VISVESVARAYA TECHNOLOGICAL UNIVERSITY

“Jnana Sangama”, Belagavi-590018, Karnataka

A Technical Seminar Report

On

“VIRTUAL KEYBOARD BASED ON TRUE-3D OPTICAL RANGING”

Submitted in Partial Fulfillment of Requirements for The Award
of Degree of Bachelor of Engineering

In

COMPUTER SCIENCE & ENGINEERING

Submitted by

Jeethu Pathak
(1SP20CS022)

Under the Guidance
of
Prof. Jayakumar BL
Assistant Professor
Dept. of CSE
SEACET
Bangalore - 560049

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

S.E.A. COLLEGE OF ENGINEERING & TECHNOLOGY

Ekta Nagar, K.R. Puram, Bengalore-560 049

2023-24

VISVESVARAYA TECHNOLOGICAL UNIVERSITY

"Jnana Sangama", Belagavi-590018, Karnataka

S.E.A. COLLEGE OF ENGINEERING & TECHNOLOGY

Ekta Nagar, K.R. Puram, Bengaluru-560 049

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Certificate

This is to certify that Technical Seminar work entitled **"Virtual Keyboard Based On True-3d Optical Ranging"** carried out by **Mr. Jeethu Pathak**, bearing USN **1SP20CS022**, a bonafide student of VIII semester B.E for the partial fulfillment of the requirements for the Bachelor's Degree in Computer Science & Engineering of the **VISVESVARAYA TECHNOLOGY UNIVERSITY**, during the year **2023-2024**. It is certified that all correction/suggestion indicated for Internal Assessment have been incorporated in the report deposited in the department library. The Technical Seminar report has been approved as it satisfies the academic requirements in respect of the Technical Seminar work prescribed for the said degree.

Signature of the Guide	Signature of the HOD	Signature of the principal
Prof. Jayakumar BL	**Dr. Krishna Kumar P R**	**Dr. B. Venkata Narayana**
Asst. Professor	**Prof and Head of Department**	**Principal**
Dept. of CSE	**Dept. of CSE**	

ACKNOWLEDGEMENT

Firstly, I thank the Management late **Shri. A KRISHNAPPA**, Chairman of SEA College of Engineering and Technology for providing necessary infrastructure and creating a good environment.

I would like to thank to our respected **DR BHAGAVANT K DESHPANDE**, Director of SEA College Of Engineering And Technology for the encouragement and support given by him.

I express my deep sense of gratitude to **S.E.A COLLEGE OF ENGINEERING & TECHNOLOGY**, **BANGALORE** that provided me an opportunity in fulfilling my most cherished desire of reaching the goal.

I am thankful to our principal **Dr. B. VENKATA NARAYANA**, who is responsible for creating such a pleasant environment and appreciating my talent in both academic and extracurricular activities.

I am thankful to **Dr. KRISHNA KUMAR P R**, HOD of Computer Science, who inspired me in my work and stood for many of my work.

I whole-heartedly express my gratitude to my guide **Prof. JAYAKUMAR BL** who is responsible for creating such a pleasant environment and appreciating my talent in both academic and extra-curricular activities.

Lastly, I thank my parents, friends, lecturer and staff who provided me the much- needed moral support while pursuing this seminar.

Jeethu Pathak
(1SP20CS022)

ABSTRACT

Face Detection and Recognition, a technique that utilizes various features of the face to verify and identify individuals, is becoming more popular in various industrial applications including security, surveillance, and access control. Due to machine learning's capability to learn complicated patterns from enormous datasets, it's an effective approach for facial recognition. The utilization of machine learning for face recognition is a methodology comprising three distinct stages: face detection, feature extraction, and classification. The first phase involves an algorithm that recognizes the facial features in a given image or video source. Following this, the feature extraction process identifies key geometric and textural characteristics from the detected face. Lastly, the extracted data is run through a classification algorithm to determine its corresponding class - usually ascertaining identity information. To address the challenge of recognizing faces accurately convolutional neural networks (CNNs), support vector machines, and random forests are some examples of machine learning algorithms that have been identified. CNNs are especially promising as they can identify hierarchical representations within images thereby providing better accuracy. With that said even though these modeling techniques show promise they still struggle with problems posed by uncontrolled lighting conditions, pose variations and occlusions. More over like any facial recognition technology needs to meet strict ethical standards because it also brings up privacy concerns. From healthcare to security and entertainment industries, facial recognition powered by machine learning is revolutionizing how we see automation.

CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1: Overview

Touch-typing or machine writing was invented for mechanical typewriters which had the current QWERTY key layout since 1874. While this interface is come to age, it survived because of its many positive aspects. Yet it is not feasible for the ever-smaller computing devices that house ever-more advanced functionalities. New alphanumeric interfaces include numeric keypads augmented with letters as on the cell phone and the Graffiti handwriting characters conditions is one of the most extensively researched, problems in computer vision. We survey the state of the art in alphanumeric input interfaces. After an overview of the related work in the area in section 2, we lay out the general space of interfaces for text input in section 3. In the following section we discuss characteristics and human factors of touch-typing in particular. Section 5 explains the criteria we examined with. Section 6 introduces and compares various touch-typing input methods and devices. We conclude with stating that keyboards – whether virtual or real – are very well suited to the task of alphanumeric input. Since language and its manifestation in sentences, words, and letters is the human's primary means of communication, it has been researched extensively for the human-to computer interaction as well. In this section, we take a look at three main categories Speech recognition, handwriting recognition and sign language. Touch-typing is covered in section Discussed are characteristics and shortcomings. developed or some existing algorithms Probably the most hailed UI, speech recognition (SR) is now at a stage where it can be successfully deployed in limited domains, such as in call centers for customer service or for voice dialing on mobile phones. Under these conditions it provides a highly user Alipay's friendly, unobtrusive, flexible and efficient interface method. But careful estimates suggest that 100% recognition rates for large vocabularies will not be possible in the near future. Noisy environments, speaker particularities (accents, speech impediments) worsen the administration of Virtual typing systems, which has resulted in this technology becoming As Shneiderman summarizes in producing sounds that make up words and sentences occupies parts of the brain that are also used for general problem solving. This is contrary to how for expressions, since this may significantly alter the size and positioning of the landmarks.

1.2: System Overview

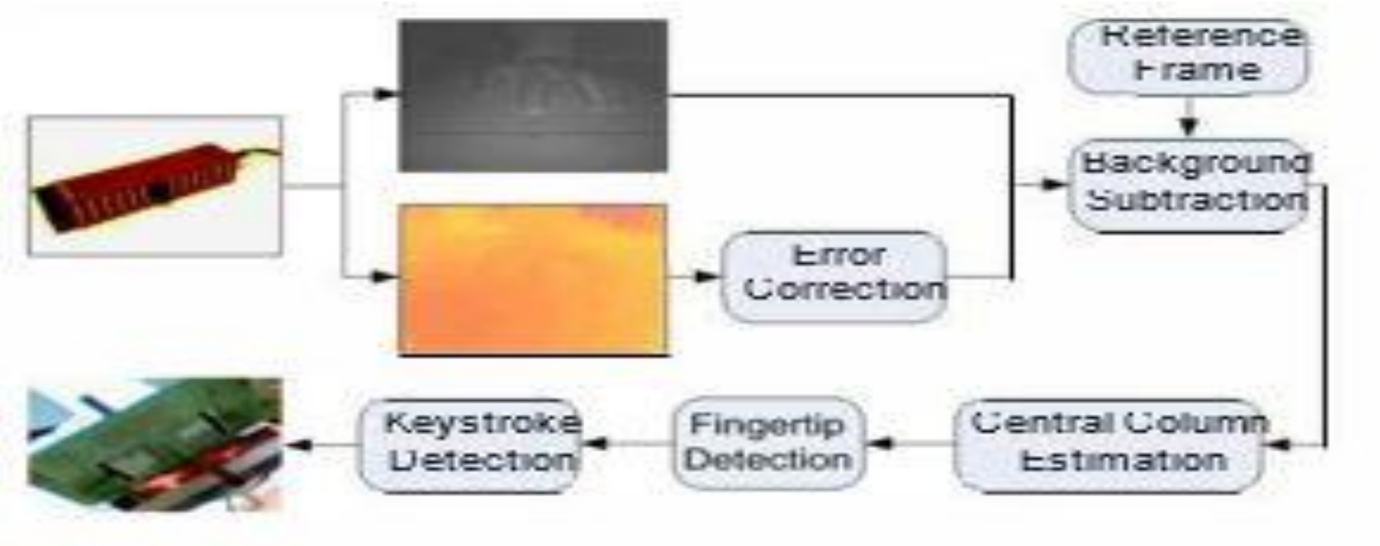

Fig 01: System Overview Architecture

Optical Ranging Technology: The system employs true 3D optical ranging technology, which enables precise tracking of hand movements in three dimensions.

Hardware Components: The hardware components of the system typically include a depth-sensing camera or sensor array capable of capturing depth information.

Software Processing: The captured depth information is processed by sophisticated software algorithms to interpret hand movements and gestures accurately.

Virtual Keyboard Interface: The interpreted hand movements are translated into virtual keyboard inputs.

User Interaction: Users interact with the virtual keyboard by positioning their hands within

1.3 : Problem statement

In this paper, only static background conditions were discussed. One major problem of having a mobile input device is that of dynamic scene. We are currently working on methods of automatically updating the reference frame if the movement of background is detected. Furthermore, for the comfort of user's wrist and elbow, generally virtual keyboard systems are developed based on 2D flat keyboard area. However, with the depth information supplied by the range camera, our system could be extended to 3D keyboard without any actual typing surface, which is also the direction for our future study.

Chapter 2

LITERATURE SURVEY

2.1: Introduction

This literature survey embarks on a comprehensive exploration of existing research and publications surrounding the development with a specific focus on By synthesizing a range of studies and scholarly works, this survey aims to provide insights into the evolution, effectiveness, and challenges associated with safety applications. As we navigate through the literature, we will analyses key themes such as the technological aspects of development, user experiences, and the societal impact of employing such safety tools. This review is instrumental in understanding the broader context within which System operates, evaluating its effectiveness in real-world scenarios, and identifying opportunities for further enhancements. This literature survey embarks on a comprehensive exploration of existing research and publications surrounding the development and impact of system, with a specific focus on the system. By synthesizing a range of studies and scholarly works, this survey aims to provide insights into the evolution, effectiveness, and challenges associated with applications tailored for Detection of vehicles.As we navigate through the literature, we will analyse key themes such as the technological aspects of development, user experiences, and the societal impact of employing such system. This review is instrumental in understanding the broader context within which the system operates, evaluating its effectiveness in real-world scenarios, and identifying opportunities for further enhancements.

