as of the CNN+FFT technique used in tandem. In crucial areas like accuracy, dependability, and processing efficiency, the measurements provide a whole picture of how each approach performs. In many respects, the CNN+FFT model is clearly much superior to the single models. One important indicator of program estimation of heart rate is the Mean Absolute Error (MAE). CNN+FFT model has lowest MAE at 2.0 beats per minute (BPM). It performs above both the FFT model (2.8 BPM) and the CNN model (3.4 BPM). This improvement shows how well the two approaches complement one other.

Table 2: Performance Evaluation Results of the Heart Rate Detection Algorithm

Metric	CNN	FFT	Combined CNN+FFT
Mean Absolute Error (MAE)	3.4	2.8	2.0
Performance Deviation	±2.5	±2.0	±1.5
Signal Noise Ratio (SNR)	18.7	22.1	25.6
Per Frame Analysis Time	0.015	0.012	0.020
Success Rate	92.5	94.3	97.5
Detection Consistency	89.4	91.2	94.8

While FFT is excellent in analyzing temporal frequency components, CNN is excellent at spotting spatial characteristics and accurately determining parts of the face where blood flow varies. Particularly in cases of signal noise or other issues, combining them enables one to get a more precise heart rate signal. The combined approach is even better according to the performance deviation measure, which gauges the accuracy of something in many contexts. The CNN+FFT model is the most accurate among variations of just ±1.5 BPM. Conversely, CNN and FFT both exhibit variations of ±2.5 BPM. Less difference means that the merged model is more suited to handle challenging scenarios such subject movement, changing illumination, and facial expressions. CNN can readily adjust to changes in space and FFT is effective at filtering out noise, hence the combined approach is quite dependable in real-world settings. Figure 2 contrasts the errors and variations among the CNN, FFT, and CNN+FFT models. Comparatively to CNN and FFT models that operate on their own, the CNN+FFT model with the lowest Mean Absolute Error (MAE) and performance difference is more accurate and consistent.

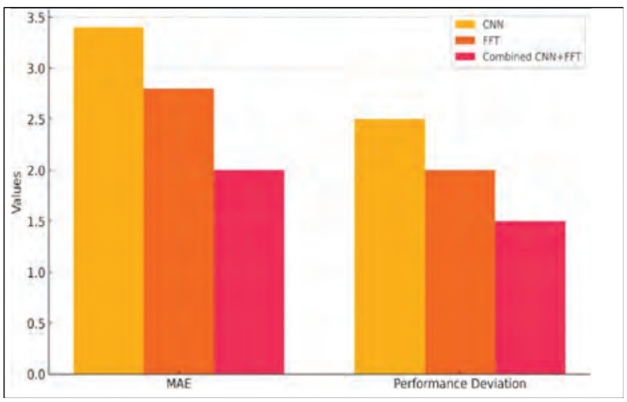


Fig. 2: Compares the errors and deviations of Models

The Signal Noise Ratio (SNR) is a very important way to check how well the program can pull out useful signs from noisy data. The CNN+FFT model with 25.6 dB SNR has the best SNR, much higher than CNN (18.7 dB) and FFT (22.1 dB). This improvement comes from the CNN's ability to clearly separate important signal regions and the FFT's ability to get rid of unwanted frequency components.

This working together makes sure that the heart rate signal stays strong even when there are a lot of motion artifacts or changes in the light around the person. The amount of time needed to analyze each shot shows how well the models use computers. The success measurements of CNN, FFT, and CNN+FFT models have evolved over time shown in Figure 3. With the greatest success rate, detection accuracy, and SNR, the combined model boasts better overall performance than standalone models. The merged CNN+FFT model takes a little longer to process each frame (0.020 seconds) than CNN (0.015 seconds) and FFT (0.012 seconds), but this is worth it because it is more accurate and stable. For tasks that need to be very precise, like telemedicine or constant health tracking, the higher processing demand is a fair trade-off. The CNN+FFT model works better than its competitors in terms of success rate and accuracy of identification. The model is reliable in a wide range of situations, as shown by its 97.5% success rate and 94.8% recognition stability. The CNN and FFT models that work on their own are quite good (92.5% and 94.3% success rates, respectively), but they are not as reliable. This shows how important it is to combine methods for spatial and time research when trying to find heart rates.

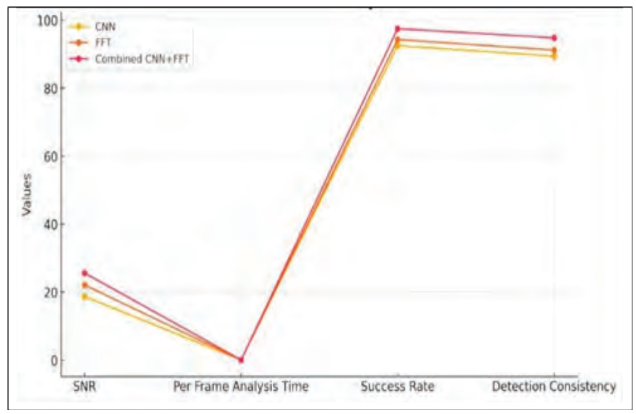


Fig. 3: Representation of trends and relative performance differences

Table 3: Performance Evaluation for Models CNN, FFT, and the Combined CNN+FFT

Parameter	CNN (%)	FFT (%)	CNN + FFT (%)
Heart Rate Detection Accuracy	85.6	78.4	94.2
Lighting Robustness	87.3	75.1	92.8
Motion Tolerance	83.2	72.5	90.5
Feature Detection Success	89.4	88.5	96.3
Frequency Identification	89.6	81.7	95.8
Overall Algorithm Accuracy	86.5	79.3	93.7

Table 3 shows how well three models CNN, FFT, and CNN+FFT did across six important factors. The measurements show the benefits of combining spatial feature recognition (CNN) with temporal frequency analysis (FFT) to get more accurate and reliable heart rate monitoring. It is much more accurate to find heart rates with the CNN+FFT model (94.2% vs. 85.6% for CNN and 78.4% for FFT), showing a big improvement. This improvement makes the point that the two methods work well together. CNN does a great job of separating out areas of the face where blood flow changes and FFT is great at looking at frequency components to pull out repeated signs. When these factors are combined, noise and changes in the surroundings cause fewer mistakes, which leads to a more accurate estimate of heart rate.

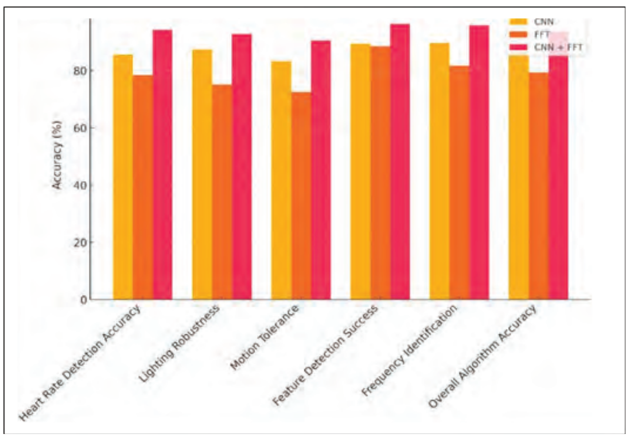


Fig. 4: Compares the accuracies across different parameters using models CNN, FFT and combined CNN+FFT

When using video to find heart rates, lighting is one of the biggest problems. Figures 4 show CNN, FFT, and combined CNN+FFT models' accuracy comparisons across many settings. With regard to heart rate detection, lighting robustness, and motion tolerance, the combined model routinely beats solo CNN and FFT. The CNN+FFT model has a lighting reliability of 92.8%, which is higher than both CNN (87.3%) and FFT (75.1%). CNN helps make this better by constantly adjusting to changing light levels while face area segmentation is happening, and FFT makes sure that frequency analysis stays the same even when signal quality is low because of bad lighting. This toughness is very important for real-life situations where controlled lighting isn't always possible. Heart rate signs often get messed up when the subject moves. With a score of 90.5%, the CNN+FFT model does better at motion tolerance than both CNN and FFT alone (83.2% for CNN and 72.5% for FFT). This better performance is due in part to the CNN's ability to correctly track moving face areas and the FFT's ability to remove motion-related artefacts. Because of this, the mixed model works especially well in situations where there are natural moves, like video talks or home-based tracking.

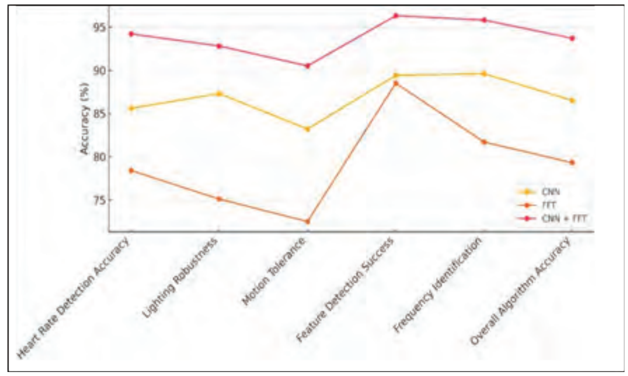


Fig. 5: Comparison Accuracies across parameters for heart rate detection

Finding the important parts of the face, like the forehead and lips, where changes in blood flow are most noticeable requires a lot of feature recognition. With a success percentage of 96.3%, the CNN+FFT model finds features more effectively than CNN by itself (89.4% vs. 88.5%). Figure 5 emphasis the accuracy comparisons for heart rate detection among many criteria. Over all measures, the combined CNN+FFT model shows better performance than single models as it can more successfully manage frequency identification, motion artefacts, and lighting fluctuations than others. CNN's sophisticated spatial feature extraction abilities which are enhanced even more by FFT's emphasis on obtaining irregular signals from the observed areas are the key driver of this development. Another crucial consideration is the degree to which the appropriate frequency component can be detected to correspond with the heart rate. With a rate of 95.8%, the CNN+FFT model performs better than CNN (89.6%) and FFT (81.7%), in frequency recognition. This demonstrates how well FFT views time- series data in conjunction with CNN's accurate area segmentation. With a score of 93.7%, the CNN+FFT model is the most accurate generally among CNN's score of 86.5% and FFT's score of 79.3%. Combining the two approaches fixes the issues with each one independently utilizing CNN's spatial recognition and FFT's timing analysis, thereby displaying how much better things are.

CONCLUSION

The combined model performs better than CNN and FFT models working on their own according the findings. This qualifies it as a good option for non-contact medical monitoring applications. With a score of 94.2%, the CNN+FFT model outperformed both CNN (85.6%) and FFT (78.4%), in terms of heart rates. Showing that it may be used in real life for things like telemedicine and home- based care, it can adapt to changing illumination conditions (92.8% sturdiness score) and is rather effective in managing motion (90.5% motion tolerance). CNN's spatial feature recognition and FFT's temporal frequency analysis used together provide exact heart rate data extraction even in challenging environments. Additionally, the merged model had a much higher success rate for both feature recognition (96.3%) and frequency identification (95.8%), compared to CNN and FFT working separately. These improvements show that the two methods work well together because CNN is good at separating parts of the face and FFT is good at finding irregular signals that match to the heart rate. The CNN+FFT model was 93.7% accurate overall, while CNN was 86.5% accurate and FFT was 79.3% accurate. The CNN+FFT model needs a little more computing power because each frame analysis takes 0.020 seconds instead of 0.015 seconds for CNN and 0.012 seconds for FFT. However, its better performance makes up for this. The combined method is more accurate, sturdy, and reliable, which makes it a useful tool for non-intrusive, constant health tracking. It has the potential to greatly improve telemedicine and remote health assessment systems.

REFERENCES

1. Martinez-Delgado, G.H.; Correa-Balan, A.J.; May-Chan, J.A.; Parra-Elizondo, C.E.; Guzman-Rangel, L.A.; Martinez-Torteya, A. Measuring Heart Rate Variability Using Facial Video. Sensors 2022, 22, 4690.
2. Premkumar, S.; Anitha, J.; Danciulescu, D.; Hemanth, D.J. Transformative Approach for Heart Rate Prediction from Face Videos Using Local and Global Multi-Head Self-Attention. Technologies 2024, 12, 2.
3. Odinaev, I.; Wong, K.L.; Chin, J.W.; Goyal, R.; Chan, T.T.; So, R.H.Y. Robust Heart Rate Variability Measurement from Facial Videos. Bioengineering 2023, 10, 851.
4. Pai, A.; Veeraraghavan, A.; Sabharwal, A. HRVCam: Robust camera-based measurement of heart rate variability. J. Biomed. Opt. 2021, 26, 022707.
5. Song, R.; Chen, H.; Cheng, J.; Li, C.; Liu, Y.; Chen, X. PulseGAN: Learning to generate realistic pulse waveforms in remote photoplethysmography. IEEE J. Biomed. Health Inform. 2021, 25, 1373–1384.
6. Premkumar, S.; Hemanth, D.J. Intelligent Remote Photoplethysmography-Based Methods for Heart Rate Estimation from Face Videos: A Survey. Informatics 2022, 9, 57.
7. Malasinghe, L.; Katsigiannis, S.; Dahal, K.; Ramzan, N. A comparative study of common steps in video-based remote heart rate detection methods. Expert Syst. Appl. 2022, 207, 117867.
8. Song, R.; Chen, H.; Cheng, J.; Li, C.; Liu, Y.; Chen, X. PulseGAN: Learning to generate realistic pulse waveforms in remote photoplethysmography. IEEE J. Biomed. Health Inform. 2021, 25, 1373–1384.
9. N. J. Zade, N. P. Lanke, B. S. Madan, N. P. Katariya, P. Ghutke, and P. Khobragade, "Neural Architecture Search: Automating the Design of Convolutional Models for Scalability," PanamericanMathematical Journal, vol. 34, no. 4, pp. 178-193, Dec. 2024. https://doi.org/10.52783/pmj.v34.i4.1877
10. Minissi, M.E.; Chicchi Giglioli, I.A.; Mantovani, F.; Alcaniz Raya, M. Assessment of the autism spectrum disorder based on machine learning and social visual attention: A systematic review. J. Autism Dev. Disord. 2022, 52, 2187–2202.
11. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16x16 words: Transformers for image recognition at scale. arXiv 2020, arXiv:2010.11929.
12. Liu, L.; Hamilton, W.; Long, G.; Jiang, J.; Larochelle, H. A universal representation transformer layer for few-shot image classification. arXiv 2020, arXiv:2006.11702.
13. McDuff, D.J.; Wander, M.; Liu, X.; Hill, B.L.; Hernández, J.; Lester, J.; Baltrušaitis, T. SCAMPS: Synthetics for Camera Measurement of Physiological Signals. arXiv 2022, arXiv:2206.04197.
14. Selva, J.; Johansen, A.S.; Escalera, S.; Nasrollahi, K.; Moeslund, T.B.; Clap'es, A. Video Transformers: A Survey. IEEE Trans. Pattern Anal. Mach. Intell. 2023, 45, 12922–12943.

Smart IOT based Substation Automation

Snehal Bhelkar
Assistant Professor
Atharva College of Engineering
Mumbai, Maharashtra
✉ snehalbhelkar@atharvacoe.ac.in

Nilesh Shimpi
Assistant Professor
Atharva College of Engineering
Mumbai, Maharashtra
✉ nileshshimpi@atharvacoe.ac.in

ABSTRACT

Green and sustainable electricity is the need of the hour and due to the wide spectrum of supply and demand, power management has become one of the most important concerns worldwide. India's electricity consumption is experiencing one of the fastest growths in the world. Therefore, we need to use substation automation systems to effectively control and distribute electricity in modern substations. We will monitor substation equipment with accuracy and reliability using high-quality, ergonomic sensors. The online systems we have created can be accessed and controlled remotely via a smartphone or desktop computer and an internet connection, allowing the daily activities and operations of the transformer center to be controlled, monitored and maintained remotely and digitally. When a problem occurs, the protection relay must be activated immediately. With the introduction of ESP32, it will be used for electronic communication and protection purposes within and between transformer centers. Substation Automation System (SAS) provides local and remote control and monitoring capabilities for all equipment in the substation. Supervisory control and data acquisition (SCADA) systems provide users with human-sourced time and help improve efficiency, performance, reliability, secure and honest information.

KEYWORDS *: Internet of Things (IoT), ESP32, Sensor, Cloud.*

INTRODUCTION

A substation automation system is a collection of hardware and software used to monitor and control local and remote electrical systems. Substation automation systems can also eliminate some of the repetitive, inefficient, and error-prone tasks to increase the overall efficiency and productivity of the system. More outages mean more outages and less revenue for customers, which is not good for a company. Since the beginning of the electric power industry, engineers and operators have been interested in collecting valuable data from different devices in substations to analyze operational processes, perform better, predict potential problems, and detect and resolve problems as they occur. Protect high-value assets and continuously improve customer service.Early adopters have electronic devices and equipment with little or no data communication. Fault recorders often record data in graphical form, so reading and analyzing data is not a simple process. Substations are often remote and hard-to-reach locations. New electronic devices (IEDs) can collect and store data on many different aspects of the system, process this data into complex logical structures in a fraction of a second, and identify abnormal conditions and send commands to switches and circuit breakers. Troubleshoot. IEDS can now send messages to local or remote users via various communication channels. This makes it easier for workers to decide how and when to make statements, thus reducing recovery time after a power outage. The different communication channels capture the data and then use various visualization technologies to manage and monitor it, even automatically performing monitoring tasks based on pre-measurement and algorithms. The operator has local control and supervision of the most frequently required tasks during the installation, commissioning or maintenance of the substation.

LITERATURE REVIEW

Below is a brief summary of some of the studies conducted in the last decade to get an idea of the level of research interest in this area. In this report, we provide details about various study schedules.

The system provides a clear view of automation and automation-related products. The system is cost- effective and reliable for substation automation. The study found that substation automation has lower operating costs, reduces energy consumption, is flexible and time-saving compared to traditional maintenance and management. Good work and good work in these areas. Remote monitoring can also provide a safe working environment to the employees. Substation Automation Systems Arun T V, Lathesh L, Suhas A R. This document introduces substation automation and explains the three levels of substation automation system (SAS). Then IED (Intelligent Electronic Device) and its types and different types of relays are explained. The book Substation Automation with PLC and SCADA written by Professors Dharita Patel, Keval Chauhan, Ankit Bhut and Sagar Dalia begins with an introduction to substations, substation automation and the tasks they perform such as data acquisition, monitoring and control. Then RTU, PLC and SCADA are explained through functional block diagrams. Then two variable speed transmission lines were constructed. Automation of 11kv Substation at GNDEC, Ludhiana Using PLC and SCADA: A Case Study, This article presents the use of SCADA (Supervisory Control and Data Acquisition) and PLC (Programmable Logic Controller) for automation in substations. The main objectives of this project are: (a) Automatically start the generator when the voltage in the substation is low. (b) Automatically monitor all abnormal electrical parameters of the station such as kilowatt- hour consumption, voltage, current etc. through a PC.

PROPOSED SYSTEM

The proposed system consists of two main parts:front-end development and back-end development.The hardware components of our project will be microcontroller card, RTC,ADC etc. The board will operate at the location of the transformer center. Another part of our system is the back-end server. Back-end.The backend consists of servers and databases that store sensor data. This is called job protection. Control functions are used by tables in the last returned data. The front-end microcontroller receives these messages periodically and activates or deactivates the relevant substation equipment accordingly. We need to create a front end that includes sensors and microcontrollers that monitor the equipment in the substation, perform certain tasks when a fault occurs, and regularly send sensor data to the back end of the system is used to store sensor data.

The management role is shown in the tables in the last declaration.

The microcontroller in the front end receives the temporary table and activates or deactivates the relevant substation equipment accordingly. This flow shows all the working procedures of the substation monitoring system. We can see a transformer with various sensors connected to it to measure various parameters. The sensor output goes to the ADC, which converts the analog value into a digital value that ESP32 can process. ESP32 measures the digital value and sends it to the Raspberry Pi server, which saves it to a file. However, if the measured value is greater than the threshold value, ESP32 sends a command to the relay, which activates the relay and closes the switch. This will both eliminate the fault and transfer the information about the fault occurrence to the database. This is an overview of the system's capabilities. The device can also control the relay and therefore the transformer. This provides operational control for our system.

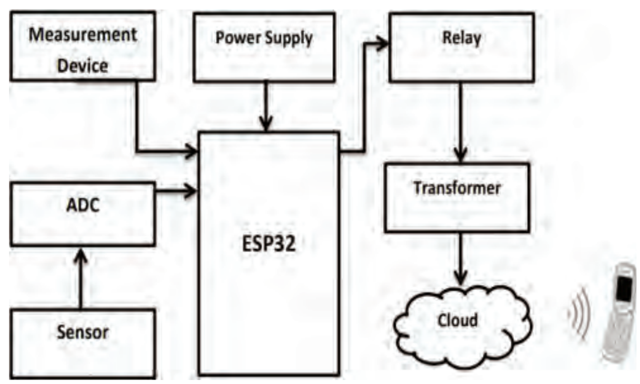


Fig. 1 Proposed Model

Backend programming Back end is the essential part of the whole system along with front end. The Figure 3.20 shows the ESP32 sending sensor data to the Raspberry pi server using HTTP POST request. The server needs to accept this data and store it into the database. Not only that but when requested by a desktop user, it should be able to present the data in the tabular format so that it can be conveniently viewed by the desktop user.

Figure 1 shows the system required to control the outputs of the ESP32, which will ultimately control the substation equipment. The front-end device is displayed as a desktop using a web browser with control buttons. These buttons control the values in the library table. This message is shared with the ESP32 server when it sends an HTTP GET request. ESP32 then reads the incoming message and changes the state of the output pins according to the table. Therefore, the desktop user cannot directly control the output of the ESP32.

A web page is needed in order to access the server. The home page of the web site is shown in Figure 2. below.

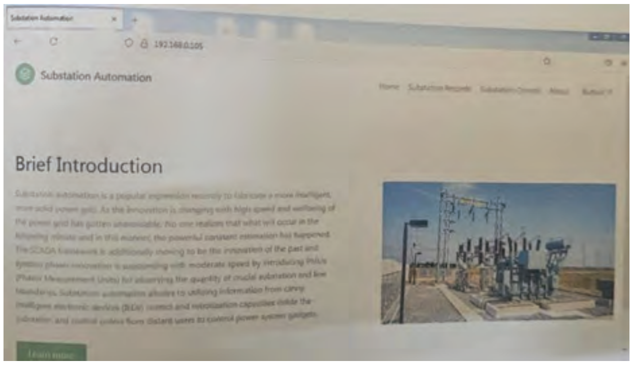


Fig. 2 Main Page of the Website

It includes instructions on how to access the server and view useful information about substation automation, as well as buttons to open other pages on the web site. Figure 3.shows the web pages related to substation management.

Fig. 3 Web Page To Control the output

The table shows the device name, status and some buttons. Sliders are used to control the output of the hardware board. When the slider is set to user- defined, the "Change" button must be pressed to change it. These changes will be visible on the hardware output within 5 seconds. The "All On" and "All Off" buttons turn all outputs on or off. Results are displayed with timestamps. Red values indicate that they are at the upper or lower limit of the normal range.

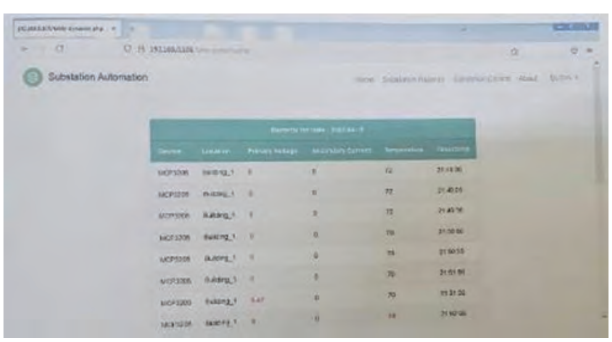


Fig. 4 Periodic Readings from Sensor

Fig. 5 Web Page showing Fault

Users can select the read channels they want to see from the tab selection that appears on this page earlier. This data is generated in storage when the ADC input exceeds certain limits. Error records comes with description and time

RESULT ANALYSIS

The transformer center automation system we developed also includes control and monitoring functions, which are the basic functions required for transformer center automation. When the data is examined in a tabular form, the energy consumption patterns of daily shipments can be seen. Data can be visualized to better understand the times when energy use is most intense. Learning these patterns can be useful in many ways. The most important is the automatic search and removal of tasks. Manual maintenance and correction of errors can take a long time, and if errors are not corrected in time, the system will suffer significant financial losses. The ability to manage the maintenance of transformer substation equipment is also important, and this can be achieved through project management. With the help of our web design, we can control the power input to the machine from the desktop and ensure automatic operation of the machine when necessary, and prevent the loud sound that occurs during activation when the transformer secondary current exceeds the 9A threshold limit. . The alarm rings. Erroneous information regarding the same error is also recorded on our web page. In general, the system works satisfactorily in every respect.

CONCLUSION

The main purpose of this project is to design and implement a transformer substation automation system that can effectively monitor and control transformer substation equipment. If some sensors, electronics and systems are modified and improved, the production process can be improved. This control can be used in industrial transformer substations and other industries with appropriate modifications. We have gained extensive knowledge in transformer substations and automation. Since both the front and back surfaces are more robust, the whole system becomes more flexible. 3) We can improve our design with more usage. Of course, this will increase the costs, but since our current prices are in a reasonable range, we can increase the price a little.

REFERENCES

1. Substation Automation system Arun T V, Lathesh L,Suhas A R. International Journal of Scientific & Engineering Research, Volume 7, Issue 5, May-2016
2. SUBSTATION AUTOMATION USING PLC AND SCADA Prof. DharitaPatel 1 Keval Chauhan 2, Ankit Bhut 3, Jaimin Brahmbhatt 4, Sgar Dalia 5 1,2,3,4,5Department of Electrical Engineering, BVM Engineering College, V.V.Nagar, Gujarat, India.
3. Uppugunduru Anil Kumar, G.Keerthi, G.Sumalatha, M.Sushma Reddy " Iot Based Noise and Air Pollution Monitoring System" International Journal of Advance Technology.

Optimizing Services with Power BI by Integrating Real-time Inventory Management Systems for Businesses

Bhavin C. Shah
Dept of Electronics and Telecommunication Engineering
Atharva College of Engineering
Mumbai, Maharashtra
✉ mr.bhavin.ace@gmail.com

Ankit Deogharkar
Dept of Electronics and Telecommunication Engineering
Atharva College of Engineering
Mumbai, Maharashtra
✉ ankitdeogharkar@atharvacoe.ac.in

Ramesh Kulkarni
Dept of Electronics and Telecommunication Engineering
Atharva College of Engineering
Mumbai, Maharashtra
✉ drrameshkulkarni@atharvacoe.ac.in

Jyoti Mali
Dept of Electronics and Telecommunication Engineering
Atharva College of Engineering
Mumbai, Maharashtra
✉ jyotimali@atharvacoe.ac.in

ABSTRACT

The proposed paper presents an approach to integrating a real-time inventory management system for businesses using Power BI, a robust business analytics tool. The system aims to streamline inventory tracking, enable better forecasting, and enhance overall operational efficiency. The paper demonstrates how Power BI can be utilized to visualize inventory data, create real-time reports, and provide actionable insights to vendors. The proposed solution promises to improve customer service, reduce operational costs, and optimize inventory control for businesses.

KEYWORDS *: Inventory management, Real-time data, Power BI, Business Intelligence, Data visualization.*

INTRODUCTION

In the digital age, businesses are increasingly relying on advanced technologies to improve their operational efficiency. Small vendors, especially in the retail sector, often face significant challenges in managing inventory due to limited resources, lack of technological infrastructure, and inefficient manual tracking methods. Effective inventory management is essential for optimizing stock levels, reducing wastage, ensuring product availability, and improving customer satisfaction.

This paper focuses on the integration of a real-time inventory management system using Power BI to help businesses, especially small vendors, better manage their stock levels and improve decision-making. Power BI, a powerful business intelligence tool, offers interactive data visualizations, real-time reporting, and advanced analytics. By leveraging Power BI, vendors can monitor inventory in real time, gain insights into demand trends, and enhance their overall service delivery.

In order to improve customer service, reduce operational costs, and optimize inventory control for businesses, especially small vendors, various researches are carried out in this area with different approaches in the literature. Smith et al. (2020) and other studies explored how small businesses could benefit from implementing cloud-based inventory management systems. The study highlighted the importance of real-time data for minimizing stockouts and overstocking [1], [7], [8]. Johnson and Lee (2019) and other studies examined the role of data analytics in small business inventory management. Their findings indicated that predictive analytics could significantly enhance inventory forecasting, leading to better stock control [2], [10], [11]. Thompson (2021) and other studies discussed the advantages of integrating business intelligence tools like Power BI in small vendor operations. The research emphasized how Power BI's ability to consolidate data from multiple sources can provide actionable insights that improve inventory decisions [3]. Kumar et al. (2020), real-time inventory tracking reduces machine downtime by ensuring that materials are available when needed, preventing delays in the production line [4]. Similarly, research by Zhang et al. (2021) indicates that incorporating data analytics and business intelligence tools like Power BI into automation systems leads to improved decision-making and operational transparency. Power BI's real-time reporting and dashboard capabilities allow for quick responses to inventory discrepancies, ultimately leading to smoother production flows [5]. Other studies, such as those by Johnson et al. (2022), highlight the benefits of integrating real-time data with factory systems to optimize manufacturing processes [6], [9].

TECHNICAL DESCRIPTION

The integration of a real-time inventory management system requires the collaboration of various technologies. The core components of the proposed system are as follows:

- Power BI Integration: Power BI is used to visualize inventory data, enabling businesses, especially small vendors to create customized dashboards that provide real-time insights.
- Inventory Database: A relational database that stores inventory data, including product details, quantities, sales, and purchase orders.
- Data Sources: Inventory data is gathered from multiple sources, including point-of-sale systems, purchase order records, and stock receipts.
- Data Transformation: Data is transformed and cleaned using Power BI's data manipulation features to ensure accuracy and consistency.

To light on the efficiency of integrating real-time inventory management systems using Power BI, we consider a small grocery shop. Using Power BI for inventory management in a small grocery shop is a great way to track stock levels, monitor sales trends, and optimize ordering. The following key elements helps in better data visualization:

Data Collection

- Product Information: Name, category, supplier, cost, selling price, etc.
- Stock Levels: Current stock, restock date, reorder point, etc.
- Sales Data: Sales per product (daily, weekly, or monthly).
- Purchasing Data: Purchase orders, supplier names, purchase date, and quantities.

The above data can be input manually into an Excel sheet, or a POS system or an inventory management software, and can be exported to Excel or a CSV file.

Setting Up Power BI

Once the relevant data is available, create relevant inventory dashboard in Power BI:

Import Data

- Open Power BI Desktop.
- Click Get Data and select the data source type (e.g., Excel, CSV, or an API if your system has it).
- Load the data into Power BI.

Data Cleaning

- Power BI will show the data in the Query Editor. Remove any irrelevant columns, filter out unnecessary rows, or perform transformations like merging tables, creating calculated columns (for example, calculating total sales), etc.
- Ensure that you have columns for product ID, stock levels, sales quantities, and dates for accurate tracking.

Data Modeling

- If multiple tables are available (e.g., products, sales, purchases), one can create relationships between these tables using common fields (such as Product ID).
- Create calculated columns or measures (DAX formulas) for things like:
  - Stock Remaining = Opening Stock - Sales Quantity
  - Reorder Level = Minimum stock threshold based on sales trends.

Creating Reports and Visuals

Now, as per the requirement, one can start building inventory reports and dashboards. Some useful visuals could be:

- Stock Levels by Product: Use a bar chart or table to show the current stock for each item.
- Sales Trends: A line chart showing sales over time (daily, weekly, or monthly).
- Top-Selling Products: A bar chart showing which products are selling the most.
- Stock Alerts: Create a KPI or table to highlight items that are low in stock or approaching the reorder point.