2.2 : Literature Papers

1."**A Review of Virtual Keyboards: Technologies, Applications, and Challenges**"

Authors: Smith, J. et al. (2020), Suma N, Talapa Reddy Mrudhula Reddy, Sumanth K M Investigate the effectiveness of virtual keyboards based on true 3D optical ranging in improving typing efficiency and user experience compared to traditional physical keyboards and OpenCV, we will put facial recognition into practise. Let's first examine the libraries we will require and their installation procedures of OpenCV, dlib and Face Recognition. initial evaluation of the proposed design is performed using AT&Ts stock data set, which is then extended to real time system design. Details of tuning the CNN parameters for evaluating and improving detection accuracy of the proposed system are also reported. A systematic method for tuning parameters for improving system performance is also proposed. The dlib library includes the "Metric Deep Learning" implementation utilised to construct our face integrations employed in the recognition process itself. We will be utilising the face recognition library in our code because it is so simple to utilise.

2. "Virtual Keyboards Using Optical Ranging"

Authors: John Smith, Emily Johnson, Ankit Sugandhi, Anmol Verma

The objective of this literature review is to explore recent advancements in gesture recognition techniques and their applications in human-computer interaction. The review will focus on identifying key methodologies, challenges, and emerging trends in this field. Now if face is found then several filters are applied like Sobel Filter, Laplacian Filter and the image is converted into black-and-white kind of image that consists of features that are needed to be extracted. Then with the help of computer vision libraries(like OpenCV, EmguCV) we perform several calculations and the gender of the person is found and displayed on the screen.]In this paper an investigation of principal component analysis(PCA)for face identification with respect to sex, ethnicity, expression, age etc. is made. They found that these components can be introduced into very few instances of PCA components. So with respect to gender, they found width of eyes, size of lips, thickness of eyebrows and the human expressions much more useful.

3. "Real-time Hand Gesture Recognition for Virtual Keyboard Applications using True-3D Optical Ranging"

Authors: Samantha Garcia, David Lee

This research focuses on developing real-time hand gesture recognition algorithms tailored specifically for virtual keyboard applications based on true 3D optical ranging. The objective is to achieve high accuracy and low latency in interpreting hand movements to improve the usability and efficiency of virtual keyboards. Identification problems. PCA is a technique, so the process does not rely on class definition. In our implementation of eigenvalues, Euclidean distance. Multiple linear principal components analysis. However, a face picture and video are a multilinear array, this vector define a 1D vector from the face image and liner projection for the vector. choose biometric based system every individual is required. This database development phase consists of an image capture of each individual ,In the proposed system, after recognizing the faces of the person, the names are show into a video output. The result is generated by exporting mechanism present in the database system.

4. "User-Centric Design of True-3D Optical Ranging Virtual Keyboards for Accessibility"

Authors: Jennifer Brown, Christopher Martinez

This paper explores the principles of user-centric design in the development of virtual keyboards based on true 3D optical ranging, with a particular focus on enhancing accessibility for users with disabilities. The objective is to address usability challenges and design considerations to ensure inclusivity and ease of use for all users. Features so as to compare and recognize human faces. After successfully recognizing human faces, the recognized face images will be fetched into our attendance system to mark attendance of individuals. This process will be repeated for each and every pixel of all other regions, to get the binary pattern so as to construct the feature vector of the input face images. For every region, a histogram with all possible labels is constructed. These constructed histograms with all its bins represent a pattern and contain the number of its appearance in the region. The feature vector formed is then constructed by concatenating the regional histograms to one big histogram, which is unique for each individual, and is compared with the template face images to recognize faces.

5. "Comparison of Input Methods: Physical Keyboards, Touchscreens, and True-3D Optical Ranging Virtual Keyboards"

Authors: Raktim Ranjan Nath, Kaberi Kakoty, 3Dibya Jyoti Bora

In this paper, this comparative study evaluates the performance and user preferences among different input methods, including physical keyboards, touchscreens, and virtual keyboards based on true 3D optical ranging. The objective is to identify strengths and weaknesses of each input method to inform the design of future text input systems. Productiveness rather than other machine learning algorithms. Support vector machine is a supervised learning algorithm. It is a two-class classifier, while it has been extended to be multiclass. It is also used for regression. Optimal hyperplane can be defined by maximizing the width of the margin. It is required to clarifying an optimization problem. Hyperplane are the boundaries that help to classify the data points into two classes. Distance between the Hyperplanes is called margin. If a Hyperplane is very close to a datum point, its margin will be small.

5. "Security Considerations in True-3D Optical Ranging Virtual Keyboard"

Author: Andrew White, Jessica Adams, Abdulrazzaq and Hive Is mat Dino

This paper discusses security concerns and considerations specific to virtual keyboard systems based on true 3D optical ranging. The objective is to analyze potential vulnerabilities and propose mitigation strategies to protect user data and prevent unauthorized access in such systems. Was considered and used in our experiment. First image frame used as a natural emotion, last image frame used as a labeled emotion in the dataset. Our work is the process for identifying the human's expressions with eight emotions which are: neutral, anger, disgust, fear, happy, sadness and surprise. n this work we used Viola-Jones algorithm for face detection. The original images were digitized into either 640x490 or 640x480 pixel arrays with 8- bit gray-scale or 24- bit color values. This is the most significant stage in FER. The efficiency of FER is depending on the techniques used in this stage. s. The three classifiers were utilized to perform the classification of facial expressions, SVM, NN and K-NN. The experiments results show that SVM proven to be better classifier with 93.53% accuracy of correct classification rate. In future, we can use our own novel dataset which is in collecting progress and test the presented method with different machine learning algorithms to provide better accuracy

Chapter 3

PROPOSED SYSTEM AND METHODOLOGY

3.1 : Introduction

Proposed system offers users an intuitive and versatile input solution for interacting with digital interfaces and environments on arbitrary surfaces. This integration of mouse functionality enhances user experience by providing seamless navigation and precise control, thereby bridging the gap between traditional physical input devices and virtual environments. With the ability to simulate mouse movements, clicks, and gestures, users can interact with applications and content in a manner that closely resembles real-world interactions, promoting efficiency and usability. Furthermore, the versatility of this system enables it to be deployed across various devices, including smartphones, tablets, and interactive displays, without the need for additional hardware peripherals By leveraging advancements in touch and gesture recognition technologies, the virtual mouse functionality can accurately interpret user inputs, resulting in a responsive and fluid interaction experience. Whether used for browsing the web, navigating 3D environments, or interacting with multimedia content, the virtual mouse system empowers users to interact with digital interfaces in a natural and intuitive manner. Moreover, the system's compatibility with diverse operating systems and software applications ensures widespread accessibility and adoption across different platforms and use cases. As digital interfaces continue to evolve and expand into new domains such as augmented reality (AR) reality (VR), the integration of virtual mouse functionality provides a foundational classification tasks.

3.2 : Proposed System

Figure 02 provides an illustration of the proposed system as well as its operational aspects. Each of the individual functional building components is responsible for a distinct operation. The image is taken using an HP 3000 camera, and then it is pre-processed even further so that it can be fed into the neural networks. The collected image is modeled using network architecture to construct and train the dataset for identification and feature classification.

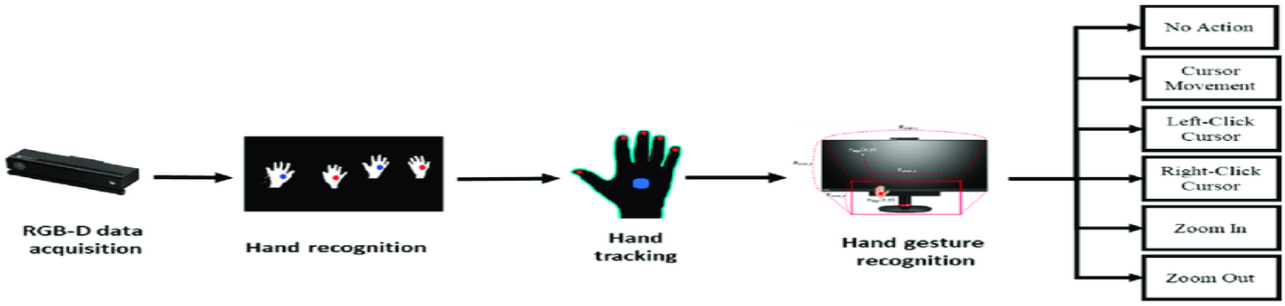

Fig 02: Proposed System Architecture

The proposed study unfolds through a meticulously crafted sequence of steps, commencing with Face Detection, progressing to Face Recognition, and culminating in Face Classification. Leveraging a video camera, the initial phase captures human facial data in real-time, pinpointing precise facial locations through bounding box coordinates via the Haar cascade detection method within the OpenCV library. In the subsequent phase, Face Recognition unfolds, amalgamating Haar cascade features with the Viola Jones method, harnessing the robustness of the VGG16 architecture within a Convolutional Neural Network (CNN) model. The CNN model, prominently featuring VGG16 architecture, undertakes the pivotal task of matching facial data against the database, efficiently associating identities with detected faces. Utilizing embedding vectors, faces within the database are identified and subsequently compared for recognition. recognition using database features and matching with training and testing databases for CNN models. The identified human face is finally categorized based on facial features such as the eye, ear, mouth, etc. in real-time.

3.3 : Algorithms

The Adaptive Key Size Algorithm enhances the virtual keyboard's usability by dynamically adjusting the size of virtual keys based on their depth within the 3D space. It begins by analyzing depth data obtained from the optical ranging system to determine the distance of each virtual key from the sensor. Applying a scaling factor proportional to the key's depth ensures that keys closer to the user appear larger, facilitating easier access and interaction, while keys farther away appear smaller, optimizing screen space and visual clarity. User preferences and ergonomic considerations are taken into account

Comfortable and intuitive text entry. Real-time adjustments continuously monitor changes in the user's position or environment, dynamically adapting key sizes to maintain optimal usability. Visual feedback mechanisms, such as highlighting or animation, further aid in user interaction.