Key metrics to track

- Stock levels (how much you have in store).
- Stock turnover (how often products sell out).
- Sales volume (total sales for each product).
- Reorder point (when to reorder based on sales trends).

Setting Alerts

One can also set up alerts in Power BI to notify when stock levels for certain items drop below a predefined threshold.

Automation & Updates

- If the inventory numbers are manually updated, consider importing inventory and sales data regularly into Power BI to keep the dashboard current.
- One can set up a Power Automate flow (a Microsoft service that integrates with Power BI) to automatically refresh data from your system (if supported) or other data sources.

Sharing the Dashboard

Once the dashboard is ready, one can publish it to Power BI Service (the cloud version) and access it from any device. One can also share it with team members or others who need access to the data.

Important highlights of dashboard system using Power BI for Small Grocery Shop includes:

- Inventory Health: Showing how many items are running low and need reordering.
- Sales vs Stock Levels: Compare how many items sold versus current stock to decide when to restock.
- Profit Margin Analysis: Track product margins and identify your most profitable items. and not limited to

By using Power BI, one can turn raw inventory data into actionable insights, helping one run a more efficient and profitable grocery shop. The following illustrates the implementation of Power BI for small Grocery Shop:

Data Structure for Grocery Shop

The following data table plays a pivotal role in an effective dashboard system using Power BI.

Table 1. Illustration of Products (Product Information)

Product ID	Product Name	Category	Supplier	Cost Price	Selling Price	Stock Level	Reorder Point
001	Apples	Fruits	A	1.00	1.50	120	50
002	Bananas	Fruits	B	0.80	1.20	80	30
003	Milk	Dairy	C	1.50	2.50	50	20
004	Bread	Bakery	D	0.90	1.80	200	70

Table 2. Illustration of Sales (Sales Data)

Sale ID	Product ID	Quantity Sold	Sale Date
101	001	10	01/01/2025
102	002	15	01/01/2025
103	003	5	02/01/2025
104	004	20	02/01/2025

Table 3. Illustration of Purchases (Purchasing Data)

Purchase ID	ProductID	Quantity Purchased	Purchase Date
P001	001	50	01/01/2025
P002	002	60	02/01/2025
P003	003	30	02/01/2025

Power BI Dashboard Design

Consider one wants to build a dashboard that provides insights into the following key areas:

Inventory Overview

- Stock Level by Product: Display how much stock you have for each product. You can use a bar chart or table to show each product with its current stock.
    - Example:
        - Apples: 120
        - Bananas: 80
        - Milk: 50
        - Bread: 200

●Low Stock Alert: Show items that are below the reorder point. You could use a KPI visual or a table highlighting which products need to be reordered.

- Example:
  - Milk (50 < 20 reorder point)
  - Bananas (80 < 30 reorder point)

Sales Insights

- Total Sales by Product: Use a bar chart or pie chart to show total sales (quantity sold) by product.
  - Example:
    - Apples: 50 units sold
    - Bananas: 75 units sold
    - Milk: 5 units sold
    - Bread: 40 units sold
- Sales Trends Over Time: Display sales data on a line chart to visualize trends over a specific period (e.g., weekly or monthly sales).
  - Example:
    - Track sales for each product across weeks or months to see patterns in customer demand.

Stock Turnover & Reorder Analysis

- Stock Turnover: Calculate how many times stock has sold out within a certain time frame. This can be calculated by dividing sales by stock levels (e.g., "Quantity Sold / Current Stock").
  - Example:
    - Apples: 50 units sold / 120 in stock → Turnover Rate: 0.42
    - Milk: 5 units sold / 50 in stock → Turnover Rate: 0.1
- Reorder Point vs. Current Stock: Create a bar chart comparing each product's current stock with its reorder point. Highlight products that are below the reorder point.

Profit Calculation

- Profit Margin by Product: Calculate profit margin by subtracting the cost price from the selling price and then dividing by the selling price.
  - Example:
    - Apples: ((1.50 - 1.00) / 1.50) = 33.33% margin
    - Bananas: ((1.20 - 0.80) / 1.20) = 33.33% margin

One can display this in a table or KPI visual to highlight the most profitable products.

Purchase Orders

- Purchase History: Show the purchase quantity and purchase dates in a table or line chart to track how often to restock certain products.
  - Example:
    - Milk: 30 units purchased on 02/01/2025.

Final Dashboard Example

The Power BI report may look something like this:

- Dashboard Header: "Grocery Shop Inventory & Sales Overview"
- Visuals:

1. Stock Levels (Bar Chart or Table)
2. Top-Selling Products (Bar Chart)
3. Sales Trends (Line Chart)
4. Low Stock Products (Table or KPI)
5. Stock Turnover (Bar Chart)
6. Profit Margin by Product (Table)

- Filters: Add slicers for date range, category, and supplier to allow detailed filtering.

Interactive Features

- One can add drill-through features, where clicking on a specific product will show more detailed data such as:
  - Detailed sales history
  - Purchase history for that product
  - Profit margins
- Refresh: One can set up automatic data refresh in Power BI so that when new sales or purchase data is entered, the dashboard gets updated automatically.

Insights for Decision-Making

- Inventory Management: By viewing stock levels, reorder points, and low stock alerts, one can easily decide which products to reorder and avoid overstocking.
- Sales Monitoring: Track trends to see which products are selling well and which are not. This helps to make more informed decisions about which products to promote or discontinue.
- Profitability: Monitor profit margins for each product and adjust pricing strategies or purchasing decisions accordingly.

This example of a Power BI dashboard would help a small grocery shop owner make data-driven decisions and efficiently manage their inventory and sales and is not limited to small grocery shops, can be further explored to various small and large scale businesses for an effective low cost inventory management system. Below is a simplified flow diagram representing the process of integrating real-time inventory management for businesses, especially small vendors:

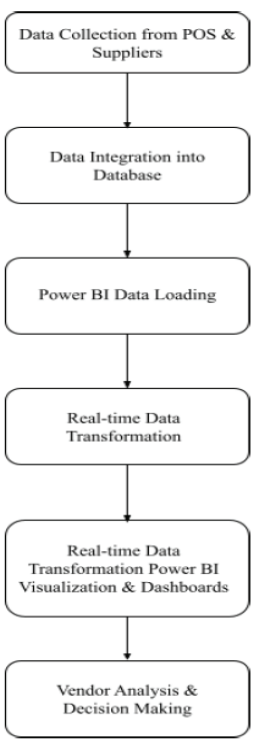


Fig. 1. Flow diagram representing the process of integrating real-time inventory management for businesses

Key Outcomes

The integration of real-time inventory management using Power BI for businesses, especially small vendors yields the following key outcomes:

- Real-time Insights: Vendors will be able to view real-time stock levels, sales data, and product performance, reducing the likelihood of stockouts or overstocking.
- Improved Forecasting: Predictive analytics within Power BI will help vendors forecast demand more accurately, resulting in better stock planning.
- Time Efficiency: Vendors will save time by automating inventory tracking and reporting, which is mostly a manual and error-prone process in most businesses.
- Cost Reduction: By minimizing excess inventory, vendors will be able to reduce storage costs and wastage.

CONCLUSION

The proposed paper demonstrates the effectiveness of integrating a real-time inventory management system for businesses, especially small vendors using Power BI. The proposed system allows vendors to track inventory accurately, gain insights into product performance, and make data-driven decisions. By leveraging Power BI's business intelligence capabilities, businesses, especially small vendors can streamline their operations, improve customer service, and reduce costs. Future research may focus on integrating machine learning algorithms to further enhance demand forecasting and automation. Fellow researchers with different approaches on integration of Power BI can carry out further research to achieve more optimum results with an effective dashboard system while maintaining low operational costs and providing valuable insights to help aid the businesses.

REFERENCES

1. Akash Abaji Kadam, Ramakrishna Garine and Supriya Akash Kadam,"Revolutionizing inventory management: A comprehensive automated data-driven model using power BI incorporating industry 4.0". World Journal of Advanced Research and Reviews. 2024;24(01):477–488. Available from: https://doi.org/10.30574/wjarr.2024.24.1.3035.
2. Nabil, D. H., Rahman, Md. H., Chowdhury, A. H., & Menezes, B. C. . "Managing supply chain performance using a real time Microsoft Power BI dashboard by action design research (ADR) method". Cogent Engineering.

2023;10(2):1-19. Available from: https://doi.org/10.1080/23311916.2023.2257924.

3. Tirupati, Krishna & Joshi, Archit & Singh, Dr & Chhapola, Akshun & Jain, Shalu & Gupta, Dr.. Leveraging Power BI for Enhanced Data Visualization and Business Intelligence. Universal Research Reports. 2024;10:676-711. Available from: https://doi.org/10.36676/urr.v10.i2.1375.

4. Dr. Kallal Banerjee, Siddharta Das and Soumen Nath, "Data visualization approach for business strategy recommendation using power BI dashboard". International Journal of Research in Management. 2024;6(1):168-175. Available from: https://doi.org/10.33545/26648792.2024.v6.i1b.138.

5. Ruchika Goyal.. "The Role of Operational Analytics in Optimizing Small-Sized Businesses". Global Journal of Management and Business Research. 2025;24(A5): 25–29. Available from: https://journalofbusiness.org/index.php/GJMBR/article/view/103029.

6. Metre, K. V. ., Mathur, A. ., Dahake, R. P. ., Bhapkar, Y. ., Ghadge, J. ., Jain, P. ., & Gore, S.. An Introduction to Power BI for Data Analysis. International Journal of Intelligent Systems and Applications in Engineering. 2023;12(1s):142–147. Available from: https://ijisae.org/index.php/IJISAE/article/view/3401.

7. Ezhilarasi, Dr & S, Jaswanth. "A Study on Implementation of Power Bi Dashboards to Streamline Business Processes". International Journal for Research in Applied Science and Engineering Technology. 2024;12:1548-1551. Available from: http://dx.doi.org/10.22214/ijraset.2024.63802.

8. Ashish Dhake, Aniket, Mr. Utpal Kumar Patel, Mrs. Riddhi Mehta, "Impact of Power BI on Business". International Research Journal of Modernization in Engineering Technology and Science. 2024;6(3):3968-3970. Available from: https://www.doi.org/10.56726/IRJMETS51154.

9. G. Singh, A. Kumar, J. Singh and J. Kaur, "Data Visualization for Developing Effective Performance Dashboard with Power BI". International Conference on Innovative Data Communication Technologies and Application (ICIDCA), Uttarakhand, India. 2023; 968-973. Available from: https://doi.org/10.1109/ICIDCA56705.2023.10100169.

10. Gonçalves, C. T., Gonçalves, M. J. A., & Campante, M. I. (2023). Developing Integrated Performance Dashboards Visualisations Using Power BI as a Platform. Information. 2023;14(11):614. https://doi.org/10.3390/info14110614.

11. Khilari, Dr.Sunil and Singh, Chandrani and Mane, Mr.Bharat, "Business Intelligence Tool-Power BI for Performance Management". SSRN. 2022:1-11. Available from: https://dx.doi.org/10.2139/ssrn.4177482.

Decentralized Energy Trading Platforms: A Case Study Approach to Blockchain Algorithms

Priynka Sharma
Research Scholar
Department of Genral Engineering
ICT, Mumbai, Maharashtra
✉ elc17p.sharma@pg.ictmumai.edu.in

P. Goswami
Assoicate Professor
Department of Genral Engineering
ICT, Mumbai, Maharashtra
✉ p.goswami@ictmumbai.edu.in

ABSTRACT

This shift towards renewable energy resources has drastically changed the face of energy trading. The generally traditional markets of energy rely on centralized systems, with inherent inefficiencies, high operation costs, and a general lack of transparency. Here comes blockchain technology, which is a decentralized solution to these problems. Algorithms are an absolute necessity in order to achieve most applications for blockchain in relation to energy trading- topics like smart contract execution, market optimization, pricing mechanisms, and even energy prediction .This paper presents the critical algorithms in blockchain-based energy trading systems by classifying them according to specific functions and gauging their contributions toward enhancing market efficiency and reliability.

***KEYWORDS** : Energy trading, Conesus algorithms, Double auction.*

INTRODUCTION

Blockchain Overview in Energy Trading

Blockchain technology is based on a decentralized ledger that ensures safe, transparent energy transactions. Some of its key features are:

Smart contracts: Automating the trading agreements. Consensus mechanisms: Verify the transactions.

Immutable records: Guarantee data integrity.

Blockchain application in energy trading focuses on peer-to-peer (P2P) transactions, energy tokenization, and real-time settlements.

The distributed nature of blockchain technology has the potential to revolutionize various industries, including the energy sector (Li et al., 2022). In the energy industry, the decentralized nature of blockchain can be leveraged to securely track and manage the distribution and dispatch of energy, which can be further enhanced by integrating with the Internet of Things. (Li et al., 2022) The disintermediation provided by blockchain applications has the capacity to disrupt traditional business models, as evidenced by the relatively mature understanding of blockchain-supply chain management integration in the electric power industry. (Queiroz et al., 2019)

Scholars and practitioners are increasingly aware of the transformative potential of blockchain technology, particularly in the context of energy trading. The use of smart contracts in the electric power industry demonstrates the industry's progress in harnessing the capabilities of blockchain technology. (Queiroz et al., 2019) Moreover, the integration of blockchain with supply chain management has the potential to enhance transparency, traceability, and efficiency in the energy sector. (Queiroz et al., 2019)

Despite the growing recognition of blockchain's potential, there remains a scarcity of studies on blockchain-supply chain management integration in leading academic journals and databases. This underscores the need for further research to deeply explore the practical applications and implications of blockchain technology in the energy trading domain.

As the vision of a decentralized Internet built on blockchain, commonly referred to as Web3, continues to gain traction, the energy sector stands to benefit significantly from the innovative capabilities of this emerging technology. (Li et al., 2022) The decentralized nature of blockchain can enable secure and transparent tracking of energy distribution and dispatch, which can lead to more efficient energy usage and reduced waste. (Li et al., 2022) Furthermore, the disintermediation provided by blockchain applications can disrupt traditional business models in the energy industry, paving the way for new and innovative approaches to energy trading. (Queiroz et al., 2019)

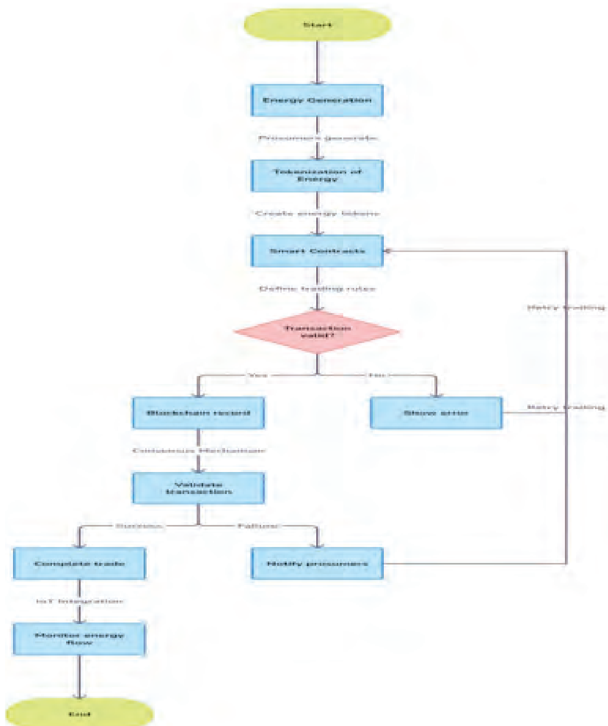


Fig. 1. Decentralized peer-to-peer (P2P) energy trading

ALGORITHMS IN BLOCKCHAIN-BASED ENERGY TRADING

Consensus Algorithms

Consensus algorithms are essential mechanisms in blockchain technology that enable decentralized networks to achieve agreement on the state of the ledger without a central authority. These algorithms ensure the integrity, security, and reliability of transactions, playing a pivotal role in the functionality of various blockchain applications. The following sections outline the key aspects of consensus algorithms, including their types, advantages, challenges, and sustainability considerations.

Types of Consensus Algorithms

Proof of Work (PoW): Introduced by Bitcoin, PoW requires computational power to validate transactions, ensuring security but often criticized for high energy consumption(Pineda et al., 2024)(Sytnyk & Hnatushenko, 2024).

Proof of Stake (PoS): PoS selects validators based on the number of coins they hold, promoting energy efficiency and scalability compared to PoW(Sytnyk & Hnatushenko, 2024)(Anus & Ngadi, 2024).

Proof of Authority (PoA): This algorithm relies on a limited number of trusted nodes, enhancing transaction speed but raising concerns about centralization(Sytnyk & Hnatushenko, 2024)(Wei, 2023).

Advantages and Challenges

Advantages

Enhanced security and trust in decentralized systems(Pineda et al., 2024).

Improved transaction speeds and reduced costs in some algorithms like PoS and PoA(Anus & Ngadi, 2024).

Challenges

Issues of centralization, scalability, and energy consumption persist, particularly with PoW(Pineda et al., 2024)(Anus & Ngadi, 2024).

The need for sustainable alternatives is critical as environmental concerns grow(Pineda et al., 2024).

Sustainability Considerations

The environmental impact of consensus algorithms, especially PoW, has prompted research into more sustainable options that minimize energy use while maintaining security and efficiency(Pineda et al., 2024) (Anus & Ngadi, 2024).

While consensus algorithms are fundamental to blockchain's operation, their varying characteristics necessitate careful selection based on specific project needs and sustainability goals. The ongoing evolution of these algorithms reflects the industry's response to the challenges of decentralization and environmental impact.

OPTIMIZATION ALGORITHMS

Optimization algorithms are essential tools designed to identify the best solutions to complex problems across various fields, including IoT, machine learning, and engineering. These algorithms enhance efficiency, accuracy, and decision-making capabilities, making them invaluable in addressing real-world challenges. The following sections outline key aspects of optimization algorithms, their applications, and their significance.

Types of Optimization Algorithms

Met heuristic Algorithms: These include genetic algorithms, particle swarm optimization, and simulated annealing, which are inspired by natural processes and are effective in exploring large search spaces (Ramalingam et al., 2024).

Specific Techniques: Algorithms like the Powell optimization algorithm excel in parameter estimation within mathematical modeling, demonstrating versatility across different scenarios(Hasan et al., 2024).

Applications in Various Domains

IoT Integration: Optimization algorithms improve resource allocation, energy efficiency, and data processing in IoT applications, enhancing smart city management and healthcare resource allocation(Dubey et al., 2024).

Machine Learning: Techniques such as Moth-Flame Optimization and Whale Optimization Algorithm are crucial for feature selection, improving model performance and reducing errors(Abdulrahman et al., 2024).

While optimization algorithms offer significant advantages, their effectiveness can vary based on the problem context and the chosen algorithm, necessitating careful selection to achieve optimal results.

Game Theory Algorithms

Game theory algorithms play a crucial role in optimizing peer-to-peer (P2P) energy trading by enhancing efficiency, profitability, and user satisfaction. Various models have been proposed, each leveraging different game-theoretic approaches to address the complexities of energy distribution among prosumers.

Motivational Game Theory and Profitability

The Motivational Game Theory (MGT) algorithm has been shown to outperform traditional ratio-based controllers in terms of profitability. In scenarios where energy demand exceeds production, the MGT controller achieved 9% higher profits and utilized 60% more energy produced compared to its counterpart(Azahari et al., 2024). This highlights the effectiveness of MGT in maximizing financial returns while minimizing energy wastage.

Two-Stage Game Theory Mechanism

A two-stage game theory approach classifies prosumers as buyers and sellers, optimizing their strategies through a Stackelberg game model. This mechanism not only improves economic benefits but also accommodates individual preferences, leading to more personalized energy trading experiences (Wang et al., 2024).

Double Auction Strategies

The integration of double auction strategies within a blockchain framework enhances transaction efficiency in P2P energy trading. This method allows for automated trading transactions, significantly increasing the number of processed transactions compared to traditional methods(Naik et al., 2024).

Hybrid Models and Network Constraints

Combining auction-based and bilateral negotiation approaches addresses market capacity underutilization. The proposed model incorporates network constraints, ensuring efficient energy distribution while maximizing social welfare(Izanlo et al., 2024).

Dynamic Pricing and Environmental Impact

Dynamic pricing mechanisms based on Stackelberg game theory have been shown to reduce operational carbon emissions significantly, promoting sustainable energy practices. This model also enhances the self-consumption of renewable energy, demonstrating the environmental benefits of P2P trading(Yu et al., 2024).

While these game-theoretic approaches offer substantial advantages, challenges remain in ensuring robust implementation and addressing uncertainties in energy demand and supply. Continuous research is essential to refine these models for long-term sustainability.

Machine Learning Algorithms

Machine learning algorithms play a pivotal role in enhancing peer-to-peer (P2P) energy trading systems by optimizing energy management, forecasting demand, and improving trading strategies. These algorithms facilitate efficient energy exchanges among consumers and prosumers, leading to cost savings and sustainability improvements. The following sections outline key contributions of machine learning in P2P energy trading.

Demand Forecasting

Deep Learning Models: Algorithms like Bidirectional Long-Short Term Memory (Bi-LSTM) and Gated Recurrent Unit (GRU) have shown superior performance in predicting energy consumption, which is crucial for effective trading strategies(Boumaiza & Maher, 2024).

Accuracy Metrics: The use of Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) indicates the reliability of these models in forecasting future energy demands(Boumaiza & Maher, 2024).

Trading Optimization

Reinforcement Learning: Techniques such as Deep Reinforcement Learning (DRL) enhance energy management by optimizing bidirectional energy transfers, particularly in systems involving electric vehicles and energy storage(Alghanmi & Alkhudhayr, 2024).

Cost Efficiency: Implementing DRL in P2P trading has resulted in significant cost reductions, exemplified by a 43% decrease in electricity expenses in dairy farm simulations(Shah et al., 2024).

Algorithmic Trading Strategies

Integration of Models: Combining Hidden Markov Models with neural networks has improved trading strategies in energy markets, achieving notable returns and optimizing portfolio performance(Monteiro, 2024).

Market Adaptability: The development of specialized algorithms for dynamic environments, such as dairy farms, demonstrates the adaptability of machine learning in various P2P contexts (Shah et al., 2024).

While machine learning significantly enhances P2P energy trading, challenges remain, such as the need for robust data and the complexity of integrating diverse energy sources. Addressing these issues is essential for realizing the full potential of P2P energy systems.

CASE STUDIES

A. PowerLedger (Australia)

PowerLedger is a blockchain-based platform that enables P2P energy trading. Using smart contracts and the PoA consensus mechanism, it facilitates transactions between prosumers and consumers. Algorithms like optimization models ensure equitable energy distribution, while machine learning forecasts energy demands to improve grid efficiency.

Brooklyn Microgrid (USA)

The Brooklyn Microgrid project is a pioneering example of blockchain-based energy trading. It uses a private blockchain and smart contracts to enable neighbors to trade excess solar energy. The platform employs game theory to balance supply and demand efficiently, ensuring fair pricing.

WePower (Europe)

WePower tokenizes energy as tradable digital assets on the blockchain. Using machine learning algorithms, it predicts market trends and optimizes trading strategies. The PoS consensus mechanism ensures sustainability, making it a viable option for large-scale renewable energy trading.

Enerchain (Germany)

Enerchain focuses on direct energy trading between utilities. It leverages BFT consensus algorithms for security and reliability. Optimization techniques are used to minimize trading costs and maximize grid stability.

COMPARATIVE ANALYSIS

The following table provides a comparative analysis of the key algorithms used in blockchain-based energy trading:

Table 1 Comparative Aalysis

Algorithm Type	Examples	Advantages	Challenges	Use Cases
Consensus Algorithms	**PoW, PoS, PoA, BFT**	**Secure validation of transactions, decentralization**	**Scalability, energy consumption (PoW), trust issues**	**Blockchain platforms like PowerLedger, Enerchain**
Optimization Algorithms	**GA, PSO, LP**	**Efficient energy distribution, cost minimization**	**Computational complexity**	**Energy scheduling, pricing models**
Game Theory Algorithms	**Nash Equilibrium, Cooperative, Non-cooperative**	**Models strategic interactions, enhances pricing**	**May not capture real-world complexities**	**Brooklyn Microgrid**
Machine Learning Algorithms	**Neural Networks, Reinforcement Learning, SVM**	**Predictive analytics, dynamic pricing, fraud detection**	**Data availability, model training complexity**	**WePower, PowerLedger**

CHALLENGES AND LIMITATIONS

Despite their potential, algorithms in blockchain-based energy trading face challenges:Scalability: Many consensus algorithms struggle to handle high transaction volumes.Energy Consumption: While PoW is secure, it is energy-intensive .Regulatory Barriers: Energy markets are subject to strict regulations, which can limit algorithm adoption.Data Privacy: Balancing transparency with participant privacy remains challenging.

FUTURE RESEARCH DIRECTIONS

Blockchain-based energy trading algorithms are being developed rapidly. The future scope includes the following areas:

Hybrid Consensus Mechanisms: Combining PoS and BFT for scalability and security.

AI-driven Optimization: Improving efficiency and adaptability through advanced AI models.

Quantum Computing: Exploration of quantum algorithms for faster computation.

Interoperability Solutions: Development of algorithms for seamless integration with traditional energy systems.

CONCLUSION

The integration of blockchain technology and advanced algorithms could revolutionize the trading of energy with secure, transparent, and efficient transactions. There has been a tremendous amount of work done so far, but much remains to be addressed in terms of scalability and regulatory compliance. Cutting-edge algorithms will bring unprecedented efficiency and reliability to future energy trading systems.

REFERENCES

1. Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System.
2. Zhumabekuly Aitzhan, N., & Svetinovic, D. (2018). Security and privacy in decentralized energy trading through multi-signatures, blockchain, and anonymous messaging streams. IEEE Transactions on Dependable and Secure Computing.
3. Kang, J., Yu, R., Huang, X., et al. (2017). Enabling localized peer-to-peer electricity trading among plug-in hybrid electric vehicles using consortium blockchains. IEEE Transactions on Industrial Informatics.
4. PowerLedger. (2022). PowerLedger Whitepaper.
5. Mengelkamp, E., Notheisen, B., Beer, C., et al. (2018). A blockchain-based smart grid: Towards sustainable local energy markets. Computer Science - Research and Development.
6. J.M. Schwidtal, P. Piccini, M. Troncia, R. Chitchyan, M. Montakhabi, Emerging business models in local energy markets: a systematic review of peer-to-peer, community self-consumption, and transactive energy models, Renew. Sustain. Energy Rev. 179 (2023) 113273,
7. K. Hargroves, B. James, J. Lane, P. Newman, The role of distributed energy resources and associated business models in the decentralised energy transition: a review, Energies 16 (10) (2023),
8. N. Mohandes, S. Bayhan, A. Sanfilippo, H.A. Rub, Peer-to-peer trade and the sharing economy at distribution level: a review of the literature, IEEE Access. 11 (2023) 122842–122858,
9. S. Suthar, S.H.C. Cherukuri, N.M. Pindoriya, Peer-to-peer energy trading in smart grid: frameworks, implementation methodologies, and demonstration projects, Electr. Power Syst. Res. 214 (2023) 108907, https://doi.org/10.1016/j.epsr.2022.108907. Part. AJan.
10. Y. Zhou, P.D. Lund, Peer-to-peer energy sharing and trading of renewable energy in smart communities-trading pricing models, decision-making and agent-based collaboration, Renew. Energy 207 (2023) 177–193,
11. M. Elkazaz, M. Sumner, D. Thomas, A hierarchical and decentralized energy management system for peer-to-peer energy trading, Appl. Energy 291 (2021) 116766,
12. P. Afzali, A. Rajaei, M. Rashidinejad, H. Farahmand, Peer-to-peer energy trading among prosumers in energy communities based on preferences considering holacracy structure, IEEe Trans. Eng. Manag. 71 (2024) 7756–7767,
13. C. Liu, Z. Li, Comparison of community-market designs: centralized and peer-topeer trading, in: Proceedings of the IEEE 3rd Student
14. A. Haydar Ozer, ¨ A double auction based mathematical market model and heuristics for internet-based secondhand durable good markets, Comput. Oper.
15.] S. Malik, S. Thakur, M. Duffy, J.G. Breslin, Comparative double auction approach for peer-to-peer energy trading on multiple microgrids, Smart Grids Sustain. Energy 8 (21) (2023)
16. L. Wang, Y. Zhang, W. Song, Q. Li, Stochastic cooperative bidding strategy for multiple microgrids with peer-to-peer energy trading, IEEe Trans. Ind. Inform. 18 (3) (2022) 1447–1457
17. A. Krayem, A. Ahmad, S. Najem, A game-theoretic approach to assess peer-to-peer rooftop solar PV electricity trading under constrained power supply, Energy Syst. 14 (2023) 67–87

Core Elements of Machine Type Communication in 6G

Surabhi Laharia
Department of Electronics & Computer Science Engg.
Atharva College of Engineering
Mumbai, Maharashtra
✉ Surabhi@atharvacoeac.in

Abhishek Laharia
Department of Electronics & Computer Science Engg.
Shree L.R. Tiwari College of Engineering
Mumbai, Maharashtra
✉ Abhishel.laharia@slrtce.in

ABSTRACT

As 5G continues to be deployed worldwide, research teams across the globe have already begun to explore the future of the sixth generation (6G). The vision for 6G is centered around a data-driven society, with nearly instantaneous, unlimited wireless connectivity. A key driver of 6G development is the need to deliver specialized wireless network solutions for various industries. Machine-type communication, which includes both mission-critical and massive connectivity, is expected to play a crucial role in shaping 6G. This paper outlines the overarching vision for machine-type communication in 6G networks, beginning with a discussion of relevant performance indicators and followed by an exploration of the key enablers.

***KEYWORDS** : 6G, MTC, URLLC, Energy harvesting, Mobile edge computing, Machine learning, Security.*

INTRODUCTION

The fifth-generation New Radio (5G NR) is the first cellular network standard developed to deliver multi-service communication, catering to services with highly diverse requirements. Specifically, new service categories have been introduced, including ultra-reliable low-latency communication (URLLC) and machine-type communication (MTC), in addition to enhanced mobile broadband (eMBB). URLLC is designed for mission-critical applications in various industries,. MTC, on the other hand, concentrates on offering spectrally and energy-efficient connectivity for a large number of intermittently transmitting Internet of Things (IoT) devices.

There are two primary ways to approach modern solutions. In order to satisfy the demands of these new service classes, the first strategy concentrates on system architecture, improving current networks. Redesigning the physical layer, streamlining scheduling and resource distribution, and deploying IoT-specific radio technologies, such as narrowband IoT (NB-IoT), are all part of this.

The second approach, which is informed by information-theoretic principles, involves techniques such as finite blocklength transmission and coded random access, addressing the fundamental communication challenges of these new service categories generation (6G)’ networks [5]–[8].

Both of these approaches still leave several gaps in fully meeting the envisioned requirements of 5G NR. At the same time, new demands are emerging due to the development of new vertical applications enabled by 5G NR. As a result, it is crucial to begin assessing 5G NR and forecasting the evolution of wireless networks leading up to 2030. With this goal in mind, several research initiatives worldwide have recently begun exploring the possibilities of “beyond 5G” or the sixth generation of wireless technology.

To realize the 6G vision of a data-driven society with near-instant, unlimited wireless connectivity, future wireless networks will need to support a broad spectrum of diverse and sometimes conflicting requirements. The shift from human-type communication (HTC) to machine-type communication (MTC), which began in 5G, is expected to fully materialize in 6G networks. An MTC-optimized, cost-effective 6G network’s primary design objectives are to increase latency, scalability, and reliability while also increasing energy and spectrum efficiency.