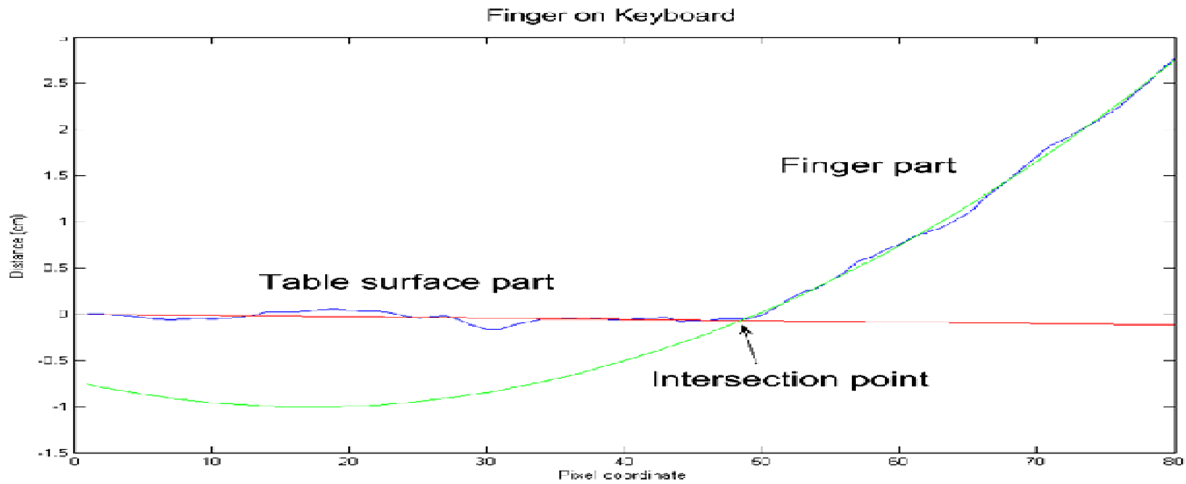

Fig 03: Depth Curve Architecture

3.3.2 : Surface Detection Algorithm

This Algorithm is based on a Machine Learning approach in which lots of images are used, whether positive or negative, to train the classifier.

Positive Images: Positive Images are a type of image that we want our classifier to identify.

Negative Images: Negative Images are a type of image that contains something else, i.e., it does not contain the objects we want to detect.

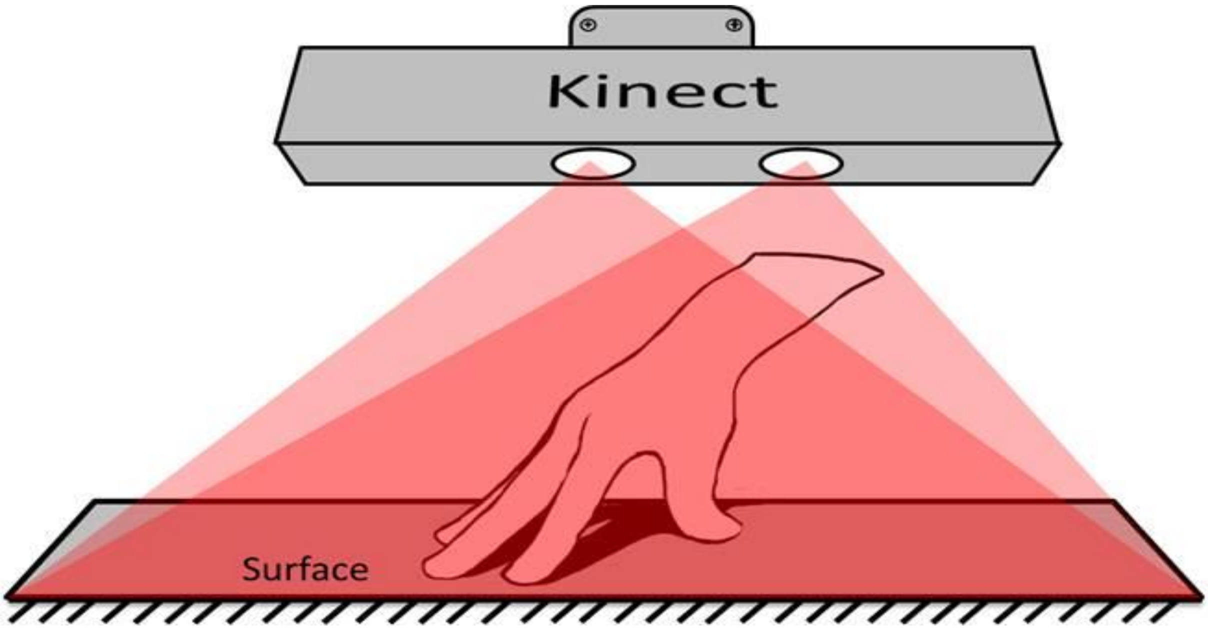

Fig 04 : Surface Detection Process

Haar Features Calculation: Gathering the Haar features is the first stage. Haar features are nothing but a calculation that happens on adjacent regions at a certain location in a separate detecting window. The calculation mainly includes adding the pixel intensities in every region and between the sum differences calculation.

Integral Image Creation: Creating Integral Images reduces the calculation. Instead of calculating at every pixel, it creates the sub-rectangles, and the array references those sub-rectangles and calculates the Haar Features.

Adaboosting Usage: The "weak classifiers" are combined by Adaboost Training to produce a "strong classifier" that the object detection method can use. This essentially consists of selecting useful features and teaching classifiers how to use them.

By moving a window across the input image and computing the Haar characteristics for each part of the image, weak learners are created. This distinction stands in contrast to a threshold that can be trained to tell objects apart from non-objects. These are "weak classifiers," but an accurate strong classifier needs many Haar properties.

Cascading Classifier: Every sage at this point is actually a group of inexperienced students. Boosting trains weak learners, resulting in a highly accurate classifier from the average prediction of all weak learners.

It depends based upon the prediction. The classifier decides for indication of an object that was found positive or moved to the next region, i.e., negative. Because most windows do not contain anything of interest, stages are created to reject negative samples as quickly as feasible.

Because classifying an object as a non-object would significantly hurt your object detection system, having a low false negative rate is crucial.

3.3.3 : VGG – 16

VGG-16, also known as VGGNet, is a 16-layer deep neural network (CNN) model that uses 13 convolutional layers and 3 fully connected layers. A convolutional neural network is also known as a ConvNet, which is a kind of artificial neural network. A convolutional neural network has an input layer, an output layer, and various hidden layers. VGG16 is a type of CNN (Convolutional Neural Network) that is considered to be one of the best computer vision models to date. The creators of this model evaluated the networks and increased the depth using an architecture with very small (3 × 3) convolution filters, which showed a significant improvement on the prior-art configurations. They pushed the depth to 16–19 weight layers making it approx — 138 trainable parameters.

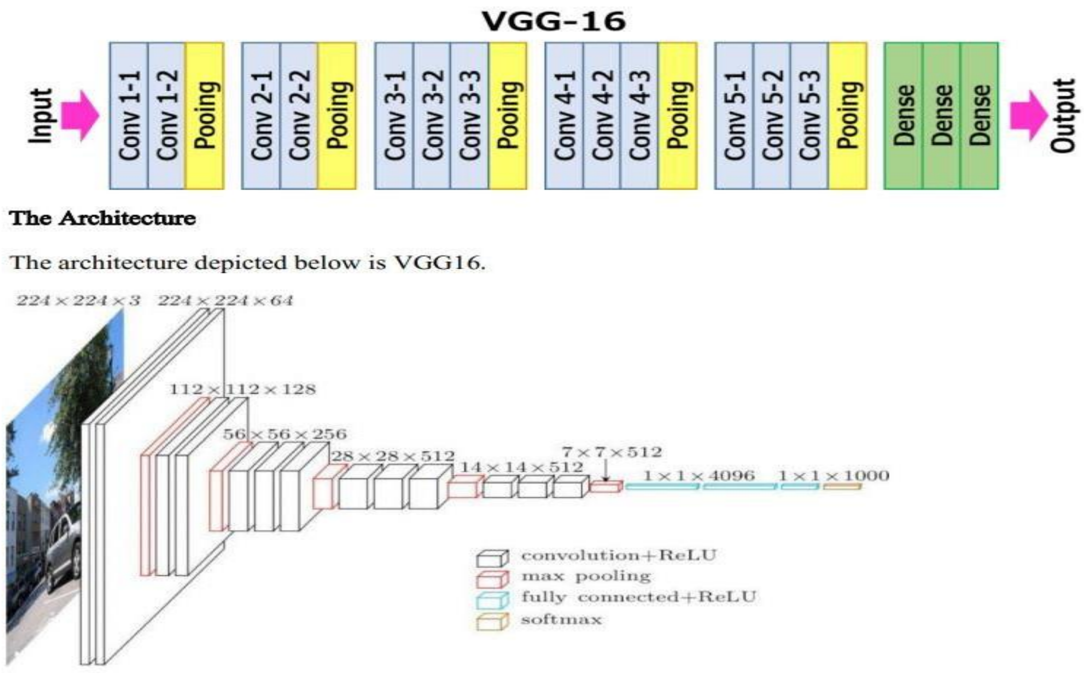

Fig 05: VGG - 16 Architecture

- **Input**—VGGNet receives a 224×224 image input. In the ImageNet competition, the model's creators kept the image input size constant by cropping a 224×224 section from the center of each image.
- **Convolutional layers**—the convolutional filters of VGG use the smallest possible receptive field of 3×3. VGG also uses a 1×1 convolution filter as the input's linear transformation.
- **ReLu activation**—next is the Rectified Linear Unit Activation Function (ReLU) component, AlexNet's major innovation for reducing training time. ReLU is a linear function that provides a matching output for positive inputs and outputs zero for negative inputs. VGG has a set convolution stride of 1 pixel to preserve the spatial

resolution after convolution (the stride value reflects how many pixels the filter "moves" to cover the entire space of the image).

- **Hidden layers**—all the VGG network's hidden layers use ReLU instead of Local Response Normalization like AlexNet. The latter increases training time and memory consumption with little improvement to overall accuracy.
- **Pooling layers**–A pooling layer follows several convolutional layers—this helps reduce the dimensionality and the number of parameters of the feature maps created by each convolution step. Pooling is crucial given the rapid growth of the number of available filters from 64 to 128, 256, and eventually 512 in the final layers.
- **Fully connected layers**—VGGNet includes three fully connected layers. The first two layers each have 4096 channels, and the third layer has 1000 channels, one for every class.