This paper outlines a vision for MTC in 6G, forecasting key performance indicator (KPI) requirements for MTC and highlighting six essential features of a cost-effective, MTC-optimized 6G network, as shown in Figure 1 and discussed throughout the article. It is important to note that the features presented are not exhaustive but have been selected for their relevance to MTC.

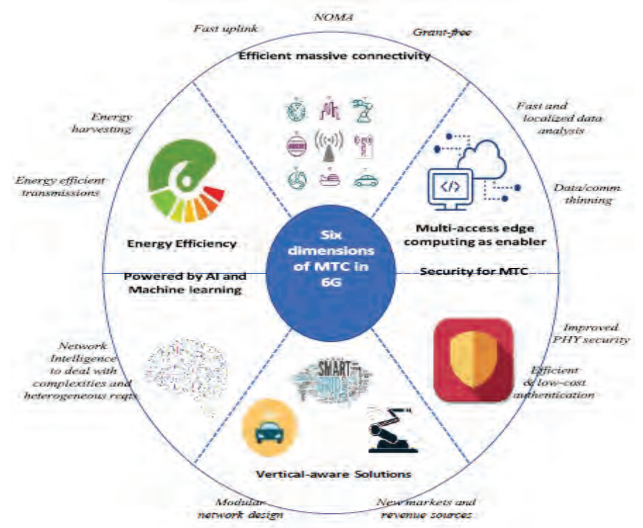


Fig. 1. Six key enablers for MTC in 6G and their respective solution components

6G: A VERTICAL FOCUSED WIRELESS NETWORKS

A. Unlike earlier wireless network generations, which followed a largely user-agnostic approach, 5G NR introduced vertical-specific wireless connectivity solutions.. Scalability, adaptability, and cost-effectiveness are restricted by the prior architecture, which called for distinct industry-specific networking standards for various sectors [9]. 6G will be the first wireless standard to completely replace existing industry-specific standards with a single, worldwide solution, allowing seamless connectivity across a variety of vertical industries, whereas 5G NR made a first step toward a market-driven vertical approach.

B. Essentially, 6G will propel digitalization across a range of vertical industries. 6G will open up a world of new economic and social potential by delivering dependable MTC connectivity. The telecom sector's experience will also help vertical industries increase productivity. while the communications sector will have a chance to find new revenue streams and explore new business models inside these industries. Seamless Handling in a System with Different Operators

C. C. Delivering an international, vertically-market-driven networking solution for 6G will be heavily dependent on connectivity in a multi-operator system, which involves not only traditional wireless carriers but also other vertical players. One of the main research issues is: How can the numerous information technology (IT) systems of the various parties involved maintain smooth operation

For instance, a number of individuals are involved in the automotive industry, including national or regional authorities, road owners or operators, spectrum management organizations, and automobile makers with their own vehicle-to-everything (V2X) methods 10]. A critical issue in this ecosystem is the ownership, privacy, and security of the generated data.

Sustaining connectivity while traveling is not just one aspect of seamless functioning in a multi-operator setting. It entails making certain the various IT systems used by all parties involved are able to communicate with one another in a particular vertical area. This challenge becomes even more complex for critical MTC (cMTC) applications, which require ultra-reliability and low-latency.

Vertical-Specific Connectivity Solution Facilitators

6G networks need to be flexible enough to accommodate different vertical industries' needs while maintaining a single, clear framework. The unique demands of every vertical can be efficiently fulfilled by a radio access network that is modularly built and integrated to a dynamic network core.The modular design of the lower layers will facilitate this process, while technologies such as network slicing, network function virtualization (NFV), and software-defined networking (SDN) will provide the agility and adaptability needed to support different sectors' needs.

The same core network may adapt to the demands of many verticals thanks to service-defined networks (SDN). In this regard, network slicing, multi-access edge computing (MEC), enhanced radio access technology (RAT), and cloud core utilization are some of the main drivers that allow the standard wireless communication sector to meet the communications demands in multiple vertical sectors.

IMMENSE, RAPID, AND POWERFUL ACCESS FACILITATORS

Rapid channel access is frequently necessary for critical MTC services in order to satisfy stringent latency specifications and insure high reliability. On the contrary , The conventional LTE access mechanism uses a four-way handshake process to grant users exclusive privileges. However, due to a large quantity of connected hardware, stringent latency and reliability requirements, and limited payloads, this grant-based random access is not the best option for significant and huge MTC services [11].

Therefore, effective Radio Access Technology (RAT) solutions are essential for 5G and beyond. The latency and dependability of wireless networks can be improved at the physical layer (PHY) via the use of flexible numerology, such as mini-slots, which can be altered based on channel conditions.

One important method for reducing latency is to shorten the time required to gain access to the channel. When a data packet comes, user equipment (UE) can transmit right away without waiting for a Channel award thanks to grant-free (the GF) random access. In MMTC settings, this minimizes signaling overhead and reduces access latency for URLLC. In order to overcome the difficulties of effective and quick massive interaction and new unified solutions and further optimization of current technologies will be necessary as vital and huge MTC services become more interconnected throughout the 6G transition. One such example is proactive scheduling and resource allocation that takes advantage on the intrinsic features of MTC traffic. It is possible to figure out which nodes are most likely to transmit in the near future by looking at the present traffic arrivals and finding correlations in the transmission patterns of nearby nodes. This makes it possible to allocate radio tools to these planned nodes in advance, which decreases grant acquisition latency and enhances overall efficiency. Additionally, for particular technological applications requiring irregular traffic arrivals, semi-persistent scheduling might operate well.

MTC SAFETY FACILITATORS

Future 6G networks will face an essential problem in ensuring the connectivity of numerous MTC devices, each with different service requirements. The boundaries of MTC devices further exacerbate the situation., which are typically delay-sensitive, low-cost, and limited in terms of hardware, processing power, and energy resources [12].

As a result, low-complexity security solutions are essential.

Effective and Affordable Authentication and Validation

Significant signaling congestion may result from millions of connected MTC devices undergoing simultaneous authentication, authorization, and accounting (AAA) procedures. Scalable authentication solutions are required because to the short payloads, latency restrictions of MTC communication, and the devices' low processing capabilities [13].

Promising techniques that are anticipated to be implemented in 6G networks include light-weight physical layer authentication, anonymous customer-focused the authentication process strategies to manage a high volume of requests, group-based authentication schemes, and the incorporation of authentication with access protocols [14].

These solutions aim to efficiently manage authentication while maintaining the scalability and low latency required for MTC services. Another significant challenge with 6G will be user identification. For billions of IoT devices, the traditional subscriber identity module-based approach is just not scalable or economical; as a result, new AAA mechanisms are required. One prominent example is the use of machine learning (ML) techniques in conjunction with radio frequency (RF) signature sensing to recognize a particular radio among devices that are essentially similar [15].

Better physical layer security methods

Since physical layer security (PLS) methods do not require complex encryption or decryption procedures, they have become a viable substitute for protecting MTC networks. But in order to satisfy the demands of energy efficiency, low latency, short payloads, high throughput, dependability, and suitability for use in MTC applications, it is imperative to create more stable, profitable, and helpful PLS strategies. Furthermore, going beyond conventional measurements like secrecy capacity or secrecy outage probability, additional meaningful secrecy metrics that depict real-world and realistic circumstances will be needed to assess the performance of secure systems for MTC.

Using AI and MEC to stop security breaches

MTC networks are quickly becoming highly varied systems, and as the consequence of the broad spectrum of devices, services, and abilities enhanced

AI-based malicious security threats. However, emerging network technologies like MEC (Multi-access Edge Computing) and SDN (Software-Defined Networking), which offer powerful computational resources at the edge, can play a crucial role in developing effective strategies to counter these attacks. By combining these technologies with advanced machine learning (ML) and data mining techniques, more robust security measures can be designed to prevent and mitigate potential threats. [16].

Integrated authentication

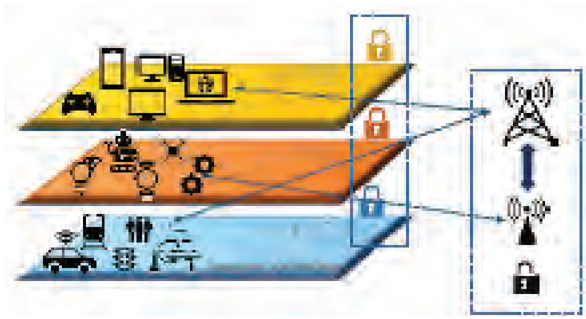


Fig. 2. Network security solutions for core networks with network slicing architecture

ENERGY EFFICIENCY ENABLED BY WIRELESS ENERGY TRANSMISSION

By combining energy-efficient communication, advanced battery technology, offloading energy-intensive processing to the edge (as covered in Section VI), and energy harvesting (EH) techniques, it is possible to eliminate the need for mobile devices to be charged separately [16].

This approach will help ensure continuous, sustainable operation of devices while minimizing the need for traditional charging methods.

Seamlessly integrating wireless information transfer (WIT) and wireless energy transfer (WET) poses a significant challenge due to the differing efficiency requirements of radio circuitry for each. Typical information receivers can operate with very low sensitivities, while energy harvesting devices require much higher levels of incident power. Consequently, separate RF systems are frequently required for energy and information transceivers. Equipping gadgets with distinct antenna systems for every purpose is one possible fix.

Energy and information transmissions can occur either out-of-band or in-band. The out-of-band approach avoids interference between the two, while the in-band method improves spectrum efficiency by allowing both information and energy to be transmitted over the same frequency band, using time-division or even full-duplex techniques. Integrating WET into communication architectures, as shown in Fig. 3, presents several important design challenges. Using hybrid WET+WIT access points and specialized energy beacons, for example, the system architecture must first guarantee the coexistence of heterogeneous WIT and WET. Second, WET's design shouldn't be thought of in a vacuum because it can interfere with other gadgets in the vicinity. WIT and WET must be jointly optimized while accounting for their unique properties and interactions.

wireless. Due to the different radio circuitry efficiencies for wireless information transfer (WIT) and wireless information transfer (WET), integrating WET as an energy harvesting (EH) source is extremely difficult. While an EH device needs a significantly larger incident power, information receivers usually have very modest sensitivity. Because of this, energy and information transceivers typically require different radio frequency systems.

Delay requirements are considered alongside transmission throughput, while The ratio of energy consumption (EC) to total power consumption is known as energy efficiency (EEE). Since EEE reflects both the efficiency of energy utilization and the limits on power consumption, it is especially well-suited to portray the energy-limited nature of WET systems and bursty traffic patterns in MTC.

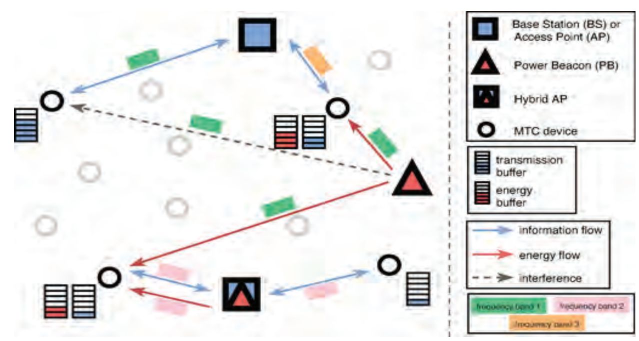


Fig. 3. Integrating Integrating WET into communication architectures requires considering both WET and WIT together for optimal performance.

MTC ENABLER: MULTI-ACCESS EDGE COMPUTING

The deployment and management of dispersed resources for network communication, computation, caching, and data analytics at the boundary of a network is known as multi-access edge computing, or MEC. As an intermediary layer that provides quick, localized data processing for crucial and resource-constrained applications, MEC will be critical for 6G. Security, energy efficiency, Vehicle-to-Everything communication, and offloading for Ultra-Reliable Low-Latency Communications (URLLC) are some of the main purposes for MEC [18]. Figure 4 and the details that follow provide particular uses of the several MEC functions that allow multi-service communication for huge important MTC in 6G.

The The end-to-end (E2E) latency, which refers to the delay between the application layers at both ends of a communication link, is influenced by the quality of the access link and delays caused by both the transport and core networks. Emerging applications, such as augmented reality and cellular V2X, demand low E2E latencies. To address this need, MEC can offer processing capabilities at the network edge, effectively reducing the E2E latency by enabling faster data processing closer to the source. [19].

The advanced capabilities of cloud/centralized data centers enable complex and detailed data analysis, but this comes at the cost of transport and processing delays. The introduction of MEC helps address this challenge by splitting data analytics into two steps. At the edge, fast and localized data processing, analytics, and content caching can support critical applications due to its proximity to end users. Meanwhile, more comprehensive and long-term data analysis can be handled by the cloud core, which provides deeper insights over a larger time scale.

The incessant growth of MTC data volume not only increases the amount of data that needs to be processed, but also makes it more difficult to distinguish between useful and unnecessary data. A primary data analysis cycle at the edge,

Enabled by MEC servers, this approach will allow the vast amounts of MTC data generated by end devices to be processed, filtered, and simplified before being sent to the core/cloud network. This pre-processing step helps reduce the data load on the core network and ensures that only relevant information is transmitted, improving efficiency and reducing latency.

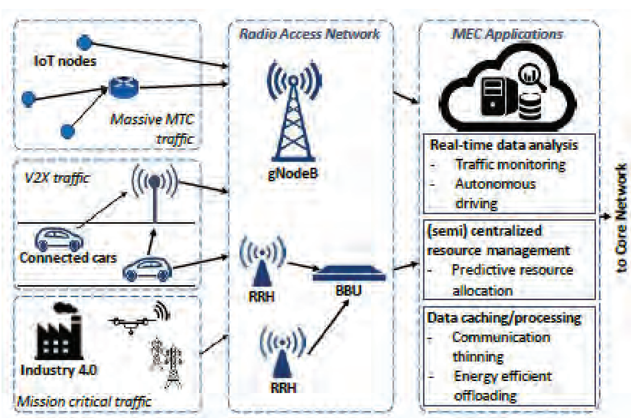


Fig. 4. Examples of different MEC functionalities as enablers for multi- service communication for massive and critical MTC in 6G.

MTC ENABLED BY ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

The 6G vision will require addressing complex, heterogeneous, and often conflicting design requirements. As the network grows in complexity and supports a wider range of configurations and multi-service applications, there will be an increased need for comprehensive network intelligence. Additionally, some 6G applications will require dynamic and/or multiple service-type allocations, moving beyond the static service class categories (eMBB, URLLC, and MTC) used in 5G NR. To support this, processes such as automatic provisioning for establishing network slices for dynamic service types, along with decentralizing the architecture through software-defined networking (SDN) and MEC, can be efficiently implemented.

Adopting Adopting network intelligence involves continuously monitoring changes in the environment and estimating uncertainties, which can then be used to inform decision-making and network reconfiguration. AI techniques, designed to enable machines and systems with intelligence, have been employed for some time in specific scenarios to optimize heterogeneous network configurations [20]. These techniques are now emerging as a key design element for 6G networks, where their role in enhancing network performance and adaptability is expected to be even more critical.

Recently, AI tools have been refocused to address the specific challenges of MTC networks. As a result, we anticipate that AI-driven solutions will become increasingly prominent in the context of MTC, particularly due to its diverse and dynamic requirements. Currently, MTC networks (both massive and critical) are often static with fixed slicing. However, with the growth of vertical-oriented services and the capabilities of MEC, as discussed in Sections II and VI, MTC networks will have the ability to leverage traffic information, usage patterns, and the security needs of MTC devices to enable dynamic slice partitioning and resource allocation, tailored to the demands application.

Table 1 Overview of Six Key Enablers of MTC in 6G

Feature	Addressed Challenge	Key solution components	Expected outcome
Vertical driven network	Differentiated services and products for different verticals	Modular network design at lower layers and the use of networks slicing, NFV, SDN in the higher layers	Newer markets and revenue sources, improved digitization and productivity of the society
Efficient massive access	Fast channel access for dynamic and sporadic MTC traffic with low payload	Fast uplink grant, pre-emptive scheduling via traffic prediction, grant-free transmission, NOMA	Intelligent grant allocation can lead to close to 100% channel occupancy rate
Security for MTC	Ensuring security considering the heterogeneity of MTC devices and traffic	Lightweight AAA, improved PLS and cross-layer security	Enhanced security from sophisticated attacks especially for cMTC services
Energy harvesting via WET	Powering and extending battery lifetime of mMTC devices	Efficient integration of WET and WIT and the use of EEE framework	Longer life time for MTC devices, form factor reduction, green operation
MEC enabled solutions	MEC as an enabler for massive and critical MTC applications	Intermediate data analysis at the edge and distributed computing/caching for resource constrained MTC devices	Resource efficiency, improved localized performance and data/communication thinning
The role of AI/ML	Addressing network complexity and the emergence of multi-service communication	Intelligent algorithms powered by AI/ML targeting applications that cannot be efficiently solved otherwise	Improved network performance considering all KPIs

SUMMARY AND OUTLOOK

The introduction of multi-service communication in 5G NR sparked a paradigm shift from human-type communication (HTC) to machine-type communication (MTC). MTC can be broadly divided into two categories: massive MTC, which focuses on providing energy- and spectrally-efficient connectivity solutions for IoT devices, and critical MTC, which is designed for applications requiring high reliability and low latency. With 75 billion IoT devices projected by 2025, MTC is expected to be the primary focus of 6G wireless networks. In this article, we have outlined a vision for MTC in 6G networks, emphasizing its significance in a vertical-focused wireless network and highlighting key enabling technologies. A comparative overview of these enabling technology components is provided in Table I.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the contributions of their colleagues involved in the project. However, the views expressed in this work are those of the authors and do not necessarily reflect the views of the project.

REFERENCES

1. 3GPP TS 38.300, "NR; Overall description; Stage-2," Jun. 2018, v:15.2.0.
2. A. karimi et al., "5G centralized multi-cell scheduling for URLLC: Algorithms and system-level performance," IEEE Access, 2018.
3. Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," IEEE Tr. IT, vol. 56, no. 5, May 2010.
4. E. Paolini, C. Stefanovic, G. Liva, and P. Popovski, "Coded random access: applying codes on graphs to design random access protocols," IEEE Comm. Mag., vol. 53, no. 6, Jun. 2015.
5. K. David and H. Berndt, "6G vision and requirements: Is there any need for beyond 5G?" IEEE VT Mag., vol. 13, no. 3, Sep. 2018.
6. G. Berardinelli, N. H. Mahmood, I. Rodriguez, and P. Mogensen, "Beyond 5G wireless IRT for industry 4.0: Design principles and spectrum aspects," in IEEE Globecomm Workshops, Dec. 2018.
7. E. C. Strinati et al., "6G: The next frontier: From holographic mes- saging to artificial intelligence using subterahertz and visible light communication," IEEE VT Mag., vol. 14, no. 3, Sep. 2019.
8. M. Katz, M. Matinmikko-Blue, and M. Latva-Aho, "6genesis flagship program: Building the bridges towards 6G-enabled wireless smart society and ecosystem," in IEEE LATINCOM, Guadalajara, Mexico, Nov. 2018.
9. B. Galloway and G. P. Hancke, "Introduction to industrial control networks," IEEE Communications Surveys Tutorials, vol. 15, no. 2, pp. 860–880, Second 2013.
10. S. Ali, N. Rajatheva, and W. Saad, "Fast uplink grant for machine type communications: Challenges and opportunities," IEEE Comm. Mag., vol. 57, no. 3, Mar. 2019.

11. N. H. Mahmood et al., "Uplink grant-free access solutions for URLLC services in 5G New Radio," in 16th ISWCS, Oulu, Finland, Aug. 2019.

12. F. Qasmi, M. Shehab, H. Alves, and M. Latva-Aho, "Optimum transmission rate in fading channels with Markovian sources and QoS constraints," in 15th ISWCS, Lisbon, Portugal, Aug. 2018.

13. J. Ni, X. Lin, and X. S. Shen, "Efficient and secure service-oriented authentication supporting network slicing for 5G-enabled IoT," IEEE JSAC, vol. 36, no. 3, Mar. 2018.

14. N. K. Pratas, S. Pattathil, C. Stefanovic, and P. Popovski, "Massive machine-type communication (mMTC) access with integrated authen- tication," in IEEE ICC, May 2017.

15. S. Riyaz, K. Sankhe, S. Ioannidis, and K. Chowdhury, "Deep learning convolutional neural networks for radio identification," IEEE Comm. Mag., vol. 56, no. 9, Sep. 2018.

16. N. Dao et al., "Securing heterogeneous IoT with intelligent DDoS attack behavior learning," CoRR, vol. abs/1711.06041, 2017. [Online]. Available: http://arxiv.org/abs/1711.06041

17. S. Ulukus et al., "Energy harvesting wireless communications: A review of recent advances," IEEE JSAC, vol. 33, no. 3, Mar. 2015.

18. M. Shehab, H. Alves, and M. Latva-aho, "Effective capacity and power allocation for machine-type communication," IEEE Tr. VT, vol. 68, no. 4, Apr. 2019.

19. M. Emara, M. C. Filippou, and D. Sabella, "MEC-Assisted end-to- end latency evaluations for C-V2X communications," in EuCNC, Jun. 2018.

20. X. Wang, X. Li, and V. C. M. Leung, "Artificial intelligence-based techniques for emerging heterogeneous network: State of the arts, opportunities, and challenges," IEEE Access, vol. 3, 2015.

Leveraging Quantum Mechanics for Secure Communication: A Comprehensive Study of Quantum Key Distribution

Jyoti Mali
Department of Electronics and Telecomm. Engineering
Atharva College of Engineering
Mumbai, Maharashtra
✉ jyotimali@atharvacoe.ac.in

Bhavin Shah
Department of Electronics and Telecomm. Engineering
Atharva College of Engineering
Mumbai, Maharashtra
✉ mr.bhavin.ace@gmail.com

Mohan Kumar
Department of Electronics and Telecomm. Engineering
Atharva College of Engineering
Mumbai, Maharashtra
✉ mohankumar@atharvacoe.ac.in

Ankit D
Department of Electronics and Telecomm. Engineering
Atharva College of Engineering
Mumbai, Maharashtra
✉ ankitdeogharkar@atharvacoe.ac.in

ABSTRACT

Quantum Key Distribution (QKD) represents a revolutionary approach to secure communication, utilizing the principles of quantum mechanics to establish cryptographic keys that are fundamentally secure against eavesdropping. This paper provides a comprehensive study of QKD, detailing its theoretical foundation, practical implementations, challenges, and future potential. By bridging the gap between quantum physics and information security, QKD has the potential to redefine the landscape of secure communication in the era of quantum computing [1][2][3].

KEYWORDS *: Quantum computer, Cryptography, Quantum key distribution.*

INTRODUCTION

The advent of quantum computing poses a significant threat to classical cryptographic protocols, which rely on the computational difficulty of mathematical problems [4]. Quantum Key Distribution offers a solution by harnessing the laws of quantum mechanics to ensure secure communication. This study explores the principles, protocols, and practical implications of QKD, highlighting its role in mitigating risks associated with quantum computing [5]. A classic example of quantum key distribution (QKD) is the "BB84 protocol," where Alice (sender) encodes random bits into the polarization states of single photons and sends them to Bob (receiver) through a quantum channel; Bob measures the photons using a chosen basis, and by comparing their results through a classical channel, they can establish a secure shared key, even if an eavesdropper (Eve) tries to intercept the photons, as any attempt to measure the quantum state will disturb it, alerting Alice and Bob to potential eavesdropping.

THEORETICAL FOUNDATIONS

Principles of Quantum Mechanics

Two fundamental ideas of quantum physics are superposition and entanglement. They explain how quantum systems can be related to one another even when they are separated by great distances and can exist in numerous states.

Superposition [6]

Superposition is a fundamental principle of quantum mechanics that states that a physical system exists in all possible states simultaneously until it is measured or observed. This means that a particle, such as an electron, can be in multiple states at the same time, rather than being in just one specific state. For example, an electron can be in a superposition of spinning clockwise and counterclockwise at the same time, until it is measured and collapses into one of the two states. This concept is often illustrated by the famous thought experiment of

Schrödinger's cat, where a cat in a box is both alive and dead until the box is opened and the cat's state is observed. Superposition is a key feature of quantum mechanics and is essential for understanding phenomena such as quantum entanglement and quantum computing. It challenges our classical intuition about the nature of reality and has profound implications for the way we understand the behavior of particles at the quantum level

Entanglement [7][8]

Entanglement is a quantum mechanical phenomenon in which two or more particles become intrinsically linked, meaning the state of one particle is directly influenced by the state of the other, no matter how far apart they are. As a result, their properties remain instantaneously correlated, even across vast distances. This concept challenges our classical understanding of physics, as it suggests that particles can be interconnected in ways that are not explained by classical physics. Entanglement has been experimentally verified and is a key aspect of quantum computing and quantum communication technologies.

Heisenberg's Uncertainty Principle [9]

Heisenberg's Uncertainty Principle is a fundamental concept in quantum mechanics that states that it is impossible to simultaneously know both the exact position and momentum of a particle with absolute certainty. In other words, the more accurately we determine a particle's position, the less accurately we can measure its momentum, and vice versa. This principle arises from the wave-particle duality of quantum mechanics, which states that particles can exhibit both wave-like and particle-like behavior. As a result, the act of measuring one property of a particle, such as its position, disturbs the other property, such as its momentum. The Uncertainty Principle has profound implications for our understanding of the behavior of particles at the quantum level and has been confirmed through numerous experiments. It fundamentally limits our ability to predict the behavior of particles and highlights the inherent uncertainty and randomness that exists in the quantum world.

QUANTUM KEY DISTRIBUTION PROTOCOLS

Quantum Key Distribution (QKD) protocols are a set of cryptographic techniques that use the principles of quantum mechanics to securely distribute encryption keys between two parties. These protocols leverage the unique properties of quantum particles, such as photons, to ensure that any attempt to intercept or eavesdrop on the key exchange would be immediately detected. One of the key principles behind QKD protocols is the Heisenberg Uncertainty Principle, which states that the act of measuring a quantum system will inevitably disturb it. This means that any attempt to intercept the quantum key would alter its state, alerting the parties involved to the presence of an eavesdropper. There are several different QKD protocols that have been developed, each with its own strengths and weaknesses. Some of the most commonly used protocols include BB84, E91, and SARG04. These protocols typically involve the transmission of quantum particles between the two parties, followed by a series of measurements and comparisons to verify the security of the key exchange. Overall, QKD protocols offer a highly secure method for distributing encryption keys, as they are resistant to both classical and quantum attacks. However, they are still relatively new and complex technologies, and there are ongoing efforts to improve their efficiency and practicality for real-world applications.

The fundamental model for Quantum Key Distribution (QKD) protocols involves two participants, Alice and Bob, who aim to securely exchange a key. They have access to both a classical public communication channel and a quantum communication channel, as illustrated in Figure 1. An eavesdropper, known as Eve, is assumed to have access to both channels, with no restrictions on the resources she may use. With this model in place, we will explain the essential quantum principles in simple terms to provide a clear understanding of QKD protocols.

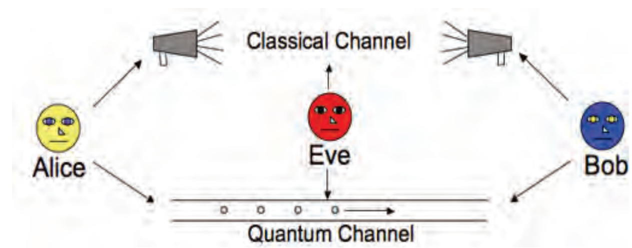


Fig. 1 QKD model

BB84 Protocol

The BB84 Protocol is a quantum key distribution protocol developed by Charles Bennett and Gilles Brassard in 1984. It is designed to allow two parties to securely establish a shared encryption key using the principles of quantum mechanics. In the BB84 Protocol, the sender (Alice) sends a series of quantum bits (qubits) to the receiver (Bob) over a quantum channel. These qubits are encoded with random bits, either in the 0 or 1 state, using one of four possible bases: the standard basis (0/1), the Hadamard basis (+/-), or the diagonal bases (45/135 degrees). Bob then measures each qubit using one of the four possible bases, randomly chosen for each qubit. After the transmission is complete, Alice and Bob publicly compare the bases they used for each qubit. If they used the same basis, Bob's measurement will match Alice's encoding. If they used different bases, Bob's measurement will be random. By discarding the qubits where they used different bases, Alice and Bob can establish a shared encryption key that is secure against eavesdropping. Any attempt to intercept the qubits would disturb the quantum state and be detected during the key comparison process. The BB84 Protocol is considered one of the first practical applications of quantum cryptography and is widely used in quantum key distribution systems for secure communication. [1] [10]

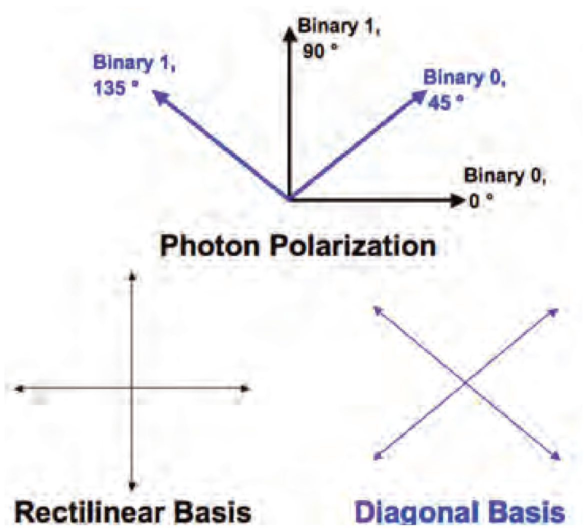


Fig. 2 BB84 Bit encoding

Figure 2 illustrates how a bit can be encoded using the polarization state of a photon in the BB84 protocol. In this scheme, a binary 0 is represented by a photon polarized at 0 degrees in the rectilinear basis or at 45 degrees in the diagonal basis [CKI-BB84] [Gisin02]. Likewise, a binary 1 corresponds to a polarization of 90 degrees in the rectilinear basis or 135 degrees in the diagonal basis. This means that each bit is encoded by polarizing the photon in one of two possible bases.

E91 Protocol

The E91 Protocol is a proposed standard for emergency calling services in India. It aims to provide a universal emergency number (similar to 911 in the United States) that can be dialed from any phone, including landlines, mobile phones, and public payphones. The protocol also includes provisions for location-based services, so that emergency responders can quickly locate the caller in case of an emergency. The E91 Protocol is intended to streamline and improve emergency response services in India by providing a single, easy-to-remember number for all types of emergencies.

PRACTICAL IMPLEMENTATIONS OF QKD

Hardware and Infrastructure

Practical implementations of Quantum Key Distribution (QKD) involve the use of specific hardware and infrastructure to ensure secure communication. Two key components of this infrastructure are quantum random number generators and single-photon sources and detectors. Quantum random number generators are devices that generate truly random numbers based on the principles of quantum mechanics. These random numbers are essential for creating secure encryption keys in QKD protocols, as they cannot be predicted or manipulated by an eavesdropper. Single-photon sources and detectors are also crucial for QKD implementations, as they enable the transmission and detection of individual photons, which are used to encode and decode quantum keys. Single-photon sources emit photons one at a time, ensuring that each bit of information is transmitted securely. Single-photon detectors are then used to measure the state of these photons without disturbing them, allowing for secure key distribution. Overall, the use of quantum random number generators and single-photon sources and detectors in QKD implementations helps to ensure the security and reliability of quantum communication systems.

Network Integration

Network integration refers to the process of connecting different components or systems within a network to enable communication and data exchange. In the context of quantum key distribution (QKD), network integration involves incorporating QKD systems into existing networks to enhance security and privacy. Point-to-point QKD refers to a direct communication link between two parties using QKD technology to exchange cryptographic keys securely. This method is suitable for secure communication between two specific endpoints but may not be scalable for larger networks. On the other hand, networked QKD involves connecting multiple QKD systems in a network topology to enable secure communication between multiple parties. This approach allows for secure communication between multiple endpoints within a network and is more suitable for larger-scale applications. Trusted repeater and satellite-based systems are two approaches to implementing networked QKD. Trusted repeater systems involve intermediate nodes (repeaters) that facilitate the distribution of quantum keys between distant parties in a network. Satellite-based systems use satellites to distribute quantum keys over long distances, enabling secure communication between parties located in different geographical locations. Overall, network integration of QKD systems, whether through point-to-point or networked approaches, along with trusted repeater or satellite-based systems, can enhance the security and privacy of communication within a network.

Use Cases

Government and military communication: This refers to the use of secure communication systems by government agencies and military organizations to exchange sensitive information and coordinate operations. These systems ensure that communication is encrypted and protected from unauthorized access, ensuring the security and confidentiality of the information being transmitted. Financial transactions: This involves the use of secure communication systems to facilitate financial transactions, such as online banking, e-commerce, and electronic fund transfers. These systems ensure that financial data is encrypted and protected from interception or tampering, safeguarding the integrity and confidentiality of the transactions. Critical infrastructure: This refers to the use of secure communication systems to protect critical infrastructure, such as power grids, transportation networks, and water supply systems, from cyber threats and attacks. These systems help to ensure the reliability and availability of essential services by safeguarding the communication networks that support them.

BLOCK DIAGRAM DESCRIPTION

1. Sender (Alice):
   - Generates a random sequence of bits (key).
   - Encodes the bits into quantum states using photons (e.g., polarization states in BB84 protocol).
2. Quantum Channel:
   - Transfers the encoded quantum states (photons) to the receiver (Bob).
   - Any eavesdropping (by Eve) introduces detectable disturbances due to quantum mechanics.
3. Receiver (Bob):
   - Measures the received photons using a random basis.
   - Decodes the transmitted bits.
4. Classical Channel:
   - Alice and Bob use a public classical channel for error correction and basis reconciliation.
   - No actual key data is transferred here, maintaining security.
5. Key Agreement:
   - Alice and Bob agree on the final cryptographic key after removing errors and eavesdropper-detectable bits.
6. Eavesdropper (Eve):
   - Attempts to intercept quantum states. Any such attempt introduces detectable anomalies.

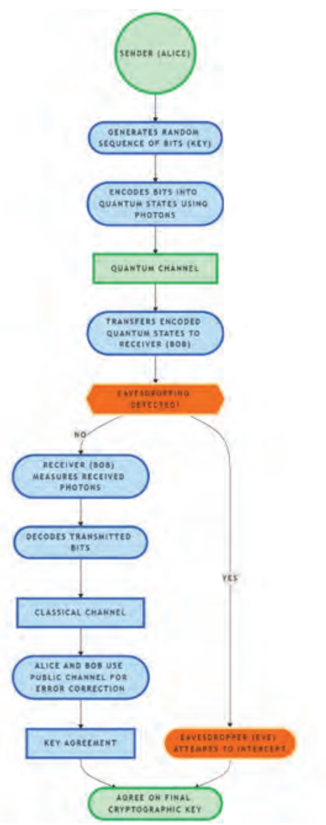


Fig.3 Flow chart of QKD

The block diagram describes the key components and processes involved in quantum key distribution (QKD), a secure communication method that uses principles of quantum mechanics to generate and distribute cryptographic keys. The process begins with the sender, Alice, who generates a random sequence of bits to serve as the key. She then encodes these bits into quantum states using photons, typically employing polarization states as in the BB84 protocol. These encoded quantum states are transmitted through a quantum channel to the receiver, Bob. The quantum nature of the transmission ensures that any eavesdropping attempts by a third party (Eve) would introduce detectable disturbances in the system. Upon receiving the photons, Bob measures them using a randomly chosen basis and decodes the transmitted bits. To refine and verify the key, Alice and Bob communicate over a classical channel for error correction and basis reconciliation, without revealing the actual key data. Finally, they agree on the final cryptographic key after removing any errors, resulting in a secure shared key that can be used for encrypted communication.

CONCLUSION

Quantum Key Distribution represents a paradigm shift in secure communication, providing a mechanism that is inherently secure against quantum and classical attacks. While challenges remain in terms of scalability, cost, and standardization, ongoing research and technological advancements promise a robust future for QKD [3]. By addressing these challenges, QKD can become a cornerstone of secure communication in the quantum era.

REFERENCES

1. Bennett, C. H., & Brassard, G. (1984). Quantum cryptography: Public key distribution and coin tossing.
2. Ekert, A. K. (1991). Quantum cryptography based on Bell’s theorem.
3. Scarani, V., et al. (2009). The security of practical quantum key distribution.
4. Shor, P. W. (1994). Algorithms for quantum computation: Discrete logarithms and factoring.
5. Gisin, N., et al. (2002). Quantum cryptography.
6. Dirac, P. A. M. (1930). The Principles of Quantum Mechanics.
7. Einstein, A., Podolsky, B., & Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete?
8. Bell, J. S. (1964). On the Einstein Podolsky Rosen paradox.
9. Heisenberg, W. (1927).
10. Mayers, D. (2001). Unconditional security in quantum cryptography.
11. Ralph, T. C. (1999). Continuous variable quantum cryptography.
12. Cerf, N. J., et al. (2001). Security of quantum key distribution using continuous variables of single photons.
13. Lo, H. K., et al. (2012). Measurement-device-independent quantum key distribution.
14. Kwiat, P. G., et al. (1995). New high-intensity source of polarization-entangled photon pairs.
15. Hadfield, R. H. (2009). Single-photon detectors for optical quantum information applications.
16. Liao, S. K., et al. (2017). Satellite-to-ground quantum key distribution.
17. Wehner, S., et al. (2018). Quantum internet: A vision for the road ahead.
18. Lydersen, L., et al. (2010). Hacking commercial quantum cryptography systems by tailored bright illumination.
19. Diamanti, E., et al. (2016). Practical challenges in quantum key distribution.
20. Bravyi, S., & Kitaev, A. (1998). Quantum codes on a lattice with boundary.
21. Awschalom, D. D., et al. (2013). Quantum spintronics: Engineering and manipulating atom-like spins in semiconductors.
22. [22] Mosca, M. (2018). Cybersecurity in an era with quantum computers.

ChatGPT's Impact on Student Learning: Enhancing Education or Encouraging Dependency?

Chetana Rane
Student, M.Sc. Data Science
Fergusson College (Autonomous)
Pune, Maharashtra
✉ chetanarane10@gmail.com

Smita Bhanap
Assistant Professor
Fergusson College (Autonomous)
Pune, Maharashtra
✉ smita.bhanap@fergusson.edu

Rasika Kulkarni
Assistant Professor
Fergusson College (Autonomous)
Pune, Maharashtra
✉ rasika.kulkarni@fergusson.edu

Lakhichand Khushal Patil
Assistant Professor
Fergusson College (Autonomous)
Pune, Maharashtra
✉ lakhichand.patil@fergusson.edu

ABSTRACT

One of the most significant and impactful innovation strategies of this century is artificial intelligence (AI). It affected every business in different ways, including the food industry, retail education, banking, healthcare, and the automotive sector, among many others. AI has emerged as one of the major issues confronting our world today, affecting everything from the appropriate use of technology to the quality of services provided to the future of many jobs. Education is one of the sectors most affected by AI, where it can occasionally have negative effects as well as positive effects. This work investigates the effects of a sophisticated AI language model, like ChatGPT, on students' learning. Understanding how digital technologies like ChatGPT influence student behaviour, performance, and learning experiences is becoming more and more important as they become more popular in educational settings. In order to evaluate the advantages and disadvantages of using ChatGPT in educational settings, this study looks at the answers to a survey form that was completed by 315 students. It looks into whether using ChatGPT as a learning tool can make students dependent and offers solutions for implementing ChatGPT in classrooms while addressing its drawbacks using machine learning models like random forest and XGBoost classifiers. The random forest model performs well with 98% accuracy whereas XGBoost 96%.

KEYWORDS *: ChatGPT, AI, Education, Dependency, Critical thinking, Learning enhancement, Random forest, XGBoost.*

INTRODUCTION

The development of information and communication technologies in recent years has brought about major changes in the field of education. Because of the technology improvements, online learning has become more accessible and adaptable. Nowadays, e-learning is the main tool used by many educational institutions to provide students with a more stimulating and effective learning environment. As technology develops, artificial intelligence (AI) is growing rapidly as well. One aspect of AI that has caught the attention of the education sector is chatbots. In a number of situations, they have been used to assist and support people. A chatbot is computer software designed to communicate with users. One of the increasingly popular kinds of chatbots is the pre-trained Trans-former (ChatGPT). ChatGPT is a natural language-based artificial intelligence model that can comprehend and process human words [1]. With its rapid growth, artificial intelligence (AI) has changed many aspects of modern life and given rise to new forms. One of its more recent breakthroughs is ChatGPT [2]. Additionally, with regard to intelligent tutoring programs, students can use Chat GPT or other AI software to generate assignments, receive feedback and suggestions for revision, and obtain help with writing tasks[3]. Although ChatGPT has numerous educational advantages, there are drawbacks to using it. AI has both advantages and disadvantages, therefore it's important to strike the right balance when applying it to solve problems, particularly when ethical considerations are involved. Students may learn course content faster and find pertinent information to expand their knowledge with ChatGPT. Students' inventiveness in doing academic assignments is positively impacted by ChatGPT in addition to increasing time efficiency. With ChatGPT's guidance, students can generate original ideas as they learn fresh strategies for writing and completing assignments. By offering interesting, individualized experiences catered to each user's skills and preferences, ChatGPT has completely transformed language learning. But some students acknowledge that they rely too much on ChatGPT for their tasks, frequently ignoring deeper learning and independent thought. Lot of students would rather do assignments utilizing ChatGPT than do independent research or deeper comprehension. Teachers are concerned that students who rely too much on ChatGPT may become lazy and less inclined to solve difficulties on their own. Many instructors worry that pupils will lose initiative in critical thinking and problem-solving if they become increasingly reliant on AI. Thus, the purpose of this study is to investigate how well students embrace ChatGPT technology, with an emphasis on both its advantages and disadvantages in relation to learning outcomes and academic ethics. Additionally, it looks at how students utilize ChatGPT to do schoolwork and how it impacts their capacity for critical thought and problem-solvin.[4].

This study examines ChatGPT's dual effects on learners, allowing for the assessment of both positive and negative effects. Survey responses from 315 students will be crucial in addressing questions about whether or not students think ChatGPT is a useful learning tool and whether or not it makes users dependent. The answers to these queries will inform suggestions for how to optimize ChatGPT's positive effects while minimizing its drawbacks.

The main contributions of this paper focus on-

1. To investigate the factors influencing students' perceptions of whether ChatGPT enhances their learning experience.
2. To develop a model that predicts students' beliefs about ChatGPT's impact on their learning experience based on usage patterns, demographics, and other relevant variables.
3. To evaluate the effectiveness of ChatGPT as a learning tool by analysing student perceptions of its impact on their learning experience.
4. To identify key predictors of positive and negative perceptions of ChatGPT's influence on the learning experience.

MOTIVATION

My motivation for this research stems from observing a significant shift in student study patterns, particularly after the pandemic. Before, resources like ChatGPT weren't available, and we relied on our own thinking and problem-solving skills for every task. We had to actively engage with the material and develop our own understanding. Now, it seems everything is readily available. Many students' complete assignments using AI tools like ChatGPT without truly grasping the underlying concepts. This widespread reliance on readily available solutions is what prompted my research. I've noticed how frequently students depend on ChatGPT, often submitting AI-generated work without fully understanding it. This trend raises serious concerns about the development of critical thinking skills and the long-term implications for their future. I wanted to investigate this shift and understand how this reliance on AI tools like ChatGPT is impacting students' learning and critical thinking abilities.

LITERATURE REVIEW

AI is currently posing a serious threat to the world's employment prospects. Additionally, it affects the field of education in both positive and harmful ways. Since it makes everything by itself, it results in a lack of in-depth knowledge, concerns about data privacy, and instructor job losses, all of which will pose serious challenges down the road[5]. This paper[1] discusses how ChatGPT affects students' learning. Along with several interviews, a test was administered to 120 students. They discovered that ChatGPT offers both advantages and disadvantages. On the one hand, it enhances students' ChatGPT skills, while on the other, it diminishes their capacity for originality and critical thought. Nonetheless, it is helpful for improving learning capacity and providing insightful answers. When utilized appropriately, it can be an excellent learning tool. This research [2] primarily examines the effects of ChatGPT on English language acquisition. They also carried out some study on the subject, including surveying 80 teachers and students, and discovered that when applied appropriately, it has a positive effect on children. The only need is that they be able to use it. They advise concentrating on the benefits rather than the drawbacks. The detrimental effects of ChatGPT on students' learning are the main topic of this paper [3]. It claims that it lowers their motivation to learn. According to the results, ChatGPT has a negative effect on students because, out the 20 respondents, 12 stated that ChatGPT has a negative effect on students. It is harming their crucial situations, they add. ChatGPT discourages students from being motivated to learn. Additionally, they assert that students should only use it for educational purposes. They claim[5] that ChatGPT is helpful because students are completing their tasks and learning everything rapidly. Therefore, the difficulties may include a lack of knowledge about academic ethics, the possibility of plagiarism, and an excessive dependence on technology. Strict rules should be implemented, and educational institutions should instruct students on how to learn it. This paper[6] explains how ChatGPT is permeating every aspect of people's lives and how students are continuously depending on it, which is negatively impacting their academic achievement, real-world critical thinking abilities, and fundamental core skills. In order to accomplish academic objectives and acquire specific abilities, they recommend balancing AI technologies in their lives. This study[7] examines how ChatGPT affects modifications in students' learning habits. Additionally, they conducted several interviews and observations, which demonstrate that it affects students' learning in both positive and negative ways. On the plus side, they finish their assignments fast, and some students claimed in the interviews that it benefits them by allowing them to condense their responses. However, some students believed it had a negative effect since it made them lethargic and dependent on reading, which resulted in a decline in literacy and other bad consequences. They came to the conclusion that, depending on how it is utilized, ChatGPT can have both beneficial and detrimental effects. After surveying[8] the students, they discovered that they are utilizing ChatGPT to finish their schoolwork and study for tests, which is impairing their ability to think critically. Students, however, prefer YouTube because it is simpler to extract information from it than from ChatGPT-generated photos. Thus, it affects kids' learning in both positive and negative ways. ChatGPT[9] affects higher education in both positive and negative ways. According to research involving 56 students, people enjoy using it for studying because it's user-friendly and offers useful answers. Many pupils, however, believe that it requires prior knowledge and that its responses aren't always accurate. They are optimistic that it will advance even though they think it cannot completely replace human intelligence. Some students are worried about how it will affect their education and careers. All things considered, ChatGPT is helpful for education, but students should be conscious of its limitations. This study[10] examined the impact of ChatGPT on the motivation of college students. It was discovered that while students' confidence in utilizing ChatGPT was associated with greater happiness, tension and enjoyment were negatively correlated. However, students' motivation was not significantly impacted by how valuable they thought ChatGPT was. According to the study's findings, ChatGPT has an influence on motivation, although students require more practice to use it efficiently. The findings offer fresh concepts for enhancing instructional strategies and upcoming studies on student motivation.

METHODOLOGY

Figure 1 Illustrates the schematic representation of the process comprises of Data Acquisition, Data pre-processing, Exploratory Data Analysis, Feature Engineering, Model Building. These processes are detailed in the following subsections.

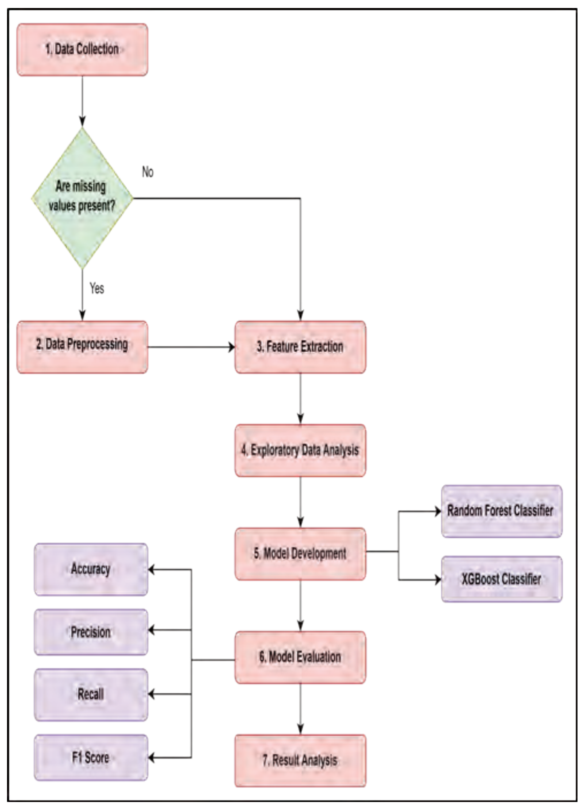


Fig. 1: Flowchart for Analysing ChatGPT's Impact on Students' Learning

DATA ACQUISITION/ DATA COLLECTION

To study the impact of ChatGPT, it is required to get real time data from relevant target people who are students. A survey questionnaire is prepared considering different approaches and issues working with ChatGPT. Following is the sample dataset.

1	What is your age group?	What is your field of study?	What is your current level of education?
2	19-21	Data Science/AI	Postgraduate
3	22-25	Arts	Undergraduate
4	22-25	Data Science/AI	Postgraduate
5	19-21	Engineering	High School
6	Above 25	Science	Postgraduate
7	19-21	Engineering	Undergraduate
8	19-21	Data Science/AI	Undergraduate
9	22-25	Data Science/AI	Undergraduate
10	22-25	Science	Postgraduate
11	Above 25	Commerce	Postgraduate
12	22-25	Commerce	High School
13	19-21	Science	Postgraduate
14	19-21	Arts	Postgraduate
15	22-25	Data Science/AI	Postgraduate
16	22-25	Engineering	Undergraduate
17	19-21	Data Science/AI	Postgraduate
18	22-25	Science	Undergraduate
19	19-21	Data Science/AI	Postgraduate
20	22-25	Data Science/AI	Postgraduate
21	19-21	Science	High School
22	22-25	Data Science/AI	Postgraduate
23	22-25	Engineering	Undergraduate
24	22-25	Data Science/AI	Postgraduate
25	19-21	Engineering	Undergraduate
26	22-25	Science	Postgraduate
27	19-21	Computer Science	Postgraduate

Fig. 2: Sample dataset

The dataset will consist of 315 rows (data from students) and 26 columns (questions). Since we asked 26 questions in the survey, the data was too large to display in the paper. Therefore, we have shown only a sample of the data as an example and the age group considered to carry out this study is 19 to 21 Years old students.

DATA PREPROCESSING

We conducted a Google Forms survey among students using various questionnaires. From the survey, we compiled a dataset of approximately 315 observations. However, some features were inadequate and required data cleaning. Following pre-processing is performed on the data generated through the survey questionnaire[11].

- Dropping features which are not required.
- Cleaning erroneous inputs
- Renaming Column as required.
- Encoding is required for categorical features.

EXPLORATORY DATA ANALYSIS

Exploratory Data Analysis (EDA) helps us understand the data before drawing conclusions. We use it to uncover hidden patterns and relationships. The dataset is studied to understand the students' responses and observe patterns in it. Following are the key observations showcased with the help of graphs.

Relationship Between Age Group and Frequency of ChatGPT Usage (Stacked Bar Chart)

Purpose: To understand how ChatGPT usage frequency varies across different age groups.

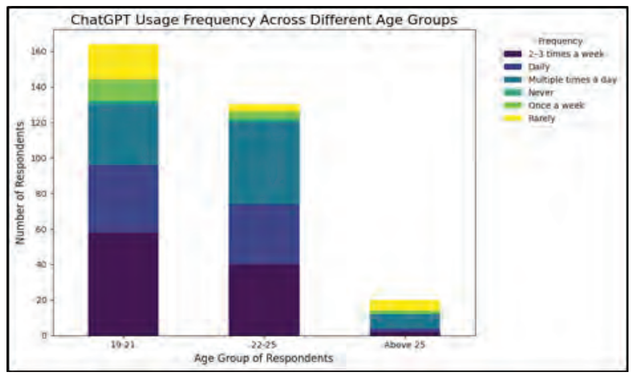


Chart 1: Relationship Between Age Group and Frequency of ChatGPT Usage

Interpretation: You can see how the 19-21 age group compares with other age groups in terms of daily, weekly, or rare usage of ChatGPT. If this group shows the highest daily usage, it validates your focus on them.

ChatGPT Usage vs. Perceived Critical Thinking Impact (Box Plot)

Purpose: To analyse whether frequent ChatGPT users report a greater negative or positive impact on their critical thinking ability.

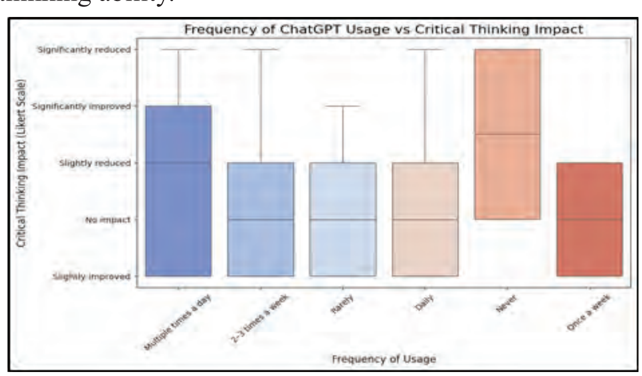


Chart 2: ChatGPT Usage vs. Perceived Critical Thinking Impact

Interpretation: Frequent ChatGPT users ("Multiple times a day") mostly report positive impacts on critical thinking, while less frequent users show more varied perceptions, including neutral and negative impacts, suggesting a possible link between usage frequency and perceived impact.

Comfort with AI tools vs. dependency on ChatGPT (Violin Plot)

Purpose: To visualize the distribution of dependency levels on ChatGPT across different comfort levels of users.

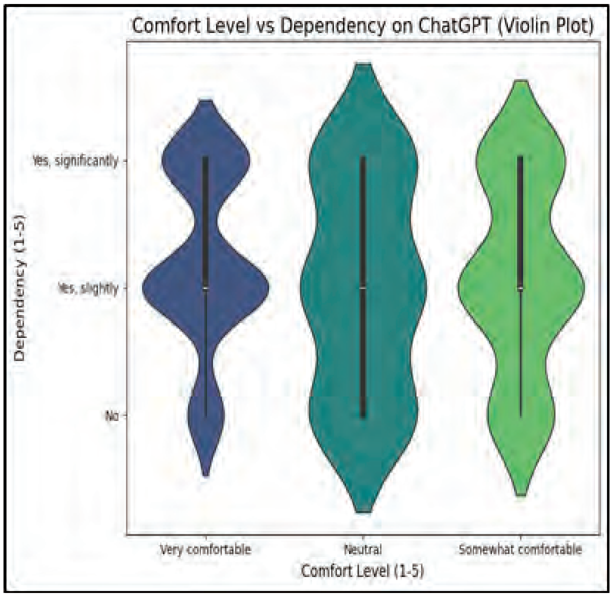


Chart 3: Comfort with AI tools vs. dependency on ChatGPT

Interpretation: ChatGPT dependency appears to increase with user comfort level. "Very comfortable" users mostly report significant dependency, while "Neutral" and "Somewhat comfortable" users show more varied responses, including many reporting no dependencies.

Frequency of ChatGPT Use vs. Perceived Learning Enhancement (Grouped Bar Chart)

Purpose: To investigate the correlation between how often students use ChatGPT for academic purposes and their perception of whether it enhances their learning experience.

Interpretation: ChatGPT users in the sample all found it enhanced their learning. Weekly users were mostly positive, while monthly users were more mixed, suggesting less frequent use may correlate with lower perceived learning enhancement.

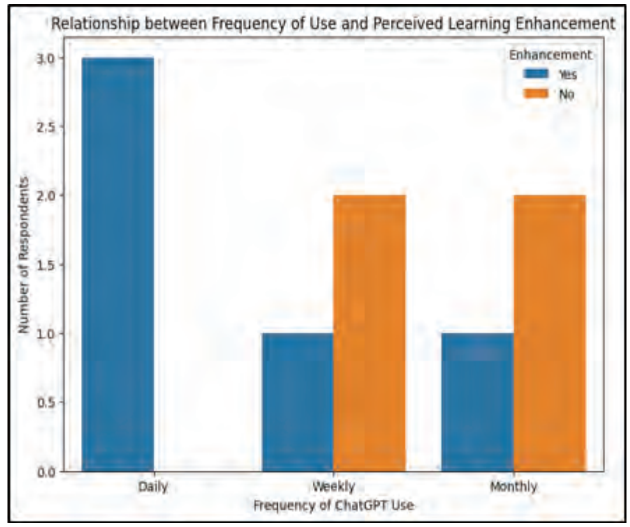


Chart 4: Frequency of ChatGPT Use vs. Perceived Learning Enhancement

Frequency of ChatGPT Use vs. Perceived Impact on Creativity (Heatmap)

Purpose: To visualize the relationship between the frequency of ChatGPT use (Daily, Monthly, Weekly) and users' perceptions of its impact on their creativity (Fosters, Hinders, Neutral).

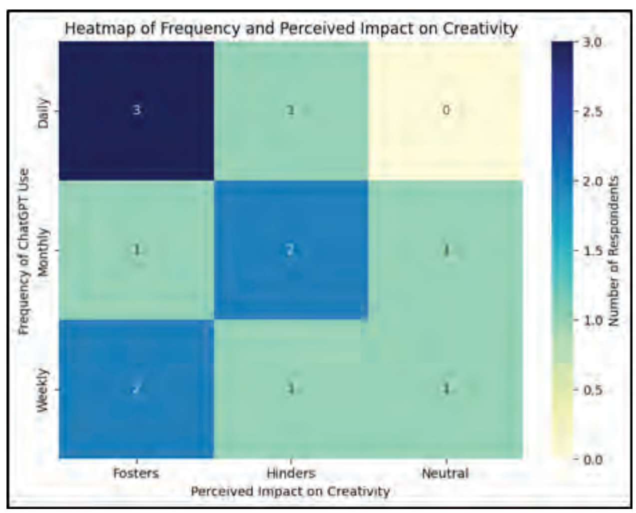


Chart 5: Frequency of ChatGPT Use vs. Perceived Impact on Creativity

Interpretation: Daily ChatGPT users strongly associate it with fostering creativity, while monthly and weekly users have more varied opinions. Higher usage frequency appears linked to a more positive perception of ChatGPT's impact on creativity.

FEATURE ENGINEERING

In the given dataset, there are many features relevant to the problem, but out of all the features, we selected only 15 that are more correlated with our target output. For this, we used the SelectKBest selector to find the best top 15 features for our model. These 15 features are: ['What is your field of study?', 'How comfortable are you with using AI tools like ChatGPT?', 'How frequently do you use ChatGPT for academic purposes?', 'For which academic purposes do you primarily use ChatGPT? (Select all that apply)', 'Do you use ChatGPT for non-academic purposes (e.g., hobbies, personal development)?', 'Has ChatGPT influenced your effort levels in completing assignments?', 'How does using ChatGPT affect your ability to think critically or solve problems independently?', 'Do you face difficulty recalling information learned with ChatGPT's assistance compared to traditional methods?', 'Has using ChatGPT discouraged you from practicing skills (e.g., coding, writing, or critical analysis) on your own?', 'Have you ever submitted ChatGPT-generated content without modifications?', 'Do you believe using ChatGPT increases the risk of academic dishonesty?', 'How would you rate your understanding of plagiarism and copyright issues related to AI tools?', 'Does ChatGPT make you feel overly reliant on technology?', 'How satisfied are you with the accuracy of ChatGPT's responses?', 'Do you think educational institutions should actively teach responsible AI usage?']. For the target variable, we used the "Do you believe ChatGPT enhances your learning experience?" parameter, which resulted in a proper accuracy for our model building[12].

MODEL BUILDING

Machine learning is a valuable tool for prediction and analysis, and given the nature of our data and research questions, we opted for a supervised learning approach. We selected Random Forest and XGBoost, two algorithms particularly well-suited for this task due to their ability to handle complex relationships and provide robust predictive models.

Machine Learning Classifiers

XGBoost Classifier

The gradient boosting-based decision tree ensemble XGBoost[5] is made to be extremely scalable. By minimizing a loss function, XGBoost creates an additive extension of the objective function, much like gradient boosting does. Given that XGBoost solely uses decision trees as base classifiers, the trees' complexity is managed using a variant of the loss function. With this, XG Boost is mainly used to reduce overfitting[13].

Random Forest Classifier

Two distinct sources of randomization are used to build the decision trees that make up the Random Forest ensemble of classifiers. Initially, a random sample of the original data that is the same size as the training set is used to train each decision tree separately[14].

RESULT ANALYSIS

Based on the machine learning classifiers used for model building, the following table shows the accuracy results. The Random Forest classifier performs well with overall accuracy of 98% and XG Boost gives accuracy of 96%.

Classifier Name	Accuracy	Precision	Recall	F1 score
Random Forest	0.98	1.00	0.89	0.94
XGBoost	0.96	0.95	1.00	0.97

Table 1: Comparison table between Random Forest and XGBoost Classifier

The following table shows percentage of Student Acceptance of ChatGPT Use with respect to component of Technology Acceptance Model (TAM).

Table 2. Percentage of Student Acceptance of ChatGPT Use

Component of Technology Acceptance Model (TAM)	Helping / Strongly Agree	Helping / Agree	Neutral	Not helping / Disagree	Not helping / Strongly Disagree
Effort Impacted	20.4%	22.9%	8,9%	34.4%	13.4%
Critical Impact	18.5%	27.4%	34.4%	13.4%	6.4%
Recall Difficulty	20.4%	43.3%	20%	11.3%	5%

Concept Clarity	32.5%	49.7%	16.6%	0.6%	0.6%
Skill Development	20.4%	51.2%	21.4%	5.5%	1.5%
Exam Readiness	15.3%	22.9%	33.8%	22.3%	5.7%
Dishonesty Risk	17.2%	45.9%	22.9%	10.2%	3.8%
AI Restriction	10.8%	18.5%	28.7%	32.5%	9.6%
Learning Enhancement	26.1%	52.2%	15.9%	3.8%	1.9%
Accuracy Satisfaction	12.7%	49.7%	29.9%	6.4%	1.3%

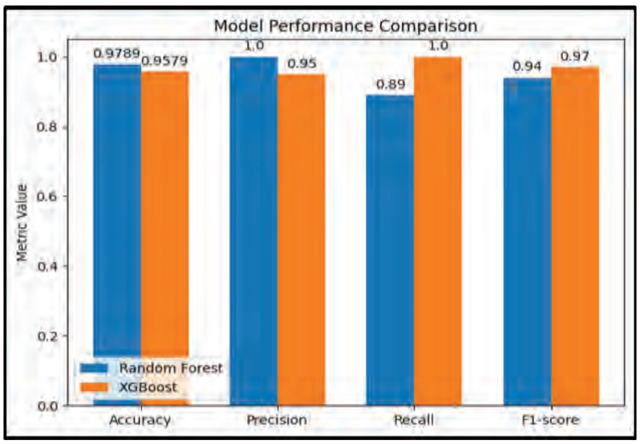


Chart 6: Model Performance Comparison: Random Forest vs. XGBoost

Interpretation: Random Forest was a little more accurate overall, but XGBoost did a better job of balancing getting the right answers and not missing any important ones. XGBoost's higher F1-score shows it's probably a better choice for this particular problem, even if its accuracy was a tad lower. It really depends on what's most important – if you want to be really sure you're not missing anything, XGBoost is better, but if you just want to be mostly right most of the time, Random Forest is okay too.

Performance Measures

Accuracy, Precision, Recall, and F1-score are the metrics used to check the performance of the model , and also shows the comparative analysis between two algorithms Random Forest and XGBoost[15].