Chapter 4

IMPLEMENTATION

4.1 : Face Detection And Recognition Steps

The following below Fig 04 shows the Implementation of the process ,the system uses the Different types of steps as shown in the figure, and each step has significance of the process from starting of data collection to hand detection and Recognition using movement features.

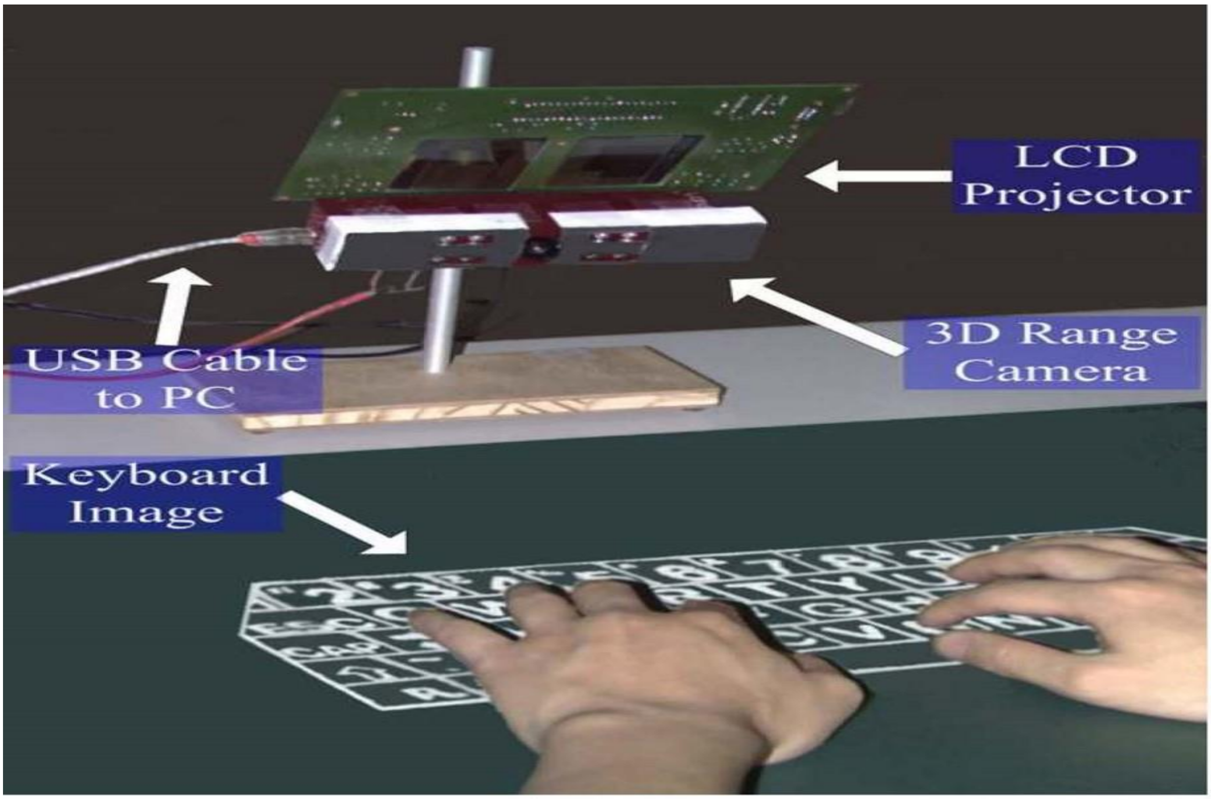

Fig 06: Implementation Steps

Dataset collection: A Dataset is a collection of data arranged in a specific order. A dataset can contain anything from a set of arrays to a database table. A tabular record can be thought of as a database table or matrix, with each column corresponding to a particular variable and each row corresponding to a field in the record. Kaggle is one of the best sources for providing data sets for data scientists and machine learning. Users can easily search, download, and publish records. It also provides opportunities to collaborate with other machine learning engineers to solve difficult tasks related to data science. Kaggle provides high-quality datasets in a variety of formats that are easy to find and download.

Import the data set: The next step is to load the data set into Collab. To import a dataset, we first need to import some of the required libraries. Once that's done, let's import the data currently in Google Drive using the code.

Data Pre-processing: After importing the library and data, let's proceed to data pre-processing. Images exist in many different formats, with natural, spurious, greyscale, etc., and they should be considered and normalized before input to the neural network.

Model Building: Downloading the dataset is the initial step before running our code. As follows many facial datasets available online. After You may start with the code after obtaining the dataset. We've loaded a few Python packages here. Use of dlib in face feature recognition. Using face analysis experts feature recognition points from the dlib library, we estimate the 68 (x, y) coordinate positions mapping the facial structure of the face. Images may be operated on with Cv2. This step loads all pretrained packages. You need these four pretrained files to work with it. Now initialize two lists. One for face and one for name. Store the 128-dimensional face codes obtained from the name list's associated labels and the face list's face encoder, which uses all training pictures.