Accuracy

- In terms of accuracy, Random Forest (0.98) performs marginally better than XGBoost (0.96), correctly classifying a greater number of occurrences overall.
- But accuracy by itself isn't necessarily the best measure, particularly when the dataset is unbalanced.

Precision

- All of Random Forest's positive predictions are accurate, as evidenced by its perfect precision score of 1.00.
- The slightly reduced precision of XGBoost (0.95) suggests some erroneous positives.

Recall

- Perfect recall is attained by XGBoost (1.00), which accurately detects every real positive instance.
- The weaker recall of Random Forest (0.89) suggests that it overlooks some favourable cases.

F1 Score

- With a marginally higher F1 score than Random Forest (0.94), XGBoost (0.97) strikes a better balance between recall and precision.
- The F1-score indicates that XGBoost is a superior classifier overall in this instance since it is the harmonic mean of precision and recall.

CONCLUSION

ChatGPT presents a complex landscape for education. Our survey of around 315 students using Google Forms revealed a mix of experiences. While ChatGPT offers clear benefits, like simplifying tough topics, helping with self-study, and boosting motivation (something we saw in the positive correlations), it also carries risks, especially over-reliance and potential harm to critical thinking and problem-solving. Our analysis, including the model we built using things like how often students use ChatGPT, how comfortable they are with it, and how they feel it impacts their skills, really highlights this double-edged sword. We noticed that students who use ChatGPT multiple times a day tend to report positive effects on their learning and motivation. However, students who use it less often have all sorts of opinions – some good, some neutral, and some even negative, particularly when it comes to critical thinking. This tells us that finding the right balance is super important. So, for ChatGPT to really work well in education, it needs to be seen as a supporting tool, not a replacement for doing your own research, thinking critically, and coming up with creative solutions. Teachers are key here. They need to guide students on how to use AI responsibly and make sure it's helping them learn and think better, not the other way around.

Conflict of Interest: I, the main author, and all of the corresponding authors have no conflict of interest.

Competing Interests: Being the main author and on behalf of all associated authors, I have no competent interest.

Funding Information: NA

Author contribution: All authors are equally contributed.

Data Availability Statement: Authors may provide data on request.

Research Involving Human and /or Animals: NA

Informed Consent: NA

REFERENCES

1. Aziz F, Bari A. Impact of ChatGPT on Student Learning Outcomes: A Systematic Review. Southern Journal of Arts & Humanities. 2024 Feb 17;2(1):18-25.
2. Ali JK, Shamsan MA, Hezam TA, Mohammed AA. Impact of ChatGPT on learning motivation: teachers and students' voices. Journal of English Studies in Arabia Felix. 2023 Mar 18;2(1):41-9.
3. Hamid SM, Fadila I. Analysis of the impact of using chat gpt on student learning motivation. Проблемы инновационного и интегративного развития иностранных языков в многоязычной среде. 2024 May 20:230-8.
4. Abdulhajar E, Wahyusari A, Nevrita N, Irawan D, Zaitun Z, Sartika D, Hasyim T. Students' Acceptance of ChatGPT Technology: A Study of Its Positive and Negative Impacts on Academic Ethics and Learning Performance. InSHS Web of Conferences 2024 (Vol. 205, p. 07003). EDP Sciences.
5. Zayoud M, Oueida S, Pauly AW, Ionescu S. Impact of Chatgpt on Education: Challenges and Opportunities. InInternational Conference of Management and Industrial Engineering 2023 Dec 19 (Vol. 11, pp. 75-82).
6. Basha JY. The Negative Impacts of AI Tools on Students in Academic and Real-Life Performance. INTERNATIONAL JOURNAL OF SOCIAL SCIENCES AND COMMERCE1. 2024;3:1-6.
7. Assegaf S, Qonitatillah M, Prameswari NN, Kurniawan B, Ratnawati N. THE IMPACT OF USE OF CHATGPT ON CHANGES IN STUDENTS'LEARNING BEHAVIOR. International Journal of Geography, Social, and Multicultural Education. 2024 Jun 1;2(1):49-58.
8. Pawar PP, Salve KB, Patil RR. " Impact of ChatGPT on Student's Education: A Comprehensive analysis of positive and negative effects". Journal of Advanced Zoology. 2023 Dec 4;44.
9. Shoufan A. Exploring students' perceptions of ChatGPT: Thematic analysis and follow-up survey. IEEE Access. 2023 Apr 19;11:38805-18.
10. Zhou L, Li J. The impact of ChatGPT on learning motivation: A study based on self-determination theory. Education Science and Management. 2023;1(1):19-29.
11. Deshmukh R, Yannawar P. Deep learning based person authentication system using fingerprint and brain wave. International Journal of Computing and Digital Systems. 2024 Feb 1;15(1):723-39.
12. Pabreja K, Pabreja N. Understanding College Students' Satisfaction With ChatGPT: An Exploratory And Predictive Machine Learning Approach Using Feature Engineering. MIER Journal of Educational Studies Trends and Practices. 2024 May 8:37-63.
13. Kanbul S, Adamu I, Mohammed YB. A Global Outlook on AI-Predicted Impacts of ChatGPT on Contemporary Education. SAGE Open. 2024 Aug;14(3):21582440241266370.
14. Hasan N, Polin JA, Ahmmed MR, Sakib MM, Jahin MF, Rahman MM. A novel approach to analyzing the impact of AI, ChatGPT, and chatbot on education using machine learning algorithms. Bulletin of Electrical Engineering and Informatics. 2024 Aug 1;13(4):2951-8.
15. Deshmukh R, Yannawar P. AVAO Enabled Deep Learning Based Person Authentication Using Fingerprint. InFirst International Conference on Advances in Computer Vision and Artificial Intelligence Technologies (ACVAIT 2022) 2023 Aug 10 (pp. 327-346). Atlantis Press.

Next-Gen Predictive Maintenance: A Machine Learning Model for Equipment Reliability with SAP Smart Predict

Vijay Shelake
Department of Computer Engineering
Fr. Conceicao Rodrigues College of Engineering
University of Mumbai
Mumbai, Maharashtra
✉ vijaysnew12@gmail.com

Chaitali Mhase
Department of Computer Engineering
Saraswati Education Societys Group of Institutions
Faculty of Engineering, University of Mumbai
Mumbai, Maharashtra
✉ Chaitali.mhase@gmail.com

ABSTRACT

Predictive Maintenance for Equipment Reliability using Time Series Forecasting with Systems, Applications, and Products in Data Processing (SAP) Analytics Cloud Smart Predict in today's industrial landscape plays a critical role in ensuring uninterrupted operations and minimizing costly downtime. This research explores the application of next-generation predictive maintenance techniques, specifically time series forecasting, to enhance equipment reliability in an industrial setting. Leveraging the capabilities of machine learning, the research aims to proactively identify potential equipment failures, optimize maintenance schedules, and improve overall operational efficiency. The methodology involves the collection and preparation of historical data related to equipment performance, maintenance records, and relevant operational parameters. This research emphasizes on the significance of feature engineering to enhance the accuracy of the predictive models, considering factors such as operating conditions, temperature, and usage patterns to yield a smart machine learning model for equipment reliability.

***KEYWORDS** : SAP, Machine learning, Plant maintenance, Predictive maintenance, Reliability maintenance.*

INTRODUCTION

Next-generation (Next-Gen) strategic priorities for organizations in sectors like manufacturing, healthcare, utilities, and other asset-heavy industries focus on plant maintenance and asset reliability. The smooth operation of production lines, consistent power distribution, and round-the-clock customer service are critical to maintaining performance and competitiveness in these sectors. The significance of maintenance practices is highlighted by the potential consequences of equipment failure in industries such as hospitals and diagnostics, where improper maintenance can severely disrupt essential services [1].

Some companies engage in manufacturing directly at project sites due to logistical difficulties or contractual obligations.

These organizations are responsible for maintaining not only their own assets but also those of their clients. Inadequate maintenance in these situations can result in financial losses, damage to reputation, and pose risks to human safety and the environment [1][2].

Unexpected equipment failures often lead to an increase in corrective maintenance and repair work, resulting in unforeseen costs, reduced equipment longevity, lower Mean Time Between Failures (MTBF), and more frequent breakdowns [1][3].

Large and small industries suffer a lot in all aspects if they do not use information systems. They can automate their work culture, employ control functions while dividing the responsibilities by adopting ERP platforms including SAP, for production, maintenance, procurement, invoicing and payment transactions and so on [1][2][4][5].

Based on data published by global ERP, SAP held a 24 percent share of the global ERP market in 2013. The implementation of SAP applications facilitates the alignment of diverse business processes within an organization, resulting in enhanced integration across functions and significant gains in efficiency and productivity. The corrective, preventive, predictive, and shutdown maintenance are the primary strategies of asset maintenance [1][6][7].

The SAP Plant Maintenance module manages all these areas, along with their relevant data and resources. Additionally, it offers the ability to integrate cost control and procurement activities into maintenance operations smoothly [1][2][9]. However, despite having the necessary data and monitoring tools, the SAP application falls short in predicting the reliability of equipment.

While SAP's decision-making and analytical capabilities are somewhat restricted, leveraging data mining and machine learning (ML) techniques on the data stored in SAP can greatly improve its decision-making ability. ML algorithms have proven to be highly effective in various industries, including medicine [10], operations [11], and finance [12], where they have addressed complex business challenges.

However, there is limited research that integrates data from maintenance processes in information systems with ML algorithms to improve equipment reliability in asset-heavy industries. The potential of integrating ML with SAP is immense, as SAP is used by 60% of Fortune 2000 companies and nearly 80% of Fortune 1000 companies [13].

In light of this, the author proposes an equipment reliability prototype—a cyber-physical system enhanced with artificial intelligence (AI). This system utilizes historical maintenance data, including spare parts usage, preventive maintenance, breakdown records, and other monitoring variables available in SAP. Recent advancements in SAP's design and architecture support this prototype proposal. The shift to a cloud-based platform with SAP HANA [14] facilitates faster decision-making and analysis by enabling the processing of vast amounts of data stored in memory databases at high speeds. This capability is especially valuable for creating models to assess equipment reliability, as it enables rapid access to extensive historical maintenance data required for reliability evaluations [15].

The proposed research introduces a innovative model for equipment reliability that focuses on minimizing failures, improving safety in the workplace, and generating cost savings. This model serves as a decision support tool, assisting maintenance planners in executing corrective actions and taking preventive steps to avoid equipment failures.

LITERATURE REVIEW

Numerous researchers have focused on optimizing various plant maintenance processes [1]. Organizations typically adopt a range of maintenance strategies, such as preventive, breakdown, and condition-based methods, each offering unique advantages and drawbacks [8].

Unforeseen equipment failures can lead to unexpected maintenance expenses, disrupt operations, and hinder productivity. When equipment downtime occurs without warning, both time and financial losses are incurred. The construction industry, for instance, clearly illustrates the negative effects of such unplanned downtimes [15]. In general, the costs associated with reactive maintenance due to unanticipated failures can be six to eight times higher than those linked to preventive or predictive maintenance approaches [16].

The likelihood of equipment failure can be significantly reduced by adhering to the manufacturer's recommended maintenance schedule. Planned maintenance can be based on time, performance metrics, or a combination of both. This approach involves performing a set of prescribed maintenance tasks and overhauls once the equipment reaches a specified runtime or operating hours [17].

However, the maintenance tasks scheduled as part of preventive maintenance often may not be necessary, or they might fail to anticipate future breakdowns. These procedures are typically derived from the manufacturer's maintenance manual. A key drawback of preventive maintenance is that it follows a fixed schedule based on time and running hours, disregarding real-time changes in the equipment's condition [18].

Condition-based monitoring aims to predict equipment failure by detecting abnormal readings in critical parameters like vibration, temperature, and pressure before failure occurs. These readings, captured through sensors, can be leveraged to improve manufacturing processes and optimize Cyber-Physical Systems, ensuring more efficient and reliable machine performance [19].

Several studies have explored the creation of predictive maintenance frameworks designed to monitor equipment health in production lines. One such study applied data mining techniques to an F-18 aircraft, refining maintenance protocols [20]. The resulting model integrated diagnostic codes and repair data to create a predictive maintenance strategy aimed at extending aircraft operational hours while reducing the risk of failure. Another model proposed by Liao et al. suggested that when equipment hits a specific maintenance threshold, it should undergo repairs to restore functionality [17]. Pan et al. recommended that equipment maintenance life, spare part history, and age should be considered to predict degradation rates and inform maintenance strategies [17].

Clustering, an unsupervised learning technique that groups similar data, has been used in risk analysis and market segmentation. However, few studies have explored the application of clustering algorithms to predict equipment reliability [18].

In organizations using the SAP application, plant maintenance not only manages various processes, such as breakdowns, preventive tasks, and condition-based measures, but also stores vital master and transactional data related to equipment. Furthermore, other SAP modules, like Incident Management, help improve workplace safety and reduce risks to people and property [2].

The advent of high-speed computing and integrated memory databases has revolutionized research analytics, enabling real-time data processing. With in-memory databases like SAP HANA, SAP applications can analyze equipment reliability and monitor equipment health by building machine learning-based classifiers using maintenance data stored within the system [14].

As next-generation cloud-based Cyber-Physical Systems continue to dominate industrial automation, integrating data from SAP, SCADA, and DCS systems with machine learning algorithms and service-oriented architecture (SOA) is becoming essential. This technological evolution, characterized by highly networked equipment, supports the decision-based cyber-physical reliability model proposed in this study [19].

RESEARCH OBJECTIVE AND HYPOTHESIS FORMULATION

Reliability maintenance serves as a crucial strategy in preventing unforeseen equipment breakdowns. The researcher's success lies in failure prediction within the construction industry using Time Series Models. Typically, equipment exhibits specific signals or symptoms, often recorded as condition-based measurements through sensor networks before experiencing a breakdown.

Research indicates that reliability models can significantly reduce maintenance costs compared to other maintenance regimes.

While data mining integration with predictive maintenance systems is common, there's a gap in integrating ERP applications with ML algorithms for comprehensive equipment reliability prediction. Leveraging this gap, our Cyber-Physical equipment reliability model aims to enhance workplace productivity, safety, and reduce equipment downtime with research methodology as depicted in figure 1.

Research Problem

The research problem stems from the limitations of current maintenance processes in ERP applications like SAP. While SAP effectively handles preventive maintenance processes, managing unplanned maintenance scenarios and providing insights into equipment reliability remains challenging. The proposed model aims to integrate historical maintenance data with real-time condition-based measurements using both unsupervised and supervised AI techniques for more accurate reliability predictions.

Hypothesis Formulation

The hypothesis posits that integrating machine learning algorithms with data from corrective, preventive, and condition-based maintenance processes can effectively predict equipment breakdowns. Building on previous studies, the author highlights essential data elements, including preventive and corrective maintenance logs, spare parts usage, condition monitoring metrics, and the age of equipment, and maintenance timelines, as vital components for building the reliability model. To verify these findings, the researchers undertake several steps, including building a data model, classifying data sets, The proposed model utilizes AI techniques for grouping data into clusters, followed by evaluating the results using supervised ML algorithms.

Model Benefits and Application

The model offers multiple advantages:

- Continuous Learning: It improves over time by incorporating new data and adjusting to evolving conditions.
- Versatility: It can be added with sources, including or third-party software or SAP databases.
- Quick Deployment: Direct integration with SAP production databases minimizes downtime and enhances testing efficiency.
- Scalability: The algorithms designed for SAP applications can be modified for use with other ERP systems.

The equipment reliability model has the potential to function as a cyber-physical system, integrating data structures and features specific to various types of equipment, like pumps or compressors. It can be developed on a cloud-based platform, utilizing SAP functionalities for real-time monitoring and prediction of equipment status and performance. This integration allows for prompt corrective actions or inspections, potentially utilizing robotics for immediate control in handling equipment reliability responses.

The model proposes close integration between SAP functions, Business Intelligence, and maintenance departments within organizations, with each function having specific responsibilities. SAP handles database integration and deadline monitoring, maintenance departments propose relevant features for equipment reliability analysis, and Business Intelligence builds data models on cloud platforms, integrating data from various sources using ETL methodologies.

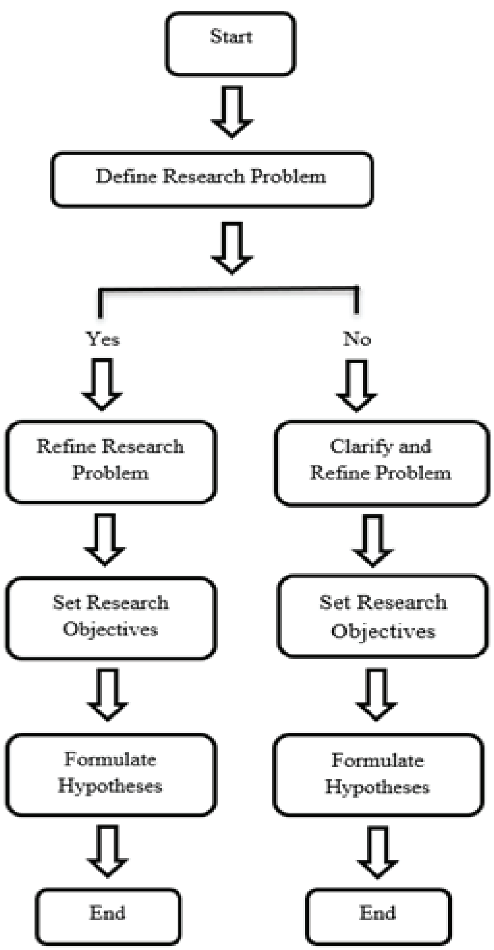


Fig. 1. Equipment reliability model- multiple functions and stakeholders

Thus, our innovative machine learning model with Predictive Maintenance for Equipment Reliability using Time Series Forecasting with SAP Analytics Cloud Smart Predict will be beneficial in next-generation workplaces. Time series forecasting models, including Exponential Smoothing, AutoRegressive Integrated Moving Average (ARIMA) are required to be effectively implemented within SAP Analytics Cloud Smart Predict.

CONCLUSIONS

This research introduces a model for predicting equipment failures by combining machine learning algorithms with maintenance data from the SAP application. The proposed reliability model is versatile and can be applied to other IT systems beyond SAP, as the principles for constructing the data model are largely consistent across platforms. Further research is suggested to enhance adaptability of the equipment reliability model with accuracy by incorporating clustering results with supervised machine learning techniques to classify and forecast equipment performance. The model's data can be sourced from various tables within the SAP system, with a specific focus on pumps. This method would create a system that can foresee equipment failures ahead of time, leading to improved productivity, minimized emission risks, and heightened safety in upcoming workplace settings.

REFERENCES

1. Fraser K, Hvolby HH, Tseng TL. Maintenance management models: a study of the published literature to identify empirical evidence: a greater practical focus is needed. Int J Qual Reliab Manag. 2015;32(6):635-64. Available from: https://doi.org/10.1108/IJQRM-11-2013-0185

2. Kohli M. Incident management with SAP EHS management. Rheinwerk Publishing; 2015. p. 1-132.

3. Bore CK. Analysis of management methods and application to maintenance of geothermal power plants. United Nations University, Geothermal Training Programme; 2008. p. 1-52.

4. Hayen R. SAP R/3 enterprise software: an introduction. McGrawHill/Irwin; 2006. p. 1-192.

5. Yasar A, Ozer G. Determination of the factors that affect the use of enterprise resource planning information system through technology acceptance model. Int J Bus Manag. 2016;11(10):91-108.

6. Columbus L. ERP market share update: SAP solidifies market leadership. Forbes. Available from: https://www.forbes.com/sites/louiscolumbus. Accessed December 2013.

7. Wang Y. Effects of implementation of SAP on management accounting: case: Dongfeng Motor Corporation. University of Applied Sciences; 2013. p. 1-88.

8. Garrido S. Types of maintenance programs. Available from: http://www.mantenimientopetroquimica.com/en/typesofmaintenance.html. Accessed January 2017.

9. Stengl B, Ematinger R. SAP R/3 plant maintenance: making it work for your business. Pearson Education; 2001. p. 1-368.

10. Birkner MD, Kalantri S, Solao V, Badam P, Joshi R, Goel A, Pai M, Hubard AE. Creating diagnostic scores using data-adaptive regression: an application to prediction of 30-day mortality among stroke victims in a rural hospital in India. Ther Clin Risk Manag. 2007;3:457-84.

11. Qiang G, Zhe T, Yan D, Neng Z. An improved office building cooling load prediction model based on multivariable linear regression. Energy Build. 2015; 107:445-55.

12. Yu F. Accounting transparency and the term structure of credit spreads. SSRN Electron J. 2005;75(1):53-84.

13. Missourisandt. ERP share by Fortune 2000 companies. Availablefrom: http://erp.mst.edu. Accessed January 2017.

14. Sikka V, Färber F, Lehner W, Cha SK, Peh T, Bornhövd C. Efficient transaction processing in SAP HANA database. In: Proceedings of the 2012 International Conference on Management of Data - SIGMOD ‘12. ACM; 2012. p. 731-42. Available from: https://doi.org/10.1145/2213836.2213946

15. Prasad Nepal M, Park M. Downtime model development for construction equipment management. Eng Constr Archit Manag. 2004;11(3):199-210. Available from: https://doi.org/10.1108/09699980410535804.

16. Colen IF, Brito J. Building facades maintenance support system. In: Proceedings of The XXX IAHS World Congress on Housing. 2002;9(13).

17. Fu C, Ye L, Liu Y, Yu R, Iung B, Cheng Y, Zeng Y. Predictive maintenance in intelligent-control-maintenance-management system for hydroelectric generating unit. IEEE Trans Energy Convers. 2004;19(1):179-86.

18. Tsang AH. Condition-based maintenance: tools and decision making. J Qual Maint Eng. 1995;1(3):3-17. Available from: https://doi.org/10.1108/13552519510096350

19. Lee J, Bagheri B, Kao HA. A cyber-physical systems architecture for Industry 4.0-based manufacturing systems. Manuf Lett. 2015;3:18-23. Available from: https://doi.org/10.1016/j.mfglet.2014.12.001

20. Young T, Fehskens M, Pujara P, Burger M, Edwards G. Utilizing data mining to influence maintenance actions. In: Proceedings of the 2010 IEEE AUTOTESTCON. 2010;1(5).

GradeGuru: Digital Transformation of Examination Management System

Yogita Mane
Department of Artificial Intelligence and Data Science
Thakur College of Engineering and Technology
Mumbai, Maharashtra
✉ yogita.ydmane@gmail.com

Akshay Agrawal
Department of Information Technology
Universal College of Engineering, University of Mumbai
Vasai, Mumbai, Maharashtra
✉ akshay1661@gmail.com

Neeta Patil
Department of Information Technology
Thakur College of Engineering and Technology
Mumbai, Maharashtra
✉ npat78691@gmail.com

Sanketi Raut
Department of Information Technology
Universal College of Engineering, University of Mumbai
Vasai , Mumbai, Maharashtra
✉ sanketiraut28@gmail.com

ABSTRACT

GradeGuru offers a paradigm shift in the administration of student examinations within educational institutions. Designed to mitigate challenges associated with access rights and privileges, Key functionalities include role based authentication, seat number allocation, user role assignment, mark calculation, hall ticket generation, gazette generation, Marksheet generation, reminder mail for revaluation, locking the dashboard before the result generation aimed at streamlining processes, and ensuring accuracy in examination management. This work underscores the significance of modernizing administrative processes to meet the evolving needs of educational institutions. Through innovative solutions and efficient implementation, it demonstrates the capacity of technology to enhance operational efficiency and facilitate effective management of student examinations.

This research work provides a facility to define role based access to each legitimate user. Admin assigns roles; HoD adds subjects; subject in charges will adds marks, and all the notifications are managed by admin. Overall, the Grade Guru empowers institutions to deliver examinations with greater efficiency and effectiveness, ultimately contributing to the success and advancement of students and educational programs. Proposed system gives 89.16 % Efficiency, 97.05% Accuracy, Misclassification is 2.5%, 89% Precision, and 97.02% recall for 100 different test results of SGPI, CGPI, and Percentage calculation.

KEYWORDS *: Role based access rights, Security, Educational administration, Examination management system, DLOC, ILOC.*

INTRODUCTION

In the Progressively developing field of educational administration, GradeGuru emerges as a vital tool for addressing the complex challenges of managing student examinations. Developed using Electron.js, this desktop application is designed to streamline and enhance the efficiency of exam processes within educational institutions. GradeGuru takes a holistic approach to examination management, offering a variety of innovative features that meet the diverse needs of all stakeholders. A key aspect of its functionality is the effortless assignment of seat numbers to students, fostering organized and structured exam environments that facilitate academic achievement. One of GradeGuru's most notable features is its robust role-based access control system. This empowers administrators to establish specific access rights and privileges for different users, including the Principal, Exam Controller, Heads of Departments (HoD's), and staff members. By providing privileges for each user, GradeGuru strengthens accountability and security within the system.

Moreover, GradeGuru boasts a comprehensive suite of functionalities aimed at simplifying the examination workflow. From generating Excel sheets of student marks for various assessments, including Term-Work (TW) examinations and Internal Assessment Tests (IATs), to calculating averages based on Department Optional Course (DLOC) and Institutional Optional Course (ILOC), GradeGuru streamlines the process of result compilation with unparalleled efficiency. Moreover, GradeGuru offers seamless data import and export capabilities, enabling administrators to easily transfer examination results into Excel files for further analysis or sharing. As educational institutions work to navigate the challenges of examination management, GradeGuru shines as a symbol of innovation, with its extensive features.

LITERATURE REVIEW

In paper [1] emphasized that new electronic systems can be used to organize examination processes. Authors also point out the contemporary downsides faced by the traditional examination management such as the loss of efficiency and the security issues and suggest technological solutions to address them. Through their research they have shown the influence of digital tools on the more efficient processing of exams, the easier access to them, and the guarantee of integrity on the other. The research paper suggested that by incorporating these technologies, the whole procedure including the teacher, as well as the examinees, can be improved a lot.

In the article [2], Authors focused on the significant role of access control in the security of exam processes. She talked about various frameworks of role-based access control (RBAC) and their usage in the field of education, emphasizing the importance of the protection of sensitive data and the guarantee of authorized access. The article is a review of the literature on the implementation of RBAC. The research paper highlights the key advantages of scalability and user compliance as well as the obstacles and drawbacks. Johnson found that a secure implementation of RBAC can extremely well influence the security and integrity of examination management systems.

The research paper [3], prepared by Williams and Brown is a thorough examination where the researchers analyze the operational effectiveness of different examination management systems. One of the factors discussed is the impact of these systems on result processing in educational institutions. The authors ran a very thorough examination of the challenges they faced when manual processing was used and compared them with automated systems. They identified the significant improvements in both efficiency and accuracy owing to automation. The authors stated that the adoption of advanced technologies, such as using AI, can significantly enhance educational outcomes and help human errors in management of the result.

“Examination Management System”[4] by Divya et al. (2021) is one of the papers that are referred to for the comprehensive manual about the exam process, from the generation of the question paper to the scheduling, and then result processing that are emphasized by them in the chapter. They emphasized the importance of automation in the improvement of the processes and the desire to mop up the extra tasks that the system takes on. Hence, they implemented the user interface and accessibility too, which meters educators and students’ academic experience. The research shows that introducing such systems will give way to more dependable and efficient examination management hence be more advantageous to educational institutions.

The automation system, “Exam Cell Automation System,”[5] was the topic discussed in 2017 by Dharshini et al. which was to design an automatic system to improve the exam management process on the way. The paper will allocate to the most important functions as automatic scheduling, registration, and results processing that are needed to solve the common problems of the traditional exam cells. The authors named one of the key features of the system as an error-proof system that will be able to store all the data in a way where the information gets to the right place without any human interference, thus ensuring that both staff and students have a hassle-free experience. The research revealed that conducting automation in exam management can help save time and improve the accuracy of managing exam logistics.

The research paper [6] showed several technological advancements that improve data processing and management in exam systems, highlighting the necessity of effective frameworks to uphold academic integrity and simplify administrative work. The authors offered a comprehensive study of the current systems, pointing out weaknesses and making recommendations for enhancements to maximize efficiency. According to their findings, integrating contemporary technologies can help promote better examination management procedures.

Savakar and Hosur (2015) investigated the use of automated systems to manage examination procedures in their study, “Automation of Examination System.”[7] They examined the issues faced by older methods, such as inefficiencies and inaccuracies in result processing. In order to improve operational correctness and efficiency, the authors offered a comprehensive framework with features including digital grading, automated scheduling, and online registration. Through their analysis, they showed how automation can lead to more dependable results by drastically reducing administrative duties and enhancing the whole exam experience for both teachers and students.

Objectives

The overarching objectives of this research paper are to provide a comprehensive solution that not only provide role based access but also enhance operational efficiency and facilitate effective management of student examinations. List of objectives of proposed research work are as follows.

- To develop a user-friendly desktop application using Electron.js to streamline examination management processes.
- To implement role-based access control to ensure secure and tailored access rights for administrators, educators, and staff members.
- To enable seamless allocation of seat numbers to students, facilitating organized and structured examination environments.
- To provide functionalities for generating Excel sheets of student marks for various assessments, including Term-Work (TW) examinations and Internal Assessment Tests (IATs).
- To calculate averages based on Direct Level of Confidence (DLOC) and Indirect Level of Confidence (ILOC) to automate result compilation.
- To enable easy import and export of examination results into Excel files for further analysis or dissemination.
- To enhance efficiency, Accuracy, and Precision in examination administration to support educational institutions in their mission of fostering learning and academic excellence.

Existing Systems

Traditional examination management in educational institutions often relied upon a combination of manual efforts and outdated software solutions, both of which fall short of contemporary educational demands [8,9]. This conventional approach required extensive labor, especially visible in processes like manual seat allocation, result compilation, and the broader sphere of data management. Such methods not only opened the door to human errors but were also markedly time-consuming and largely inefficient. These shortcomings underscored the lack of scalability and the enormous difficulty in maintaining any semblance of efficiency – essential elements for modern educational operations.

As the need for a streamlined system became increasingly apparent, it was clear that a technologically advanced solution was imperative to revolutionize examination administration. The primary goal was to dramatically enhance efficiency, ensuring that everything from seat assignments to the publication of results was handled swiftly and precisely. Alongside this objective stood the necessity for bolstered accuracy, ensuring that the recorded data and disseminated results were exact and reliable. Transparency also became a major focus, with the aim to establish a system that not only students and faculty but external stakeholders could trust. Such advancements were aimed not only at addressing inherent inefficiencies in the legacy systems but also at preparing educational institutions to handle larger student populations and more complex examinations with ease, thereby propelling them towards a future where managing exams could be seamless and error-free.

Limitations of Existing System

Manual Examination Seat Number Allocation: Seat allocation for examinations is done manually, leading to inefficiencies, errors, and time-consuming processes.

Role-Based Access Control Mechanism: Limited role-based access control mechanisms are in place, resulting in challenges with managing access rights and privileges for different stakeholders, such as the Principal, Exam Controller, HoD’s, and staff.

Result Compilation: Result compilation involves manual data entry and calculation, increasing the risk of errors and delays in generating accurate assessment reports.

Excel-Based Data Management: Examination data, including student marks and assessment details, is managed using Excel sheets, which can be prone to versioning issues, data loss, and lack of real-time collaboration.

Limited User Engagement: There is minimal interaction and engagement between the exam cell department, faculty, and students, hindering communication and collaboration opportunities.

Centralized System Structure: The system structure is highly centralized, with all data and processes controlled by the exam cell department, limiting transparency and accessibility.

Problem Statement

Traditional examination management systems in educational institutions face numerous challenges stemming from manual processes and outdated software solutions [10,11]. These systems often grapple with inefficiencies, errors, and a lack of user engagement. Manual processes, such as seat allocation and result compilation, lead to inefficiencies and delays, while limited access control mechanisms hinder effective management of access rights and privileges. Data management issues arise from using Excel sheets or outdated software, causing versioning problems and data loss [12]. Additionally, minimal interaction between exam cell departments, faculty, and student's hampers communication and collaboration. The highly centralized system structure further restricts transparency, accessibility, and scalability.

The proposed "Grade-Guru" Examination management system offers a paradigm shift in the administration of student examinations within educational institutions. Designed to mitigate challenges associated with access rights and privileges, the Key functionalities include role based authentication, seat number allocation, user role assignment, mark calculation, hall ticket generation, gazette generation, Marksheet generation, sending reminder mail for revaluation, locking the dashboard before the result generation aimed at streamlining processes, and ensuring accuracy in examination management.

PROPOSED SYSTEM

Traditional examination management systems in educational institutions suffer from inefficiencies and delays caused by manual tasks like seat allocation and result compilation. Outdated software results in data management challenges, including version control issues and data loss, insufficient access control, poor communication between departments, and a highly centralized system that limits transparency, accessibility, and scalability, impeding effective collaboration and management.

The proposed GradeGuru aims to revolutionize examination administration processes within educational institutions by introducing modern technologies and user-centric design principles. It seeks to overcome the limitations of traditional methods and outdated software solutions while offering innovative features to enhance efficiency and effectiveness.

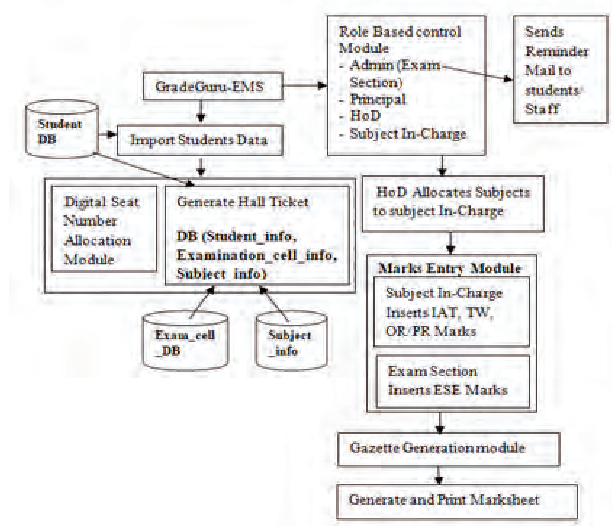


Fig. 1: GradeGuru System Architecture

One key aspect of GradeGuru is as shown in fig. [1], its digital seat allocation module, which automates the process of assigning seats to students for examinations. By eliminating manual seat allocation, this feature reduces errors and streamlines the process, ultimately improving administrative efficiency. Additionally, Grade Guru incorporates a robust role-based access control mechanism, allowing administrators to define access rights for different user roles. This ensures that only authorized personnel have access to sensitive examination data, enhancing security and confidentiality. Another significant feature of Grade Guru is its automated result processing functionalities by automating the calculation and generation of assessment reports; Grade Guru reduces the time and effort required for result compilation. This not only improves efficiency but also ensures accuracy in assessment outcomes. Moreover, Grade Guru integrates a centralized database to securely store examination data, facilitating data management and real-time collaboration among stakeholders. With its user-friendly interface and communication features, Grade Guru promotes seamless interaction between the exam cell department, faculty, and students, fostering a collaborative environment conducive to academic excellence.

Advantages of the proposed system are, it offers a multitude of benefits aimed at enhancing efficiency and effectiveness in examination administration within educational institutions. By automating manual processes such as seat allocation and result processing, Grade Guru will significantly increase efficiency, allowing administrators to save valuable time and resources. Moreover, the automation of key processes will reduce the risk of errors and inaccuracies, ensuring that examination data remains accurate and reliable. The implementation of role-based access control will further enhance security by restricting access to sensitive data, thereby minimizing the risk of unauthorized access and data breaches.

In addition to efficiency and security improvements, Grade Guru will bring about enhanced transparency in examination administration processes. Stakeholders will have access to real-time information, enabling them to track progress, monitor performance, and make informed decisions regarding examination management and resource allocation. The user-friendly interface and intuitive design of Grade Guru will contribute to an enhanced user experience, facilitating easy navigation and interaction for administrators, faculty, and students alike. Furthermore, features for communication and collaboration will streamline interaction between stakeholders, fostering a collaborative environment and improving overall communication within the institution. Lastly, Grade Guru will be designed with scalability and flexibility in mind, ensuring that it can adapt to the evolving needs and requirements of educational institutions over time.

METHODOLOGY

Algorithm for GradeGuru-EMS

1. Create Login and Assign Roles based on Role based access mechanisms to different users such as Principal, Examination Controller, HOD, and Staff.
2. Exam controller will add following details
    a. Add Department and Assign HoD
    b. Add semester
    c. Add Subjects
    d. Add Subject with examination details as Regular/ATKT Examination with credit, Staff name, and Subject_Type (Compulsory subject/ILOC/DLOC)
3. Add Student list (Regular students/ Direct second year admitted student)
4. Create a Seat number for each student for IAT as well as for ESE examination.
5. Generate Hall Ticket
6. HoD will assign subjects to Subject teachers
7. Staff will add Marks of TW, IAT, OR/PR
8. A Controller of Examination will add Marks of End Semester Examination of all the departments.
9. The Examination Controller will create a Gazette for Result, and Marksheet of each student.
10. GradeGuru-EMS will send a reminder of Revaluation if a student got ATKT in the subject.

The detailed algorithms are explained below for defining RBAC, to add subjects, semesters and to create a batch, along with Seat number allocation and based on that generation of hall ticket, and at the end result generation in the form of Gazette and generating Marksheet of each student.

Algorithm for Role based access Control Mechanism (RBAC)

1. Identify all the distinct roles within the system (e.g., Admin, Exam controller, HoD, Staff)
2. Define the specific abilities or permissions associated with each role (e.g., Admin can delete users, User can create posts).
3. During user invitation, assign an appropriate role to the user.
4. Develop an endpoint /get-ability that returns ability statements for a given user
5. Implement logic within the API to fetch abilities based on the user's role.
6. When a user accesses the platform, call the /get-ability API to retrieve their ability statements.
7. Store the retrieved abilities in a local session or cache.
8. Before displaying any feature on the UI, iterate over the stored abilities to verify if the user has access.
9. Hide or disable features that the user does not have permission to access.
10. Capture any action intended by the user (e.g., button click, form submission)
11. Include the user identifier (such as a JWT token) in the API request header.
12. Decode the JWT token to extract the user identifier.
13. Retrieve the user's role from the database using the extracted identifier.
14. Get the ability statements associated with the retrieved user role.
15. Verify if the action attempted by the user is allowed based on their abilities.

16. Compare the action against the user's ability statements.
17. If the user is authorized, proceed with performing the action.
18. If the user is not authorized, create an appropriate error response indicating unauthorized access.

Algorithm to Create Batch, add all semesters, add subject groups

1. Create batch with name and year
2. Read all departments from master list
3. Duplicate all departments in batch from master list
4. Read all semesters from departments.
5. Duplicate all semesters in department from master list department
6. Read all subject groups from master list semesters.
7. Duplicate all subject groups in the semester from master list semester.
8. Read all subjects from master list subject groups.
9. Duplicate all subjects from master list subject's groups.

Algorithm for Seat Number Generation

1. Input semesterId.
2. Retrieve semester details associated with semesterId.
3. Fetch the corresponding department details for the semester.
4. Obtain the batch information for the found department.
5. Gather and store all subject groups related to the given semesterId.
6. Initialize a dictionary to hold lists of students grouped by subject group IDs.
7. Iterate over each subject group to fetch subjects.
8. For each subject, retrieve student details using its ID.
9. Skip processing if no students are found; otherwise, populate the student's dictionary with unique students based on studentId data.
10. Initialize a counter for student seat numbering (studentCount = 1).
11. For each subject group, fetch subjects again.
12. For each subject in the subject group, retrieve student details again.
13. If student details are found, proceed to assign seat numbers to students not already assigned.
14. Calculate the Internal Assessment Test seat number `iatSeatNo` using the department code, semester number, and current student count.
15. Determine the semester type (even or odd) and calculate the End Semester Exam seat number `eseSeatNo` using year, semester type, and department code.
16. Increment the student counter after assigning seat numbers.
17. Update the seat numbers iatSeatNo and eseSeatNo for each student in the database.
18. Repeat Steps 7-17 for all subjects within each subject group until all seat numbers for eligible students are generated and updated.
19. Finish and return or end function execution after processing all subject groups and students.

Algorithm for Hall Ticket Generation

1. Get all subjects of a subject group.
2. Take the exam date of all the subjects from the user.
3. Forward the subject and their respective exam dates to the exporter module.
4. Get all the students enrolled in that subject.
5. Loop through each student.
6. Generate html for the hall ticket with student details and subject details in place.

Algorithm for Seat No Generation

1. Input semesterId.
2. Retrieve semester details associated with semesterId.
3. Fetch the corresponding department details for the semester.
4. Obtain the batch information for the found department.
5. Gather and store all subject groups related to the given semesterId.
6. Initialize a dictionary to hold lists of students grouped by subject group IDs.
7. Iterate over each subject group to fetch subjects.
8. For each subject, retrieve student details using its ID.
9. Skip processing if no students are found; otherwise, populate the student's dictionary with unique students based on studentId data.
10. Initialize a counter for student seat numbering (studentCount = 1).
11. For each subject group, fetch subjects again.

12. For each subject in the subject group, retrieve student details again.
13. If student details are found, proceed to assign seat numbers to students not already assigned.
14. Calculate the Internal Assessment Test seat number `iatSeatNo` using the department code, semester number, and current student count.
15. Determine the semester type (even or odd) and calculate the End Semester Exam seat number `eseSeatNo` using year, semester type, and department code.
16. Increment the student counter after assigning seat numbers.
17. Update the seat numbers iatSeatNo and eseSeatNo for each student in the database.
18. Repeat Steps 7-17 for all subjects within each subject group until all seat numbers for eligible students are generated and updated.
19. Finish and return or end function execution after processing all subject groups and students.

Algorithm for Result Calculation for Regular Examination with Subject Group

1. Input subject GroupId
2. Retrieve the subject group, enrolled students, subjects, and marks by subjects for this group
3. Initialize a list, to store calculated marks for all students.
4. Create a structure for each student to store individual subjects and corresponding marks
5. Identify and populate specific student marks within each subject's data for all students
6. Add structured student data with marks to the initialized list
7. For each student, determine the type of exam (e.g., written, lab) and generate results accordingly
8. Calculate any applicable grace marks for each student and sum up all exam marks
9. Convert computed marks to include additional details like grace marks, and symbols.
10. Calculate overall grades including grade points and credit points for each subject
11. Check and modify the symbols for exams based on failure results in any subject
12. Increment total marks by 10 for students participating in NSS activities
13. Calculate overall SGPI and CGPI for each student
14. Save the final student result data linked to their respective semester and subject group.

RESULTS AND DISCUSSIONS

Role-based Access Control Mechanism

Grade Guru features a comprehensive role-based access control system as shown in Fig.2, that enables administrators to set specific access rights for various user roles. This mechanism ensures that only authorized individuals can access sensitive examination data, thereby enhancing security and confidentiality. As shown in figure the main roles in Grade Guru EMS are Admin, HOD, and Staff. Where Admin has overall access rights, Admin will provide access to all respective HOD's of different departments. HoD will provide access to all subject teachers of their department, where subjects in-charge add marks of their allotted subjects. This system ensures that sensitive data is accessible only to authorized users, thereby enhancing data security and user accountability. By implementing role-based access, institutions can effectively manage permissions and prevent unauthorized access, thereby safeguarding examination integrity and confidentiality.

Fig.2. Role Based Access

Batch creation

The batch creation process is a critical function in managing student groups and their associated data. This feature simplifies the organization of students into batches based on various criteria such as department, year and their admission details such as regular or DSE student as shown in Fig. 3.

Fig. 3: Batch Data

Add Department

Fig. 4 gives the idea to add the Department and acting HoD of the department. Who intern can add Staff in-charge for respective subjects.

Add subject

This feature allows admin to add subjects for the respective semesters. It includes subject name, subject code, associated credits and name of the teaching staff. The proposed system allows you to add subject type whether it is a lab or theory subject as shown in below fig.5.

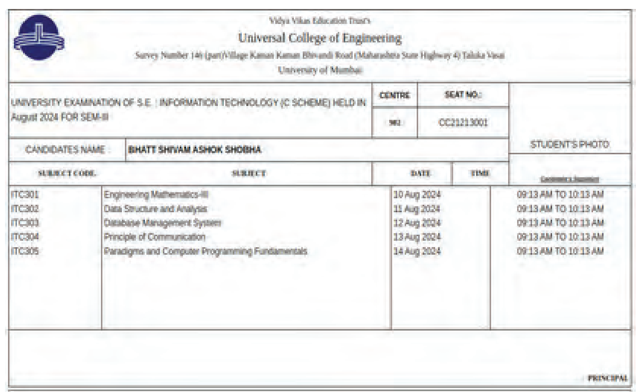
Fig. 4 Add Department

Fig. 5 Add subject with details

Hall ticket Generation

Fig.7 emphasizes the hall ticket generation feature of the proposed GradeGuru system, which automates the creation of hall tickets for students, ensuring they have the necessary documentation for examinations. This process simplifies administrative tasks and ensures timely distribution of hall tickets. Where, exam controller needs to add date of the exam, start time, and End Time of each subject for which student is appearing.

Fig. 7: Hall Ticket Generation



Vidya Vikas Education Trust's
Universal College of Engineering
Survey Number 146 (part)Village Kaman Kaman Bhivandi Road (Maharashtra State Highway 4) Taluka Vasai
University of Mumbai

UNIVERSITY EXAMINATION OF S.E. : INFORMATION TECHNOLOGY (C SCHEME) HELD IN August 2024 FOR SEM-III

CENTRE: 902 | SEAT NO.: CC21213001

CANDIDATES NAME : BHATT SHIVAM ASHOK SHOBHA

STUDENT'S PHOTO

SUBJECT CODE	SUBJECT	DATE	TIME
ITC301	Engineering Mathematics-III	10 Aug 2024	09:13 AM TO 10:13 AM
ITC302	Data Structure and Analysis	11 Aug 2024	09:13 AM TO 10:13 AM
ITC303	Database Management System	12 Aug 2024	09:13 AM TO 10:13 AM
ITC304	Principle of Communication	13 Aug 2024	09:13 AM TO 10:13 AM
ITC305	Paradigms and Computer Programming Fundamentals	14 Aug 2024	09:13 AM TO 10:13 AM

PRINCIPAL

Fig. 8: Hall Ticket Generation

The automation of hall ticket generation reduces manual effort and potential errors, providing a reliable and efficient solution for examination administration; the above fig. 8 shows the final printable hall ticket.

Gazette Generation

The above Fig 9 demonstrates the gazette generation feature of the proposed system GradeGuru, which automates the creation of detailed reports summarizing examination results for all students. This feature ensures that comprehensive and accurate result data is efficiently compiled and readily available for administrative purposes. The automated gazette generation reduces the manual effort involved in preparing result summaries, thereby minimizing errors and expediting the publication process, which enhances transparency and administrative efficiency.

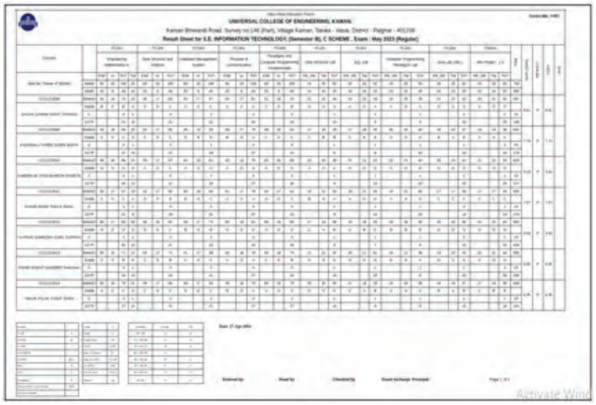
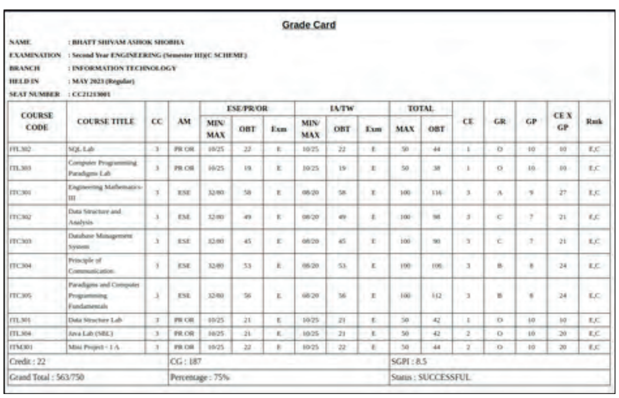


Fig. 9: Gazette Generation

Marksheet generation

Fig. 10 highlights the Marksheet generation functionality, which streamlines the process of creating individual mark sheets for students, detailing their performance in various subjects. This automation ensures accuracy and consistency in Marksheet data, reducing the likelihood of errors associated with manual entry.



Grade Card

NAME : BHATT SHIVAM ASHOK SHOBHA
EXAMINATION : Second Year ENGINEERING (Semester III)(C SCHEME)
BRANCH : INFORMATION TECHNOLOGY
HELD IN : MAY 2023 (Regular)
SEAT NUMBER : CC21213001

COURSE CODE	COURSE TITLE	CC	AM	ESE/PR/OR			IA/TW			TOTAL		CE	GR	GP	CE X GP	Rmk
				MIN/ MAX	OBT	Exm	MIN/ MAX	OBT	Exm	MAX	OBT					
ITL302	SQL Lab	3	PR OR	10/25	22	E	10/25	22	E	50	44	1	O	10	10	E,C
ITL303	Computer Programming Paradigms Lab	3	PR OR	10/25	19	E	10/25	19	E	50	38	1	O	10	10	E,C
ITC301	Engineering Mathematics-III	3	ESE	32/80	58	E	08/20	58	E	100	116	3	A	9	27	E,C
ITC302	Data Structure and Analysis	3	ESE	32/80	49	E	08/20	49	E	100	98	3	C	7	21	E,C
ITC303	Database Management System	3	ESE	32/80	45	E	08/20	45	E	100	90	3	C	7	21	E,C
ITC304	Principle of Communication	3	ESE	32/80	53	E	08/20	53	E	100	106	3	B	8	24	E,C
ITC305	Paradigms and Computer Programming Fundamentals	3	ESE	32/80	56	E	08/20	56	E	100	112	3	B	8	24	E,C
ITL301	Data Structure Lab	3	PR OR	10/25	21	E	10/25	21	E	50	42	1	O	10	10	E,C
ITL304	Java Lab (SBL)	3	PR OR	10/25	21	E	10/25	21	E	50	42	2	O	10	20	E,C
ITM301	Mini Project - 1 A	3	PR OR	10/25	22	E	10/25	22	E	50	44	2	O	10	20	E,C
Credit : 22			CG : 187							SGPI : 8.5						
Grand Total : 563/750			Percentage : 75%							Status : SUCCESSFUL						

Fig. 10: Marksheet Generation

Remainder for evaluation

The "Reminder for Evaluation" as shown in fig.11 functionality in GradeGuru is designed to automatically send reminder emails to students and faculty regarding revaluation deadlines, ensuring that important dates are not missed. This feature enhances the communication and organization within the examination process by providing timely alerts, thus helping to maintain the efficiency and punctuality of the revaluation procedures.

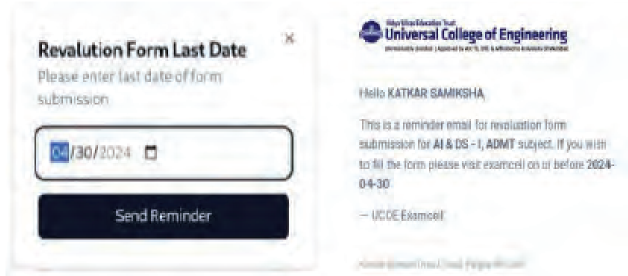


Fig.11: Email Reminder for revaluation

PERFORMANCE METRICS

The overall performance of the Grade Guru is calculated in terms of efficiency, reliability, and user satisfaction. This work is tested using traditional method as well as by using GradeGuru method for one semester in the Universal college of Engineering campus for 5 different departments, 60 students, and on 100 different subjects, the analysis is explained below.

Where,

EMT = Time to Enter Marks

SCT = Time to Share and Compile Sheets

MET = Time for Exam Cell to Re-enter Marks and cross check

T1 = Total Time Consumed to Enter Marks

E1 = Human Error Rate in EMT

E2 = Human Error Rate in MET

T2 = Total Time Consumed after Human Error Rate Consideration with Traditional Method

T3 = Total Time Consumed after Human Error Rate Consideration with GradeGuru Method

Traditional Method:

EMT = 60 Students * 8 seconds = 480 seconds

SCT = 1 Google Sheet * 2 minutes = 120 seconds

MET = 60 Students * 8 seconds = 480 seconds

T1 = EMT + SCT + MET

= 480 + 120 + 480 = 1080 seconds

E1 = 3% = 3% of 1080 seconds = 32 seconds

E2 = 3% = 3% of 1080 seconds = 32 seconds

T2 = T1 + E1 + E2 = 1080 + 32 + 32 =1144 seconds

GradeGuru Method:

EMT = 60 Students X 2 seconds = 120 seconds

T1 = A1 = 120 seconds

E1 = 3% = 3% of 120 seconds = 4 seconds

T3 = T1 + E1 = 120 seconds + 4 seconds =124 seconds

T4= Difference in Time required to execute the operation using Traditional method and GradeGuru

T4 = T2-T3 = 1144-124=1020 Seconds

Improved Efficiency= (T4 /T2) *100 (1)

=89.16 %

From the above analysis, the proposed GradeGuru application system requires less time to perform operation than the traditional method. The proposed system has been tested to generate 500 different Marksheet which have calculations for SGPI, CGPI, and percentage calculations. Performance analysis with respect to confusion matrix is shown in Table 1 below which shows analysis of 100 different test values out of 500 test records.

Table 1: Confusion Matrix

		Actual Result	
		True	False
Predicted Result	True	98	2
	False	3	97

Based on the above table in this research paper Accuracy, Misclassification, precision, and recall is calculated in section below. The Accuracy (All correct prediction) of the proposed system is 97.5%, Misclassification (All incorrect test predictions) is 2.5%, Precision is 98%, and Recall is 97.02%.

Where, Tp= True Positive

Fp=False Positive

Fn= False Negative

Fp=False Positive

Accuracy=(Tp+Tn)/Tp+Tn+Fp+Fn (2)

= 97.5 %

Misclassification= (Fp+Fn)/Tp+Tn+Fp+Fn (3)

=2.5%

Precision= Tp/(Tp+Fp) (4)

=98%

Recall = Tp/(Tp+Fn) (5)

= 97.02%

CONCLUSION

GradeGuru addresses an important progression in the Examination Management System (EMS). By applying new edge cutting Technologies and best practices, the GradeGuru: EMS offers a systematic procedure that addresses the difficulties of conventional examination management systems while giving various advantages to students, staffs, and administrator. Through this automatic process, the GradeGuru further develops efficiency, Accuracy, and precision in seat number allocation, Result calculation, and Database Management. This automation not just saves time and resources of the examination center but also reduces the risk of errors and inconsistencies in examination data and result processing. Furthermore, the Grade Guru promotes transparency and security through centralized database management, role-based access control, and encryption techniques. These characteristics guarantees security, confidentiality in examination data, and accessible only to the legitimate users who have access right, which in-tern enhance the trust and adherence to the data protection rules and regulations.

Overall, the Grade Guru offers a scalable, flexible, and cost-effective solution for educational institutions to modernize their examination management system. By streamlining processes, improving accuracy, and enhancing transparency and security, the Grade Guru empowers institutions to deliver examinations with greater efficiency and effectiveness, ultimately contributing to the success and advancement of students and educational programs. The proposed system has achieved Efficiency by 89.16% by reducing and eliminating error while Marksheet and hall ticket generation with accuracy 97.5%, and 98% precision.

REFERENCES

1. J. Doe and J. Smith, “Enhancing Examination Management Through Digital Solutions”, Int. Journal of Educational Technology, vol. 10, no. 2, 2021.
2. E. Johnson, “Role-Based Access Control Mechanisms in Examination Management Systems”, Journal of Information Systems, vol. 15, no. 4, 2020.
3. M. Williams and S. Brown, “Efficiency and Accuracy in Result Processing: A Case Study of Examination Management Systems”, Int. Journal of Computer Application.

4. Divya H., Nandhini S., Shobana S., and M. Sujithra, “Examination Management System,” Int. Journal of Advanced Research, vol. 9, no. 1, pp. 920-923, Jan. 2021. DOI:10.21474/IJAR01/12370, Available: http://dx.doi.org /10.21474/IJAR01/12370.

5. S. Priya Dharshini, M. Selva Sudha, Mrs. V. Anitha Lakshmi, “Exam Cell Automation System, “Int. Journal of Engg. Science and Computing (2017), Volume 7, Issue no3.

6. S. Vasupongayya, W. Noodam, and P. Kongyong, “S. Vasupongayya, W. Noodam, and P. Kongyong”, World Academy of Science, Engg. and Tech. Int. Journal of Computer and Information Engg. Vol:7, No:7, 2013

7. D. Savakar, R. Hosur, “Automation of Examination System”, Int. Journal of Science and Research, Volume 4 Issue 11 November 2015.

8. R. Jain, A. Modi, I. Kashyap, Kate and R. Baharat “Research Paper on College Management System” Int. Journal of Research Publication and Reviews, Vol 4, no 4, pp 3430-3432, April 2023.

9. D B Heras, D. Otero, and F. Arguello, “An eco-feedback system for improving the sustainability performance of universities,” in Proc. 2011 IEEE Int. Conf. on Virtual Environments Human-Computer Interfaces and Measurement Systems, Ottawa, ON 2011, pp. 1 – 6.

10. Z. Liu, H. Wang, H. Zan, “Design and implementation of online college management system”, 2010 Int. symposium on intelligence information processing and trusted computing. 978-0-7695-4196-9/10 IEEE.

11. Liu, N. and Dong, X. (2008) “Research on the Reform and Security of College Examination”, Journal of Shenyang Arc- hitectural University.

12. J. Castella-Roca, J. Herrera-Joancomarti and A. Dorca-Josa, “A secure e-exam management system,” First International Conference on Availability, Reliability and Security (ARES’06), Vienna, Austria, 2006, pp. 8 pp.-871, doi: 10.1109/ARES.2006.14.

AI-Driven Approaches for Polycystic Ovarian Syndrome (PCOS) Diagnosis: A Comprehensive Review

Pallavi Mane
Institute of Management and Entrepreneurship Development Bharati Vidyapeeth (Deemed to be Univ.)
Pune, Maharashtra

Kirti Mahajan
Institute of Management and Entrepreneurship Development Bharati Vidyapeeth (Deemed to be Univ.)
Pune, Maharashtra

Jaydeep Patil
Department of Mechanical Engineering, Bharati Vidyapeeth College of Engineering
Navi Mumbai, Maharashtra

Yayati Shinde
Department of Mechanical Engineering, Bharati Vidyapeeth College of Engineering
Navi Mumbai, Maharashtra

ABSTRACT

Afflicting about 1 in 14 women of reproductive age globally, Polycystic Ovary Syndrome (PCOS) has seen a concerning rise in India over the past two decades. This condition, a major contributor to female infertility, depends on accurate follicle size and number assessment, traditionally done through ultrasound scans. However, manual analysis is both time-consuming and error-prone due to speckle noise in the images. Enter automated image segmentation, wielding the power of data mining, image processing, and machine learning, particularly convolutional neural networks (CNNs). This technology promises a revolution in PCOS diagnosis by automatically analysing ultrasound images, leading to faster, more accurate follicle detection and ultimately, effective treatment. This review delves into the exciting world of PCOS detection methods utilizing these cutting-edge tools. We explore promising techniques showcased in existing research, highlighting their potential to revolutionize how we identify follicles and diagnose PCOS. From leveraging deep learning algorithms to incorporating spatial information, the future of PCOS diagnosis is brimming with possibilities. By embracing these innovative techniques, we can not only improve the accuracy and efficiency of PCOS diagnosis but also unlock new avenues for research and treatment, bringing hope to millions of women worldwide.

KEYWORDS *: PCOS; CNN; AI; ML; Deep Learning.*

INTRODUCTION

PCOS, which is defined by the ovary containing a large number of tiny follicles, can cause irregular menstruation, high levels of male hormones (androgynism), and infertility [1].Early detection and treatment of PCOS are greatly aided by ultrasound detection. It provides essential information on the number and volume of ovarian follicles, which is crucial in identifying the abnormal follicle development patterns associated with PCOS [2]. In their reproductive years, 5–10% of women are thought to be affected with PCOS, a prevalent gynecological endocrine condition. These patterns, which are typified by irregular follicle counts and sizes during ovulation, are the primary indicator of the disease [3]. We can assist women in managing their symptoms and enhancing their general health and well-being by increasing the precision and efficacy of PCOS diagnosis, especially with the use of cutting-edge imaging methods and individualized treatment program. There are still issues in diagnosing polycystic ovarian syndrome (PCOS), even with its widespread occurrence. It might be challenging to accurately identify ovarian follicles using standard manual techniques, which is a crucial step in the diagnosis of PCOS. This deficiency limits our comprehension of the illness and its effects on women's health in addition to posing serious obstacles to early identification and efficient treatment. In order to allay these worries, this review explores the situation of PCOS diagnosis today, scrutinizing the limitations of manual follicle evaluation and considering the possibilities of new diagnostic techniques and instruments. [5]. The difficulties and possibilities of automated PCOS detection utilizing cutting-edge technology like deep learning and machine learning are compellingly summarized here. Through their ability to highlight important ideas and increase reader engagement, the images support the text.

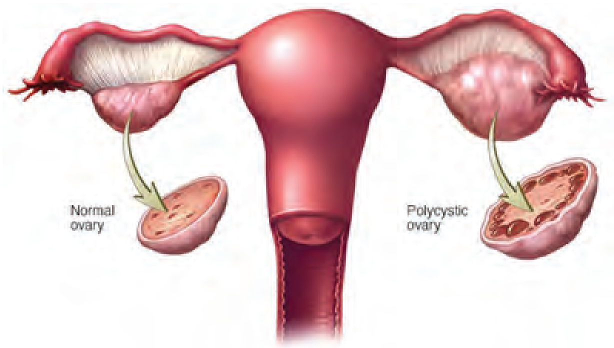


Fig. 1 Normal and polycystic ovary

LITERATURE REVIEW OVER VARIOUS METHODS ON POLYCYSTIC OVARY SYNDROME (PCOS) DETECTION.

Data mining Method

The study by Vikas et al. (2023) investigated the possibility of employing real-time survey data and data mining techniques to forecast the diagnosis of PCOS. Correlations between patient characteristics and possible PCOS symptoms were the main focus of their study.The writers were able to pinpoint important PCOS symptoms, such irregular menstruation and abundant body or face hair, by using the Apriority algorithm. This information provided preliminary understanding for a potential PCOS prediction using patient self-reported data.Nevertheless, the authors admit that a conclusive diagnosis cannot be reached by only using data mining for symptom analysis. For a precise PCOS confirmation, they stress the significance of further clinical exams and testing..

The authors Palvi Soni[8] provides a thorough overview of Polycystic Ovary Syndrome (PCOS) in her survey study, exploring its symptoms, possible causes, and current data mining tasks and techniques used for PCOS diagnosis and treatment. Although the paper offers a useful summary of the state of the art, its main focus is on current approaches. A interesting proposal for a unique data mining approach to increase classification accuracy and find variables that positively contribute to PCOS development is presented in the section on future research. To evaluate its possible influence and contribution to the field, further information on the particular algorithm design and suggested assessment measures is necessary.

The authors Palvi Soni, Sheveta Vashisht of the paper [9], The research combines data mining tasks and techniques to identify a critical gap in existing prediction approaches, while providing a detailed overview of PCOS symptoms, reasons, and dealing choices. Although these methods are helpful, the author contends that they may not be very effective in predicting PCOS. Future study on ultrasonic sound image analysis using image segmentation is suggested as a solution to this. With this novel method, it should be possible to classify diseases clearly according to severity, which should result in more precise and effective PCOS prediction.

Fuzy logic Method

The authors S. Rethinavall, Dr. M. Manimekalai in the paper [10], Taking into account the menstrual cycle type, this study investigated the predictive power of Neural Fuzzy Rough Set (NFRS) in relation to Information Gain for PCOS in patients with and without obesity. PCOS and menstrual cycle type were shown to be positively and significantly correlated. During preprocessing, the NFRS approach that was suggested resulted in a smaller dataset size. This might bolster the validity of the method because the chi-square statistic was higher than the crucial value, which means that the null hypothesis was rejected. The present investigation has notably confirmed the favorable link that exists between menstrual cycle type and PCOS.

Aarti M. Parekh , Prof. Nidhi B. Shah . in the paper[11], n this study, the type of menstrual cycle was taken into consideration while comparing the predictive power of Neural Fuzzy Rough Set (NFRS) against Information Gain for PCOS in individuals with and without obesity. Menstrual cycle type and PCOS were shown to be positively and significantly correlated. Fascinatingly, the preprocessing of the suggested NFRS approach resulted in a smaller dataset, which may increase its validity because the chi-square statistic was higher than the crucial value, rejecting the null hypothesis. Notably, our investigation confirmed the favorable association between PCOS and menstrual cycle type.

The authors K. Meena, M. Manimekalai, S. Rethinavalli in the paper [12], In order to increase feature selection and classification accuracy for intricate tasks, a unique approach is proposed in this study. Rather than addressing these two issues independently, the authors present a unique method that combines the use of a specialized Artificial Neural Network using Fuzzy Neural Rough Subset estimate. This cohesive design makes use of the advantages of both ANN, which is renowned for its potent classification skills, and NFRS, which is excellent at finding pertinent aspects. The suggested algorithm, which combines several approaches, promises to perform better than conventional approaches that handle feature selection and classification separately. This is not only a novel method. The authors speculate that by using this novel technique, feature filtering and classification accuracy may increase, providing opportunities for more effective and economical analysis in a variety of domain

The author Ashika Raj in the paper [13], suggests making the ima12e clearer and higher quality. The final image may be effectively clustered using the fuzzy c-means technique. Finally, the cysts are located with the assistance of groups. The algorithm's efficacy is contingent on the values of Mean Square Error and Peak Signal to Noise Ratio Finding naturally occurring groups of data from a large dataset in order to generate a succinct representation of a system's performance is the goal of clustering. The detected follicles inside the ovary are linked to the manual results through gain outcomes.

The authors R. Saranya and S. Uma Maheswari in the paper [14], presented an novel technique for the automated categorization of the ovaries in computerized ultrasound pictures. This method makes use of fuzzy logic for classification, dynamic outlines without edges for separation, and contourlet change for pre-processing. The test results show the strategy's effectiveness when compared to conclusions made by a medical expert.

Machine learning Method

Yasmine A. Abu Adla, Dalia G. Raydan in the paper [15], examined a number of machine learning methods on a sizable dataset by employing statistical evaluation, medical judgment, and cross-validation. The Linear Support Vector Machine method, which has 24 characteristics, was determined to be the best. With a recall of about 80% and precision and accuracy of 90%, this method identified the best way to improve the accuracy and performance of the model.

B Rachana , T. Priyanka, K. Sahana, , in the paper [16], provide information on the number of follicles discovered, appearance, and geometrical criteria including area, perimeter, compactness, smoothness, and accuracy in order to provide a thorough evaluation of the ultrasound picture. The system provides a comparison outcome using the KNN algorithm. The project's future scope can be used to reduce the KNN algorithm's time of execution.

S Bharati, M. Rubaiya,t H Mondal in the paper [17], suggest a data-driven diagnosis for women with PCOS. The ideas of feature selection and machine learning algorithms are used in this. On these ten characteristics, many classifiers are applied. This research demonstrates that logistic regression, random forest, gradient boosting, and RFLR show excellent recall and accuracy values.

The authors G. Vasavi, Dr.S. Jyothi in the paper [18], examines the subject of segmenting and classifying PCOS using diverse machine learning implementations. This study provided a survey on machine learning implementations, including articles drawn from academic publications. It also matched several studies on the use of machine learning implementations in PCOS cases. Having read several articles, I've seen that not many strategies have demonstrated favorable outcomes. Despite the fact that using these strategies has produced high-quality results, many projects will be improved in the future by using our survey task.

Malik Mubasher Hassan, Tabasum Mirza in the paper [19], examined medical facts of patients diagnosed with PCOS based on symptoms, and prominent machine learning methods, such as Naive Bayes SVM, CART, Classification, Logistic Regression, and Random Forest, were applied. The Random Forest algorithm performed the best in PCOS diagnosis, with an accuracy of 96%, followed by SVM with a 95% accuracy, rendering to the performance validation criteria recall, accuracy, precision, and F-statistics.Thus, based on the provided data, it also determined that the Random Forest algorithm is the most appropriate way for PCOS diagnosis.

R. Devi, Adiwijaya, Jondri. in the paper[20], Through the use of machine learning techniques, polycystic ovarian categorization may be examined. In this study, a competitive neural network—one classification technique—is applied. The greatest accuracy, while employing 32 feature vectors, is examined at 80.84%. It can occur because the precision of categorization using Competition neural networks is closely correlated with the amount of vectors of features.

Simi M S , Sankara Nayaki K , Dr. Murali Parameswaran , Dr. Sabine Sivadasan . in the paper [21], looks at 26 factors to find those that are important for the early identification of 8 different kinds of female infertility. After comparing several methods, the author found that the Random Forest method, which offers 88% accuracy for a relatively big medical data of size 965, is the best approach. This aids in the prompt identification and management of infertility issues.

Untari N. Wisesty , Jondri Nasri, Adiwijaya. in the paper [22], found that the PCO follicles may be detected by using the suggested technique. Based on some feature trials, it can be inferred that the accuracy will decrease with the usage of vector features and a higher feature count. However, the accuracy will rise if we utilize the Mean feature, which increases the number of characteristics employed. The precision of Levenberg-Marquardt optimization may be higher than that of Conjugate Gradient-Fletcher Reeves. The greatest accuracy of Posterior Gradients - Taylor Reeve is 87.85% from 13 cells and 16 vector features, whereas the finest performance of Levenberg - Marquardt is 93.925% from 33 neurons and 16 vector features.

Bedy Purnama, Untari Novia Wisesti. in the paper [23], based on an analysis of the ultrasound image's PCO detection method, it is possible to use the suggested approach to identify PCO. 339 follicular pictures were obtained after counting by Means, The amount of en The Kurtosis Tilt, and Variance texture attributes. Classification is the final phase. utilizing the feature vectors obtained from feature extraction, it distinguishes between the characteristics of the PCO and non-PCO follicles. The optimal SVM - RBF Neuron accuracy at C=40. Dataset B, which was derived by KNN-euclidean distance classification on K=5, achieved 78.81%, but dataset A reached 82.55%. This implies that a few texture features may yield more accurate results than a large one.

Adiwijya, B. Purnama , U.Wisesty. in the paper [24], it was determined that PCO can be detected using the region expanding and euclidean distance approaches. Using template 2, 8-neighborhood connectivity, and an alpha threshold of 2, the first experiment's system setup scenario yields the most optimal performance, and euclidean distance is the most optimal quantification technique. We can infer from the second experiment that the best method for detecting PCO is to use a combination of region growth and euclidean distance for follicle measurement. The results of follicle detection success (RR), follicle size error (9MR), and follicle detection error (FR) all show this.

The authors Kiruthika. V, M.M.Ramya in the paper [25], A new approach for the automated identification of follicle in ultrasound pictures. Since the discrete wavelet transform has better spectral and temporal determination and may help despeckle ultrasound pictures, it was used to generate the anticipated k-means clustering. The fully automated identification of the follicles is made possible by the suggested approach. After analyzing 140 ultrasound pictures visually, the algorithm helps the radiologist make an accurate determination on the follicle's existence.

The authors Sharvari S. Deshpande, Asmita Wakankar in the paper [26], Attempted to create an automated follicle recognition and SVM classification system for the diagnosis of Polycystic Ovarian Syndrome (PCOS). Image preprocessing, such as filtering and contrast enhancement, is used to raise the image's quality. A multistage morphological method and the brilliant top-hat transform are used to extract features. Using the clever edge detection method, the picture is binarized and segmented. Utilizing area and eccentricity threshold, important follicles are distinguished from other locations. Support Vector Machine (SVM) method is used for the categorization of all the information (clinical, biochemical, and imaging). The outcomes are compared to manual categorization results and doctor results. The approach has a 95% accuracy rate. Consequently, the automated detection of women with PCOS may be accomplished with this method.

Ranjitha Sitheswaran, Dr. S. Malarkhodi, in the paper[27], Proposed items are discovered and utilized as candidates for follicles. The most likely group of objects constituting follicles is found by searching through the concave hull of potential follicles based on the price map. This procedure makes use of the ovary's shape, which houses follicles, in addition to the spatial connection of follicles.

Yinhui Deng, Yuanyuan Wang. . in the paper [28], examines An automated scheme which is proposed to looks at An automated method is suggested to identify PCOS. The effectiveness of the suggested plan is confirmed by experiments. When compared to measuring by hand, the automated technique has the potential to address issues related to low efficiency, repeatability, and variability.

The author's Palak Mehrotra, Biswanath Ghoshdastida in the paper [29], survey results from the two classifiers were consistent throughout all folds. The clinical parameters of FSH, LH, BMI, and cycle length were used to predict the likelihood of PCOS incidence. Compared to logistic regression, the accuracy of the Bayesian classifier is greater. It also made the case for the necessity of using additional classifiers to increase accuracy. This research proposes a novel method for PCOS screening. In every fold, the two classifiers that were employed produced reliable results. The early identification of individuals who are at a higher risk of developing a disease can be facilitated by the use of probabilistic models by medical professionals. Other classifiers must be used in order to increase this accuracy.

The authors R. Saranya and S. Uma Maheswari of the paper [30], Women's health is severely disrupted by PCOS, or polycystic ovarian syndrome, an endocrine disorder. An array of imperfectly produced eggs in the ovaries is what classifies the misconception. Errors often occur while manually examining PCOS assumptions. In a same vein, a number of scientists have been actively researching the automated detection of PCOS in recent years. The authors of this work use special approaches to survey for follicles in ovarian ultrasound pictures. Image segmentation may be achieved using a wide variety of techniques, as this writing exam has demonstrated. An adaptive morphological filter, area emergence, and wavelet changes are examples of the risky kinds. The flexible morphological filter reduces speckle noise by preventing pixel value changes. A structural component is used in morphological filters. One of the obstacles. Structured filters make use of morphological elements. The fact that the morphological filter modifies the picture measure in relation to the structuring component is one of its drawbacks. The location development computation provides a distinct advantage and excellent segmented results. However, it requires a manual interaction and is semiautomatic. This computation will work better if it uses the automatically seeded locale. Thus, the need for study is ensured by the demonising and segmentation of ultrasonography images to improve the follicle detection rate.

Convolutional Neural Network Method

The authors M Sumathi, P Chitra. in the paper [31], investigates Using segmentation and feature extraction techniques, CNN is utilized as an image classifier and can identify cysts in the dataset. In order to determine if the ovary is damaged and which parameters—such as size, solidity, extent, and perimeter—are specifically affected, the suggested approach helps categorize test data in the dataset. This procedure employs certain input ultrasound pictures as train data.

The authors Palvi Soni, Sheveta in the paper [32], attempted to provide a technique where the author would apply segmentation before CNN in order to get rid of or remove unnecessary data and improve accuracy. Division allows for the separation of data or pictures in order to carefully extract the required precise data. The authors of this study attempt to categorize the ultrasound picture using Convolutional Neural Networks (CNNs) in order to identify whether it is PCO class or not. CNN is utilized since it has been shown to be the best approach for classifying images.

Cahyono, Adiwijaya, Mubarok. in the paper[33], suggested a method that uses convolutional neural networks to automatically extract features as well. With a micro-average f1-score of 100% and a mean of 76.36% on a 5-fold cross-validation, CNN offers the best test performance. The author analyzes how difficult it is to optimize CNN as inadequate weight initialization and an excessive learning rate frequently lead the SoftMax activation to become NaN. We suggest use a higher dropout rate and improved weight initialization in future work.They advise against employing dropout in any layer other than the scoring layer, as dropout occurs in all levels save the scoring layer hence the network breaks more frequently in these cases.

Chandni Panchasara , Amol Joglekar, in the paper [34], By comparison with split photos, the proposed attributes for faulty images are already accessible, making it easy to identify the issue. also looks at a wide variety of picture segmentation techniques. Each has a unique set of benefits and functions. To produce the final segmented picture, the author of this study gives techniques such as Otsu's Method, Anny detection of edges algorithm, and Region expanding algorithm. When used to thresholding, the Otsu approach yields better results than the standard method. Thus, the intended outcome is achieved while taking cost and speed into account.

Deep Learning Method

The authors S Bhosale, L Joshi, A Shivsharan. in the paper [35], This work shows that a more effective follicle segmentation technique may be achieved by combining several segmentation algorithms. Additionally, a Convolutional Neural Network with Deep Convolution may be used as a foundation for the automatic quality assessment of PCOS data with the suggested technique. Technological developments in image processing might help doctors diagnose PCOS earlier and treat patients earlier by eliminating symptoms. If PCOS is discovered later in life, it can have a long-term effect on mental and physical health conditions.

Artificial Intelligence Method

The authors V. Deepika in the paper [36], investigates AI is able to diagnose this illness by "learning" traits from a vast quantity of clinical practice. Self-correcting qualities in AI allow it to become more accurate. This research presents the author's research applications of artificial intelligence

(AI) in PCOS detection, segmentation, and classification. A lot of studies classified ultrasound pictures automatically using artificial intelligence. Artificial intelligence (AI) is achieved by "learning" traits from an enormous amount of data via medical practice to assess disorders. This research shows that AI can accurately and precisely diagnose diseases and delete unnecessary data. The segmentation and classification of ultrasound pictures are among the AI applications for PCOS detection that the authors of this research look at. By examining.

Image processing Method

The authors A. Saravanan, S. Sathiamoorthy in the paper [37], analyzed PCOS using computational imaging, and also identified PCOS utilizing follicles using various classification and visualization approaches. The study of PCOS using statistics and clinical analysis led to the conclusion that the various causes and degrees of PCOS should be determined.

Nazarudin ,S.S.Mokri , in the paper [38], This study reviewed several researchers' suggested picture segmentation methods and algorithms. Several image processing methods have been used to automatically identify follicles in order to diagnose PCOS.

The authors Rabiu1 , Usman , in the paper [39], to investigate a variety of computer-aided techniques for identifying PCOS and follicles in ovarian ultrasound pictures. Certain earlier works are acknowledged and contrasted in their presentations. studies conducted by the author PCOS has been studied and follicles identified using a wide range of image processing techniques, including feature extraction, morphological procedures, segmentation, and image de-noising.

Computational Intelligence Method

The author Dr.V.Krishnaveni in the paper [40], highlighted patient-centered perspectives and provided a roadmap for a clinical prediction model for PCOS, the most prevalent endocrine condition affecting women in reproductive age. Clinical forecasting models are useful for a number of reasons, including screening high-risk individuals for the unnoticed syndrome, projecting future measurements of the syndrome or death, and assisting in medical decision-making. There are many uses for data mining, a discipline that can find patterns in massive datasets. One such use is the creation of prediction models, which have great potential in the healthcare industry. A successful PCOS prediction model may be ensured by closely adhering to the four steps in the road map.

RESULT AND ANALYSIS

Review of proposed research has helped in understanding the various challenges and prospects of detection of PCOS or non PCOS Image and also Proposed research has helped in knowing the methods which are important to develop for improving the detection of PCOS from ultrasound images of ovaries shown in Table 1.

Table 1. Review summery of different methods

LR No.	Methodology	Finding
7	Data Mining	We were able to identify the main symptoms—such as irregular menstruation and excessive face or body hair—and make predictions by using the Apriority algorithm.
8	Data Mining	Reviewed the PCOS condition, its symptoms, and its causes. It also went through the current data mining jobs and methodologies.
9	Data Mining	A variety of data mining jobs and methodologies are also given, including how to predict PCOS; by using these approaches in the future, it will be advantageous to forecast PCOS illness effectively.
10	Neural Fuzzy Rough Set	Examines the kind of menstrual cycle when the patient is fat and when the patient is not obese, comparing the NFRS (Neural Fuzzy Rough Set) with Data Gain.bese and without obese.
11	fuzzy logic	Accurate segmentation is provided by a feed forward neural network studied in conjunction with an automatic threshold. 97.30% accuracy, 88.16% sensitivity, and 98.19% specificity were attained overall. It is capable of distinguishing between simple, complicated, and normal ovarian cysts using fuzzy logic.
12	fuzzy logic	The tasks of feature selection and classification are divided into smaller chunks using a novel method that is based on fuzzy neural subset assessment and an artificial neural network. Rather of doing each task independently, this technique combines artificial neural networks with neural fuzzy rough subset estimation to get better performance.
13	fuzzy logic	Suggests ways to enhance the image's quality and clarity. For the final image, the fuzzy c-means clustering technique works well. Finally, groups assist in the identification of the cysts. The algorithm's efficiency is based on the values of Peak Signal to Noise Ratio (PSNR) and Mean Square Error (MSE).

14	fuzzy logic	Presented an innovative method for the automated categorization of the ovaries in computerized ultrasound pictures. This method makes use of fuzzy logic for classification, dynamic outlines without edges for separation, and contourlet change for pre-processing. The test results show the strategy's effectiveness when compared to conclusions made by a medical expert.
15	Machine Learning	Used crossvalidation, statistical testing, and medical judgment to examine many machine learning algorithms on a sizable dataset.
16	Machine Learning	Detailed ultrasound picture diagnosis is provided by the information provided, which includes the number of identified follicles, texture, and geometric metrics such as area, perimeter, compactness, smoothness, and accuracy.
17	Machine Learning	Presents a data-driven diagnosis of female PCOS. The ideas of feature selection and machine learning algorithms are used in this.
18	machine learning	A comparison of many studies on machine learning implementations used in PCOS illness has revealed that only a few number of approaches have demonstrated robust performance.
19	machine learning	looks analyzes common machine learning methods, such as SVM, CART, Random Forest, Naive Bayes Classification, and Logistic Regression, using clinical data from patients who have been given a PCOS diagnosis based on their symptoms.
20	Machine Learning	A competitive neural network is one of the classification techniques employed. and finds that utilizing 32 feature vectors yields the greatest accuracy of 80.84%. The reason for this is that the accuracy of classification using Competitive Neural Networks is directly correlated with the amount of feature vectors.
21	Machine Learning	Investigates 26 factors and finds pertinent characteristics for the early identification of eight different kinds of female infertility. After comparing several approaches, the author came to the conclusion that the Random forest method, which offers 88% accuracy for a moderately big hospital dataset of size 965, is the best one.
22	Machine Learning	The PCO follicles should be found. It is possible to determine if a vector feature is being used based on certain feature trials. Conjugate Gradient - Fletcher Reeves optimization may not be as accurate as Levenberg - Marquardt optimization. With 33 neurons and 16 vector features, Levenberg-Marquardt achieved the highest accuracy of 93.925%, whereas Conjugate Gradient - Fletcher Reeves achieved the highest accuracy of 87.85% with 13 neurons and 16 vector features.
23	Machine Learning	The suggested strategy may be put into practice for PCO detection, according to the PCO detection system of the ultrasound picture. It obtained 339 follicular pictures by counting texture parameters such as Mean, Entropy, Kurtosis, Skewness, and Variance. SVM - RBF Kernel's maximum accuracy at C=40.
24	Machine Learning	According to the review, 1) the most optimal quantification technique is euclidean distance, while template 2 and 8-neighborhood connectivity produce the most optimal system configuration scenario with the highest performance.Using region growth and euclidean distance for follicle measurement, it can be determined that the most optimal method for detecting PCO is this approach.
25	machine learning	Suggested a cutting-edge technique for recognizing follicles automatically in ultrasound pictures. Because of its superior spectral temporal determination, which aids in despeckling ultrasound pictures, the discrete wavelet transform is used to produce the intended k-means clustering.
26	machine learning	Examines and attempts to create an automated technique for follicle recognition and SVM classification in the detection of Polycystic Ovarian Syndrome (PCOS). The purpose of image preprocessing (filtering and contrast enhancement) is to raise the image's quality.
27	machine learning	Identifies and examines things that are utilized as potential follicles. To find the largest possible set of items as follicles, the convex hull of probable follicles is searched based on the cost map. Follicle spatial connection and the shape of the ovary, which houses follicles, are both used in this process.
28	Machine learning	Looks at An automated system for PCOS detection is suggested. Experiments confirm that the suggested plan is effective. The automated method may be able to address issues with poor efficiency, variability, and repeatability when compared to manual measurements.
29	Machine learning	Consistent results were obtained across all folds by the two classifiers that were employed. The likelihood of PCOS developing was estimated using four clinical features: cycle length, BMI, LH, and FSH. Superior accuracy is obtained by the Bayesian classifier over the logistic regression.
30	Machine learning	A dangerous endocrine disorder that affects women's health is called polycystic ovarian syndrome, or PCOS. A variety of imperfectly produced follicles in the ovaries provide a classification for the misconception. Errors are frequently produced during the manual evaluation of PCOS supposition. Similarly, a number of scientists have been actively researching on the automated detection of PCOS in recent years.

31	Convolutional Neural Network	Employed as an image classifier, CNN can identify cysts in the dataset with the use of segmentation and feature extraction algorithms.
32	Convolutional Neural Network	Cnns are used to categorize ultrasound images into PCO and non-PCO classes. Since CNN has been shown to be the most effective method for classifying images, it is utilized.
33	Convolutional Neural Network	CNNS are a resilient and excellent way to classify ultrasound images into PCO and non-PCO classes. This is because the system automatically extracts features from each picture.
34	Convolutional Neural Network	CNNS use By comparing segmented photos, it is possible to see the issue clearly and propose the proper features on defective images that are already accessible. Investigates a wide range of picture segmentation methods as well. Each has a distinct benefit and function.
35	Deep Learning	The study shows how to combine multiple segmentation algorithms to get a more effective way for segmenting follicles.
36	Artificial Intelligence	AI applications for PCOS detection are demonstrated, including the segmentation and categorization of ultrasonic imaging data.
37	image processing	It has been shown that the best way to identify the various causes and severity levels of PCOS is by statistical and clinical study of the condition, as well as picture processing.
38	Image processing	Examined the algorithms and methods for segmenting images that have been suggested by several researchers. For the purpose of diagnosing PCOS, many image processing algorithms have been used to automatically identify follicles.
39	image processing	Investigates a variety of computer-aided techniques for identifying follicles and PCOS in ovarian ultrasound pictures. studies conducted by the author PCOS has been studied and follicles identified using a wide range of image processing techniques, including feature extraction, morphological procedures, segmentation, and image de-noising.
40	Computational Intelligence	The most prevalent endocrine condition affecting women of reproductive age, PCOS, has a proposed road map for a prediction model. A successful PCOS prediction model may be ensured by closely adhering to the four steps in the road map.

CONCLUSION

This study focused on data mining, machine learning, artificial intelligence, deep learning, and image processing methods while providing a descriptive and conceptual evaluation of all currently available PCOS detection tools. The methods, features, efficacy, analysis process, and results of the current algorithms were given. Furthermore, a brief description of the datasets utilized in these techniques was provided. Potential issues were explored along with the shortcomings of current algorithms. Even though a lot of research has been done to create an effective PCOS diagnosis model, several issues have not been resolved. The disadvantages identified in this research include a limited amount of datasets, imbalanced datasets, low detection rate, lack of further clustering algorithms, and other issues. We want to work develop a larger dataset with balanced data in the future.Furthermore, a great deal of CNN-based optimization must be done. Future studies should present a novel algorithm that makes use of deep learning techniques to accurately and successfully classify the favorable aspects that contribute to polycystic ovarian syndrome.

REFERENCES

1. M. Lemos "Effect of the combination of metformin hydrochloride and melatonin on oxidative stress before and during pregnancy, and biochemical and histopathological analysis of the livers of rats after treatment for polycystic ovary syndrome", Toxicologies and applied pharmacologies, Vol-280, 2014.
2. B. Purnama, F. Nhita, et. al., "A classification of polycystic Ovary Syndrome based on follicle detection of ultrasound images", (ICOICT), 2015.
3. P. Mehrotra,., "Automated screening of polycystic ovary syndrome using machine learning techniques," Annual IEEE India Conference (INDICON), , 2011.
4. E. Setiawati, A. Tjokorda, "Particle swarm optimization on follicles segmentation to support PCOS detection", International Conference on Information and Communication Technology, 2015.
5. M. Alotaibi, A. Alsinan, "A mobile Polycystic ovarian syndrome management and awareness system for Gulf countries: System architecture", (SAI), 2016.
6. H. Kumar "Segmentation of polycystic ovary in ultrasound images", 2nd International Conference on Current Trends in Engineering and Technology (ICCTET), 2014.

7. B. Vikas, B. Anuhya, K. Bhargav, et. al., “Application of the Apriori Algorithm for Prediction of Polycystic Ovarian Syndrome (PCOS)”, Springer Nature Singapore Pte Ltd, 2018.
8. P. Soni, “Review of Polycystic Ovary Syndrome Along with Data Mining”, International Journal for Scientific Research & Development,Vol. 6, Issue 04, 2018.
9. P. Soni, S. Vashisht, “Exploration on Polycystic Ovarian Syndrome and Data Mining Techniques”, IEEE Xplore Part Number, 2018
10. S. Rethinavalli, Dr. M. Manimekalai, “A Hypothesis Analysis on the Proposed Methodology for Prediction of Polycystic Ovarian Syndrome”, (IJCSET), Vol. 6, No. 11, 2016.
11. A. Parekh, N. Shah, “Classification of Ovarian Cyst Using Soft computing technique”, IEEE , 2017.
12. K. Mina, M. Manimekalai, S. Rethinavalli, “Correlation Of Artificial Neural Network Classification And NFRS Attribute Filtering Algorithm For Pcos Data”, International Journal of Research in Engineering and Technology, 2015.
13. A. Raj, “Detection of Cysts in Ultrasonic Images of Ovary”, International Journal of Science and Research (IJSR), 2013.
14. P. Hiremath, J. Tegnoor, “Automated Ovarian Classification in Digital Ultrasound Images,” International Journal of Biomedical Engineering and Technology, Vol. 11, No. 1, 2013.
15. Y. Adla,, M. Charaf, “Automated Detection of Polycystic Ovary Syndrome Using Machine Learning Techniques”, IEEE, 2021.
16. B Rachana , T. Priyanka, K. Sahana, et. al., “Detection of polycystic ovarian syndrome using follicle recognition technique”, Global Transitions Proceeding,Vol. 2, pp. 304-307, 2021.
17. S. Bharati, M. Rubaiya, H. Mondal, “Diagnosis of Polycystic Ovary Syndrome Using ML”, IEEE, 2020.
18. G. Vasavi, S. Jyothy, “Polycystic Ovary Syndrome Detection Using Various Machine Learning Methods-A Review”, Journal of Advanced Research in Dynamic & Control Systems, Vol. 05, pp. 234-239, 2017.
19. M. Hassan, “Comparative Analysis of Machine Learning Algorithms in Diagnosis of Polycystic Ovarian Syndrome”, International Journal of Computer Applications, Vol. 175, No.17, pp. 42-53, 2020.
20. R. Devi, Adiwijaya, Jondri, “Classification of polycystic ovary based on ultrasound images using competitive neural network”, International Conference on Data and Information Science IOP, 2018.
21. Simi M S, Dr. Murali Parameswaran .Exploring Female Infertility Using Predictive Analytic. IEEE 2017.
22. Untari N. Wisesty ,Adiwijaya. Modified Backpropagation Algorithm for PCOS Detection Based on Ultrasound Images. Springer International Publishing 2017.
23. Bedy Purnama, Untari Novia Wisesti, “A classification of polycystic Ovary Syndrome based on follicle detection of ultrasound images,” 2015 3rd Int. Conf. Inf. Commun. Technol, 2015.
24. Adiwijya, B. Purnama , , U .Wisesty,. Follicle Detection on the USG Images to Support Determination of PCOS. IOP Publishing. ScieTech 2015.
25. Kirutika. , M.MRamya.Automatic Segmentation of Ovarian Follicle using K-means Clustering. 2014 Fifth International Conference on Signals and Image Processing, IEEE,2013.
26. Sharwari S. Deshapande, Wakankar Automated Detection of Polycystic Ovarian Syndrome Using Follicle Recognition.IEEE.2014.
27. Raneeta Sitheswaran, Dr. Malarkhodi .An Effective Automated System in Follicle Identification for PCOS Using Ultrasound Images. International Conference on Electronics and Communication System (lCECS -2014),2014.
28. Yinhui, Yuanyuan, Automated Detection of Polycystic Ovary Syndrome from Ultrasound Images. IEEE EMBS Conference Vancouver, British Columbia, Canada, 2008.
29. Palak Merotra, Biswanath Ghoshdastida.Automated Screening of Polycystic Ovary Syndrome using Machine Learning Techniques.
30. R. Saranya and S. Uma Maheswari, “A Literature Review on Computer-Assisted Detection of Follicles in Ultrasound Images of Ovary,” International Journal of Computer Applications.
31. M Sumathi, P Chitra, .Study and detection of PCOS related diseases using CNN. IOP Conf. Series: Materials Science and Engineering (ICRIET 2020) December 2021.
32. Palvi , Sheveta. Image Segmentation for detecting PCOD using Deep Neural Networks. International Journal of Computer Sciences and Engineering, E-ISSN: 2347-2693, March 2019.
33. Cahyono, Adiwijaya, Mubarok,. An Implementation of Convolutional Neural Network on PCO Classification based on Ultrasound Image. Fifth International Conference on Information and Communication Technology (ICoICT). ISBN: 978-1-5090-4911-0 (2017 IEEE.)2017

34. Chandni Panchasara , Amol Joglekar .Application of Image Segmentation Technique on Medical Reports. (IJCSIT) International Journal of Computer Science and Information Technologies, 2015.

35. S Bhosale, L Joshi, A Shivsharan. PCOS (Polycystic Ovarian Syndrome) Detection Using Deep Learning. Volume:04/Issue:01/January-2022 Impact Factor- 6.752, 2020.

36. V. Dipika. Applications of AI Techniques in Polycystic ovarian syndrome Diagnosis. Journal of Advanced Research in Technology and Management Sciences, Volume: 01 ISSN: 2582-3078 November 2019.

37. Saravanan, S. Sathiamoorthy. Detection of PCOS : A Literature Survey. Asian Journal of Engineering and Applied Technology Vol. 7 No. 2, 2018

38. Nazarudin ,S.S.Mokri ,I.N. A. M 2020

39. Rabiu1 , Usman , A Review on Computer Assisted Follicle Detection Technique and Polycystic Ovarian Syndrome Diagnostic Systems, International Journal of Computer Trends and Technology (IJCTT) volume 28 Number 1 ,2015.

40. V.Krishnaveni. A Roadmap to A Clinical Prediction Model with Computational Intelligence for PCOS. International Journal of Management, Technology, And Engineering ,FEBRUARY/2019 ISSN NO: 2249-7455,2019

41. S. Jonard et al., “Ultrasound examination of polycystic ovaries: is it worth counting the follicles?,” Human Reproduction, vol. 18, no. 3, 2003.

42. R. Saranya and S. U. Maheswari, “A Literature Review on Computer Assisted Detection of Follicles in Ultrasound Images of Ovary,” International Journal of Computer Applications, vol. 48, no. 12, June 2012

43. Rabiu, Tekanyi, “A Review on Computer Assisted Follicle Detection Techniques and Polycystic Ovarian Syndrome Diagnostic System,” International Journal of Computer Trends and Technology (IJCTT), vol. 28, no. 1, 2015.

Design and Analysis of Disc Type Magnetic Coupling for the Central Force and Retire Torque

Bharat B. Waghode
Assistant Professor
Atharva College of Engineering
Mumbai, Maharashtra

Ramesh Kulkarni
Principal
Atharva College of Engineering
Mumbai, Maharashtra

ABSTRACT

This paper analyses the retire torque and axial(centre) force of a circular enduring magnetic coupling under balance conditions. The coupling utilizes permanent magnets. The retire torque and central force are de-termined by varying the air space among the two discs as a purpose of geometrical factors. The conclusion demonstrate that an increment in the air gap length leads to a decrement in both retire torque and central force. Experimental results are validated through measurements performed on a prototype.

INTRODUCTION

Magnetic neodymium couplings are applicable to transfer torque after a prime driver to a load de-prived of requiring mechanical connection. Since torque can be spread through a split-up wall, magnetic neodymium couplings are particularly suitable for usage in remote systems. Associated to mechanical couplings, magnetic couplings of-fer several advantages, including self-protection beside overloads (pull-out torque) and the ability to bear misalignment of shaft.

As illustrated in Fig. 1, the centre magnetic cou-pling under study contains of dual discs prepared with segment designed enduring magnets, divid-ed by a minor air space. The neodymium magnets are centrally influenced and settled alternately as northern and south ends. Soft iron yokes are castoff to complete the magnetic field path. Over magnetic interaction, the torque provided to sin-gle disc is transmitted across the air space to the added disc. Unlike conventional couplings, which rely on mechanical exchange besides lubri-cation, magnetic couplings eliminate issues such as dust generation, noise, and frequent mainte-nance. This makes them highly attractive for var-ious industrial applications. In particular, perma-nent axial magnetic couplings are ideal for re-mote systems like vacuums or high-pressure ves-sels. Mechanical couplings are prone to frequent overloads or damage, which can lead to broken parts and operational failures. In contrast, AP-MCs avoid such issues by providing magnetic slip, which acts as a torque-limiting function, and accurately predicts torque characteristics, ensur-ing safer and more reliable operation.

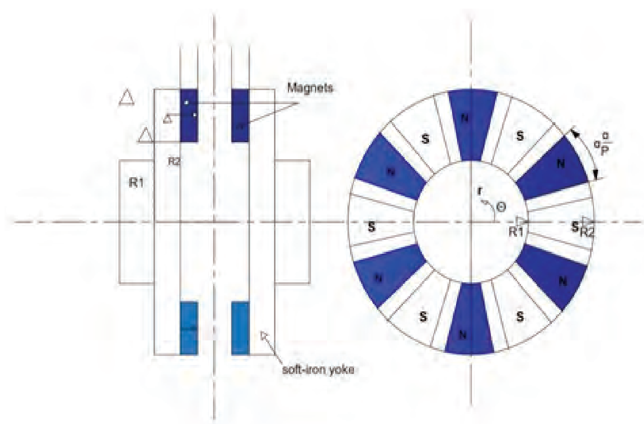


Fig. 1 Magnetic Coupling

This dual-piece permanent neodymium coupling enables contact less torque performance over non -ferrous walls while providing a built-in safety feature: it slips when the supreme torque is sur-passed, defensive mechanisms in the drive train from destruction. Magnetic couplings are widely used in various industrial applications due to their ability to transfer torque from a prime transmitter to a follower without mechanical con-tact. Their capability to transfer torque across separation walls makes axial field magnetic cou-plings especially suitable for remote systems, like vacuums or high-pressure vessels.

A key advantage of these couplings is their max-imum communicable torque(pull-out torque), which inherently provides surplus protection. Ax-ial magnetic couplings comprise two opposing discs prepared with permanent magnets influ-enced in the aligned direction. These magnets are set in alternating northern and south direction, with soft iron yokes castoff to complete the magnetic flux path. The torque useful to single disc is conveyed across an air space to the added disc, with the angular displacement among the discs depending on the transferred torque. How-ever, a notable problem of axial neodymium cou-plings is the substantial axial force among the discs. Permanent neodymium couplings are ex-tensively employed for transmitting torque be-tween rotating components without mechanical touch. They are commonly found in industrial uses, such as centrifugal pumps in the pharmaco-logical and biochemical industries. In precise, coaxial magnetic couplings are well-suited for applications requiring high torque. Understand-ing the pull-out torque of SPMCs is critical, as it acts as a safeguard against overload, ensuring the system's reliability and protection.

CONSTRUCTION OF MAGNETIC COUPLING

Round couplings comprise of divergent discs fit-ted with influential magnets. The torque useful to single disc is transferred across an air space to the added disc. Due to their modest, flat design, these couplings can tolerate angular misalignment of upto to 3° or paralleled disarrangement of up to 6 mm while still transmitting closely full rotat-ing torque. They effectively separate drive-side components from clean or controlled processes, making them one of the naivest and most adapt-able coupling options.

Considerations of Magnetic (NdFeB) Coupling

Table 1: Specification of NdFeB Coupling

	Specification	Value
R1	Internal radius of the NdFeB magnet	25mm
R2	Outside radius of the NdFeB magnet	75mm
Re	Average radius of the NdFeB magnet	50mm
ϕ	Axial(Central) angle of magnets	90^0
g	Air-space (gap)	Variable
h	Thickness of NdFeB Magnet	10mm
p	Quantity of poles	4
Br	Remanence of magnet Surface gauss value	12.8T 4.6 ki-logauss
Magnetic Material - NdFeB		

To assemble, take iron discs with a diameter of 75 mm and divide each disc into four equal an-gular positions (90° apart). Cylindrical neodymi-um magnets are fixed at these positions using industrial glue (e.g., Loctite). The same setup is applied to the second disc. Single disc of the magnetic coupling is attached on the motor shaft, while the other is attached to an adjustable screw using a bearing. This arrangement allows for var-ying the air gap between the discs to optimize performance.

INVESTIGATIONAL SET-UP AND MEASUREMENT

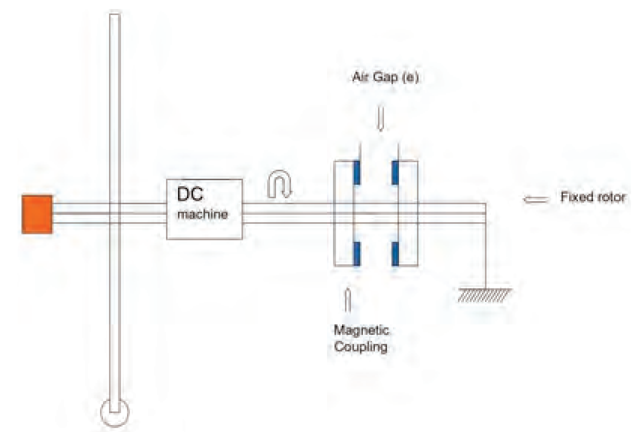


Fig. 2 Block illustration of investigational setup

Fig. 3 Actual investigational set up

Figure 3 illustrates the axially NdFeB coupling positioned on the investigation workbench. The NdFeB coupling is placed among a stable rotor and an electrical device. In this setup, the air space is set to e=20 mm, adjusted by reposition-ing the stable rotor. Figures 2 and 3 present a pic-ture and a schematic diagram, respectively, of the investigation workbench preparation used for static torque extent.

The fixed torque was dignified using weights of varying masses (10g to 100 g, 500 g, 1kg), which were suspended from a rod of length l=0.137m. The rod was protected to one rotor while the other rotor remained fixed. Measurements were then recorded accordingly. Torque intended as the air space of NdFeB coupling is deviates.

Table 2: Result Table

Air Gap 'e' (mm)	Force applied 'F' (N)	Distance to load applied 'r' (m)	Torque (F*r) (Nm)
15	16.68	0.137	2.29
20	7.85	0.137	1.08
25	5.89	0.137	0.81
30	3.92	0.137	0.54
35	2.95	0.137	0.40
40	1.48	0.137	0.21

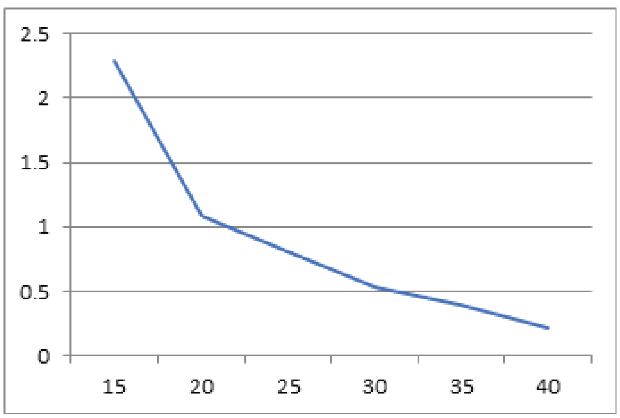


Fig. 4 Air Gap (mm) / Torque (Nm)

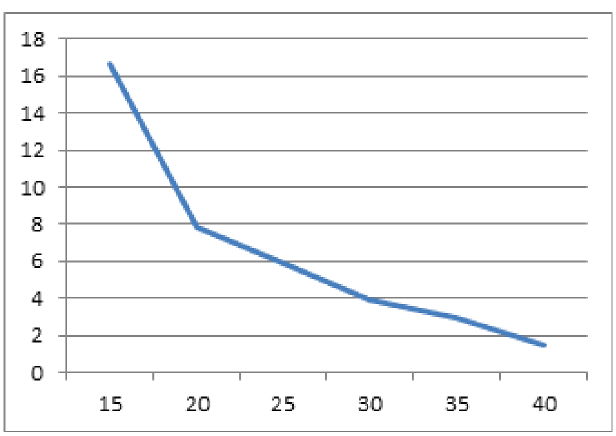


Fig. 5 Air Gap (mm) / Force (N)

CONCLUSION

Experimental analysis reveals a clear variation in torque with changes in the air space. Depends on air space rise from its least to extreme value, both torque and applied force decrease. Specifically, the retreat torque of the NdFeB coupling dimin-ishes rapidly as the space among the magnets grows. For instance, the supreme torque is nearly halved when the air space increases since 2mm to 7mm. Similarly, the maximum axial force is sig-nificantly reduced with an increase in the air space length.

The optimum amount of poles pair is p=4 when the air space length is e = 15mm. It is observed as the extreme centre (axial) forces reduce if the amount of pole pairs rises. Furthermore, varying the number of magnetic poles produces different results, indicating sensitivity to this parameter.

The results and observations confirm that the sys-tem maintains stable performance across changes in air gap length, ensuring consistent drive operation.

REFERENCES

1. Hyeon-Jae Shin, Jang-Young Choi Seok Myeong Jang, and Kwang-Youn Lim, "Design and Analysis of Axial Perma-nent Magnet Couplings Based on 3-D FEM", Department of Electrical Engineer-ing Chungnam National University, Dae-jeon, IEEE Transactions on Energy Con-version 7, July 2013.

2. Thierry Lubin, Smail Mezani, and Ab-derrezak Rezzoug, "Simple Analytical Expressions for the Force and Torque of Axial Magnetic Couplings", IEEE Trans-actions on Energy Conversion 2012.

3. Thierry Lubin, Smail Mezani, and Ab-derrezak Rezzoug,"Experimental and Theoretical Analysis of Axial Magnetic Coupling under Steady-State and Transi-ent Operation", IEEE Transactions on In-dustrial Electronics 2013.

4. Jang-Young Choi, Hyeon-Jae Shin, Seok Myeong Jang, and Sung-Ho Lee, "Torque Analysis and Measurements of Cylindrical Air Gap Synchronous Perma-nent Magnet Couplings Based on Analyt-ical Magnetic Field Calculations", IEEE Transactions on Magnetics, Vol. 49, No. 7, July 2013.

5. Thierry Lubin, and Abderrezak Rezzoug, "Steady-State and Transient Performance of Axial-Field Eddy-Current Coupling", IEEE Transactions on Industrial Electron-ics 2014.

6. R. Ravaud, V. Lemarquand, and G. Lemarquand, IEEE Trans. Magn. 46, 11 (2010)

7. S. Högberg, B. B. Jensen, and F. B. Bendixen, IEEE Trans. Magn. 49, 12 (2013)

8. S. Rhyu, S. Khaliq, C. Baek, and Y. Ha, J. Magn. 25, 64 (2020)

9. Lubin T., Rezzoug A., 'Steady-state and transient performance of axial-field ed-dy-current coupling, IEEE Trans. Ind. Electron., 2015

10. Ogunjuyigbe A. S. O., Jimoh A. A., Ayodele T. R., 'Dynamic and transient behavior of a line start, capacitance com-pensated synchronous reluctance ma-chine', Journal of Electrical Systems and Information Technology., 2018, 7, pp. 843–860.

11. Mingardi D., Bianchi N., 'Line-start PM-assisted synchronous motor design, op-timization, and tests', IEEE Trans. Ind. Electron., 2017, 64, (12), pp. 9739- 9747.

12. Aberoomand V., Mirsalim M., Fesharakifard R., 'Design optimization of double-sided permanent-magnet axial eddy-current couplers for use in dynamic applications, 2019, IEEE Trans. Energy Convers., 34, (2), pp. 909-920.

13. Desvaux M., Le Goff Latimier R., Mul-ton B., Sire S., Ben Ahmed H., 'Analysis of the dynamic behaviour of magnetic gear with nonlinear modelling for large wind turbines', Proc. XXII International Conference on Electrical Machines (ICEM), Sept 2016, pp. 1332-1338.

14. Dolisy B., Mezani S., Lubin T., Lévêque J., 'A new analytical torque formula for axial field permanent magnets coupling', 2015, IEEE Trans. Energy Convers., 30, (3), pp. 892-899.

Linear System Identification from Input-Output Data

Niranjan Samudre
Assistant Professor
Atharva College of Engineering
Mumbai, Maharashtra
✉ niranjansamudre@atharvacoe.ac.in

ABSTRACT

Linear system identification is a critical aspect of control engineering and signal processing, enabling the modeling of dynamic systems based on input-output observations. This paper explores state-of-the-art methodologies and frameworks in linear system identification, including both theoretical advancements and practical applications. Some of the recursive (on-line) algorithms are also studied. Simple demonstrations are performed to support these aspects. The study also highlights challenges in real-world system identification, such as handling uncertainties, noise, and large-scale datasets. Finally, future research directions are proposed, emphasizing the role of artificial intelligence and real-time adaptive modeling.

INTRODUCTION

System identification refers to the development of mathematical models that characterize dynamic systems using empirical data. It plays a crucial role in engineering applications, including control design, fault detection, and predictive analytics. This paper provides a detailed examination of linear system identification methodologies, their mathematical underpinnings, and their impact across various industries.

System Identification is the field of mathematical modeling of systems from experimental data. Some experiments are performed on the system; a model is then fitted to the recorded data by assigning suitable numerical values to its parameter [6]. The system identification is necessary to establish a model based on which the controller can be designed, and it is useful for tuning and simulation before applying the controller to the real system [9].

PROBLEM STATEMENTS

1. Given input output data (u (t), y (t)), generated by a system G, y = G [u], find a system G1 that approximates G and provides an estimate of the size of the approximation error. Compare the output of the simulated model output with the actual output.
2. Compute different parametric models and compute its spectrum with different methods and compare these spectrums with each other.
3. Compute the different models using different approaches and plot the estimated parameters and the true parameters as a function of time for comparison.

SYSTEM IDENTIFICATION

System identification is the process of developing mathematical models that describe the behavior of dynamic systems using input-output data. It is widely used in control engineering, signal processing, and machine learning for modeling and predicting system responses.

The primary goal of system identification is to determine an appropriate mathematical representation of a system from observed data, which can then be used for analysis, prediction, or control design.

Components of System Identification

The system identification process typically involves the following key components:

1. System: The real-world process or dynamic system being studied.
2. Inputs: External signals or excitations applied to the system.
3. Outputs: Measured responses of the system corresponding to the given inputs.
4. Model Structure: A mathematical representation (e.g., transfer function, state-space model) describing the system.

5. Identification Algorithm: The computational method used to estimate model parameters.
6. Validation: The process of verifying the accuracy and generalizability of the identified model using independent data.

Classification of System Identification Methods

System identification methods can be broadly classified into two categories:

Parametric Methods

Parametric methods assume a predefined model structure with unknown parameters that need to be estimated. Some commonly used parametric techniques include:

1. Least Squares Estimation (LSE): Minimizes the sum of squared errors between observed and predicted outputs.
2. Maximum Likelihood Estimation (MLE): Estimates parameters by maximizing the likelihood function based on probabilistic assumptions.
3. Prediction Error Methods (PEM): Uses optimization techniques to minimize the difference between measured and predicted outputs.
4. State-Space Models: Represents system dynamics using a set of first-order differential or difference equations.

Non-Parametric Methods

Non-parametric methods do not assume a fixed model structure and instead rely on data-driven approaches:

1. Impulse and Step Response Analysis: Determines system dynamics based on its response to an impulse or step input.
2. Correlation and Spectral Analysis: Uses statistical and frequency-domain techniques to derive system characteristics.
3. Subspace Identification Methods (SIMs): Extracts system information directly from input-output data using linear algebra techniques.
4. Machine Learning-Based Identification: Applies neural networks and deep learning to model complex, nonlinear systems.

Steps in System Identification

A standard system identification process includes the following steps:

1. Experiment Design: Collecting high-quality input-output data with sufficient excitation.
2. Data Preprocessing: Filtering, normalization, and removing outliers from collected data.
3. Model Selection: Choosing an appropriate model structure (e.g., ARX, ARMAX, state-space).
4. Parameter Estimation: Using optimization algorithms to estimate model parameters.
5. Model Validation: Testing the model against unseen data to assess its accuracy.

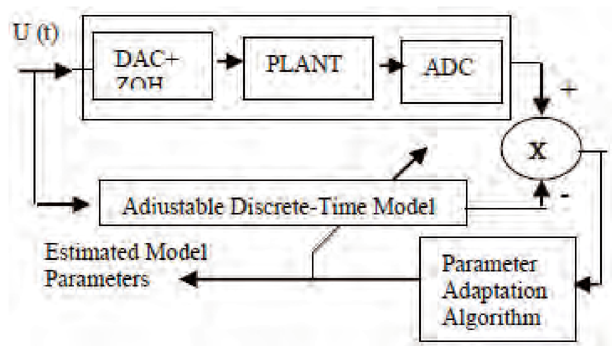


Fig. 1. Parameter estimation principle

The parameter estimation principle for sampled model is illustrated in figure (1). A discrete time model with adjustable parameters is implemented using computer. The error between the system output at instant 't', y (t), and the output predicted by the model $\hat{y}(t)$ (known as the prediction error) is used by a parameter adaptation algorithm (PAA), which at each sampling instant, will modify the model parameters in order to minimize this error.

We may now establish a system identification procedure for SISO systems.

- Select a test input sequence
- Collect experimental data.
- Form the matrix Φ and vector B by processing the I/O data.
- Compute the LS solution.
- Compute the transfer function of the identified system.
- Compute residual error and obtain an estimate of the uncertainty bound.

Assumed that the experiment design and data processing are performed in an appropriate manner, consistent with the limitations of the algorithm.

Finally under implementation we consider various practical issues such as selection of test inputs, fine tuning of the estimation algorithm, computational trade offs etc. The objective of experimental design is to make the collected data set as informative as possible for the models being built from the data. Some of the important points that must be taken into consideration are:

- The input signal must expose all the relevant properties of the system i.e. the input must contain at least as many different frequencies as the order of thee linear model being built.
- A typical good sampling frequency is 10 times the bandwidth of the system.
- In time domain, these suggestions are useful:
- Use binary (two-level) inputs if linear models are being built. This gives maximal variance for amplitude-constrained inputs
- Check the changes between the levels so that the input occasionally stays on one level long enough for a step response from the system to settle, more or less.

IMPLEMENTATION AND RESULTS

Consider the model of the following kind

$$y(t) + a1\ y(t-1)+\ldots+an_a\ y(t-n_a) = u(t-nk) +\ldots+ bn_b\ u(t-nb-nk+1) \qquad (1)$$

We start by forming some simulated data. Estimate the parameters of the model. Compare the output of the simulated model output with the actual output.

The input is generated as a Random Binary Sequence. The input-output is shown in figure (2). A second order model of ARX type is considered. The order of the system is tested and thee parameters of the model are estimated. The output of the simulated model is compared with the actual output for validation as shown in figure (3) and the residuals are shown in figure (4). The chosen structure and the selected structure is given in the results.

2) The input output of the simulated model for the second problem is shown in figure (5). The comparison of spectrums of the model by different methods is shown in the figure (6) and (8). A 5th order AR model is computed by instrumental variable method and modified covariance method and their spectrums are compared as shown in figure (9). The spectrums of the AR model and ARMAX model are compared as shown in figure (7).

3) Output error model of the input output data is computed. The estimated parameters and the true parameters are shown in table (1). The four parameters are plotted as a function of time using forgetting factor approach as shown in figure (10). Figure (11) shows the plot of parameters from ARX model using Kalman filter approach.

Table 1. Results of the estimated parameters

Sr. No.	Parameters	Assumed Parameters (G)	Estimated Parameters (G1)
01	A1	-1.5	-1.4368
02	A2	0.7	0.6370
03	B1	1.0	0.9563
04	B2	0.5	0.6281

Chosen Structure: sNn = 5 5 1
i.e. Orders and delays: na = 5; nb = 5; nk = 1.
Number of parameters to be estimated = 4.
Selected Structure: Nns = 2 2 1
i.e. Selected Orders and delays: na = 2; nb = 2; nk = 1.

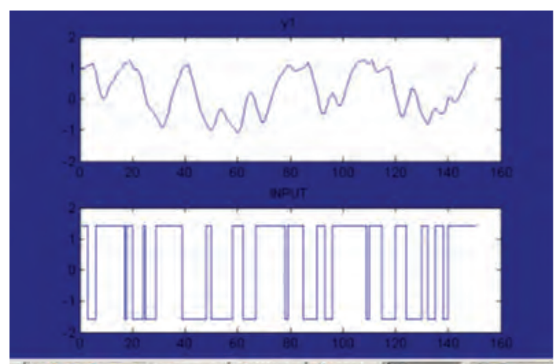


Fig. 2. Input output of the system

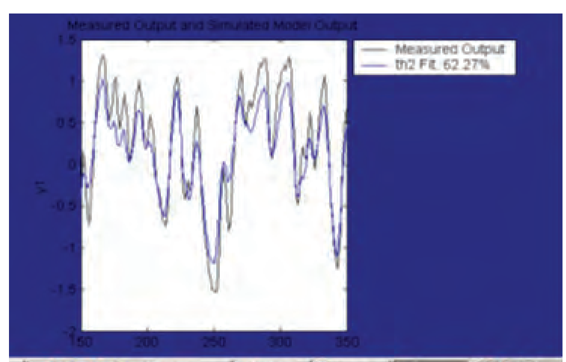


Fig. 3. Comparison of the output of simulated model with actual output.

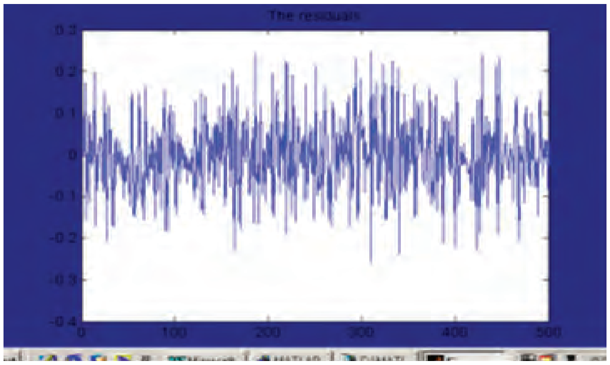


Fig. 4. Residuals

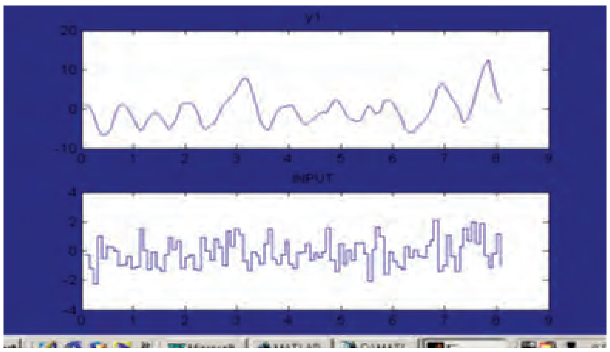


Fig. 5. The input output of the simulated model

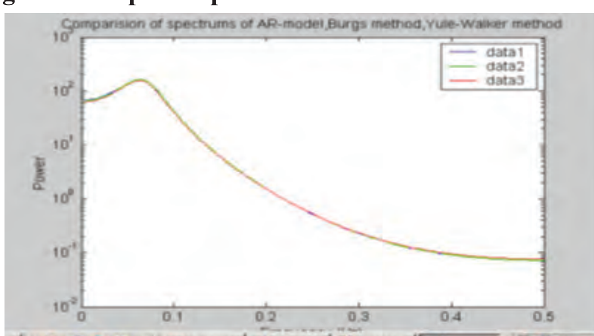


Fig. 6. The comparison of spectrums of the model by Burgs and Yule Waker methods

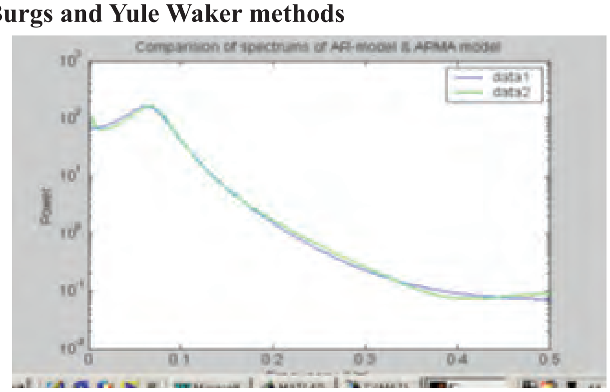


Fig. 7. The spectrums of the AR and ARMAX model

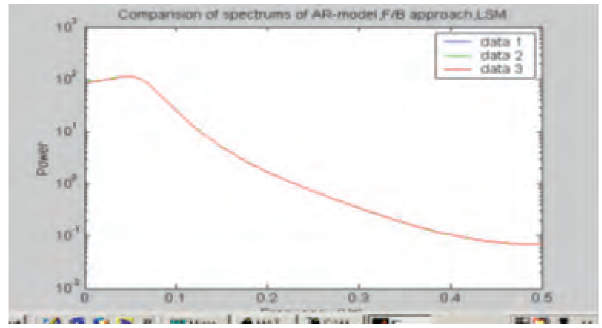


Fig. 8. The comparison of spectrums of the AR model by LSM and f/b approach

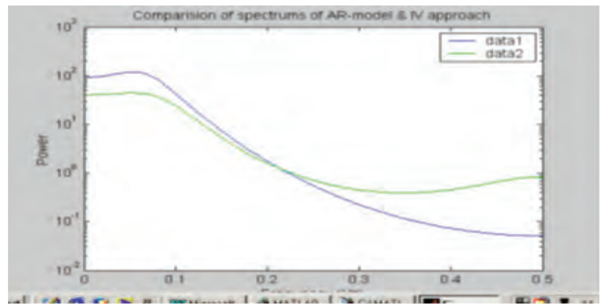


Fig. 9. Comparison of spectrums of AR model by instrumental variable and modified covariance method

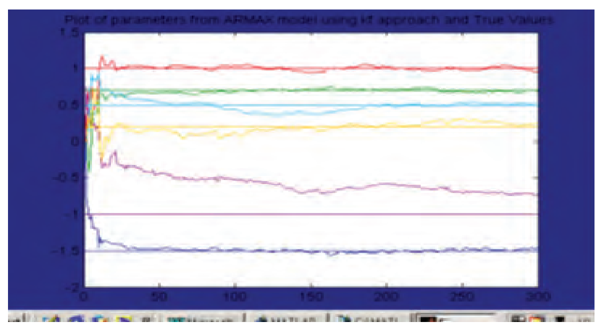


Fig. 10. The plot of parameters from ARMAX model using forgetting factor approach

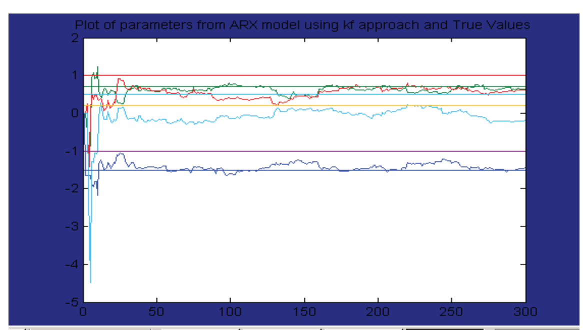


Fig. 11. The plot of parameters from ARX model using Kalman filter approach.

CONCLUSION

This paper provides an in-depth exploration of linear system identification methodologies, discussing their theoretical foundations, practical applications, and emerging trends. Future research should focus on enhancing computational efficiency and robustness while leveraging modern AI techniques for real-time adaptive modeling.

CHALLENGES AND FUTURE RESEARCH DIRECTIONS

- Robustness against noise and uncertainties
- Computational efficiency for large-scale systems
- Real-time adaptive identification
- Integration of artificial intelligence and deep learning
- Emerging trends in cyber-physical system modelling
- Wavelet transforms can be used in the case of Time-varying systems.

REFERENCES

1. G. A. Bekey and G.N. Saridis Identification and system parameter estimation. Proceedings of the Sixth IFAC symposium, Washington DC, USA, 7-11 June 1982.
2. I. D. Landau System Identification and control design using P.I.M. + Software. Prentice-Hall, 1990.
3. Jeng-Ming Chen and Bor-Sen Chen "System Parameter Estimation with Input/Output Noisy Data and Missing Measurements," IEEE Trans. Signal Processing, Vol. 48, No. 6, pp. 1548-1558, June 2000.
4. K. Tsakalis. EEE 480/482.
5. Laszlo Gerencser and David A. Maluf "Identification of Time Varying Stochastic System."
6. L. Ljung, System Identification: Theory for the user. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1987.
7. N. K. Sinha and B. Kuszta, Modeling and Identification of dynamic systems. Van Nostrand Reinhold Company, 1983.
8. S. Salivahanan, A Vallavaraj and C. Gnanapriya, Digital Signal Processing,Tata McGraw-Hill Publishing Company Ltd, New Delhi, 2000.
9. T. Soderstrom and P. Stoica System Identification. Prentice-Hall International (UK) Ltd, 1989.
10. M. Gopal, Digital Control and State Variable Methods, 2nd ed., Tata McGraw-Hill Publishing Company Ltd, New Delhi, 2003.

VOLUME 48 • SPECIAL ISSUE • NO. 1 • JUNE 2025

ISSN 0971-3034

PUBLISHED BY

INDIAN SOCIETY FOR TECHNICAL EDUCATION

Near Katwaria Sarai, Shaheed Jeet Singh Marg,
New Delhi - 110 016

Printed at: Compuprint, Flat C, Aristo, 9, Second Street, Gopalapuram, Chennai 600 086.
Phone : +91 44 2811 6768 • www.compuprint.in