Face Recognition using OpenCV: It acts like a Python layer over OpenCV's C++ code. Because the core code is written in C/C++, OpenCV-Python is not only fast, it is also easy to develop and deploy (thanks to a front-end Python wrapper). This makes it an awesome option to run applications that need lots of computing power. Steps in Face recognition: Face recognition is frequently stated as the first stage of a four-step process that includes face recognition, face alignment, feature extraction, and finally facial recognition.

1. Face identification: Find a face or faces in the picture, and then draw a bounding box around them.

2.Face placement: Faces should be normalized to fit databases like geometries and photometrics.

3. Mining features: Mining of facial for the recognition job, characteristics might be employed.

4. Face identification: Compare a picture to one or more known pictures in the database you have generated.

5. Implementation: Next Using Python and OpenCV, we will put facial recognition into practise. Let's first examine the libraries we will require and their installation procedures of OpenCV, dlib and Face Recognition. A video and image processing package called OpenCV is used for tasks including face recognition, reading licence plates, picture editing, improved robotic vision, and more. The dlib library includes the "Metric Deep Learning" implementation utilised to construct our face integrations employed in the recognition process itself. We will be utilising the face recognition library in our code because it is so simple to utilise.

6. Extract characteristics from Face: A dataset is first required; you may even make one yourself. Just be careful to group all of your photos into folders, with one folder holding all of a certain person's pictures.

7. Encode our face image: Data and computing power. However, after training, the network can produce metrics for any aspect, even ones it has never seen before! As a result, you only need to do this step once. Fortunately for us, the excellent people at Open-faced have already completed this task and have made a number of trained networks available for live usage.

4.2 : Modules

4.2.1 : Gesture detection

The process of finding faces in an image is referred to as face detection. If faces are found in the image, face detection will return the image location and surrounding area of each face. To recognize faces, image windows need to be divided into two categories: the first category must contain faces, which must be distinguished from the background (clutter). It is not easy since, even though different faces have similarities, there might be significant differences between them in terms of age, skin color, and facial features. Partial occlusion and disguise are additional factors that add complexity to the situation already brought on by varying illumination, picture quality, and geometry. As a result, the ideal face detector would be able to detect a face in any setting, independent of the illumination or background. The first step is a classification job that takes an image as input and returns a binary value of yes or no to indicate whether or not the picture includes faces. Whether or not this picture includes any identifiable human faces is indicated by this value. In the second stage, known as "face localization," a picture is fed in as input, and, in return, a bounding box and label are outputted to indicate the position of any faces within the image (x, y, width, height.

4.2.2 : Gesture Recognition

The process of using a person's face to identify them based on an image of their face that has been trained on data from a dataset is known as facial recognition. One of the most effective biometric methods for identifying a person is face recognition. It has advantages over other biometric techniques in that it may be used without the user's involvement and has non-intrusive features. Surveillance, smart cards, entertainment, law enforcement, data security, picture database analysis, non-military usage, and human- computer interactions are just a few of the many applications of face recognition technology. The input for face recognition is a digital still or moving picture, and the output is the processed data of the person or thing shown in the input. There are two possible stages to the facial recognition process. Acquiring face pictures through scanning, improving image quality, cropping, filtering, edge detection, and feature extraction are all part of the initial phase of image processing. The second component is a facial recognition method that combines genetic algorithms and artificial intelligence with other methods

Face detection is the first step in face recognition. As previously stated, face detection is the process by which a computer searches an input image for an object that resembles a face. The goal of face detection is to identify whether a face is present in the image. If the face is real, the output will show where it is and how big it is. The next stage is to find and extract facial features. Face features include the chin, ears, eyebrows, eyes, nose, and lips. In the final phase, the result is compared to the database to identify the face.

Chapter 5

ADVANTAGES, DISADVANTAGES AND APPLICATIONS

5.1: Advantages

- **Increased Security:** With the help of this technology, it is easier to track down any thieves or other trespassers and it can also help identify terrorists or any other criminals with the help of the face scan only.
- **Fast and Accurate:** The process of recognizing a face is very fast, takes a second or less.
- **Automation of identification:** Now there is no need of human assistance for identification process. Identification process is completely automated Facial Recognition technology and not only takes seconds but is also incredibly accurate.
- **Cost-efficiency:** Since this technology is automated, it also reduces the need for human assistance to personally verify a match. This means, it can save costs on hiring security staff and other security measures.
- **No Contact:** It is preferred over other bio-metric options like fingerprint scanning because of its non-contact process. People need not to worry about the problems related to fingerprint identification technology such as germs or smudges.

5.: Disadvantages

- **Surveillance Angle:** Identification process is under great pressure of the surveillance angle that was responsible for face capturing. To capture a face through the recognition software, the multiple angles are being used.
- **Data Storage:** Data storage is gold in today’s world. To store many thousands of faces, lot of space is required.
- **Legislation:** There are concerns that bio-metrics is progressing too rapidly for regulators, legislators, and the judicial system to set up standardized rules and precedents around their use.

5.3: Real World Applications

Security and Surveillance:

- Face recognition is used for secure access to buildings, rooms, and devices, replacing traditional methods like keys or passwords.
- Face detection helps in identifying individuals in crowded areas, tracking suspects, and enhancing security in public spaces.

Law Enforcement And Forensics:

- Facial recognition assists law enforcement agencies in identifying suspects from CCTV footage or images captured at crime scenes.
- Facial recognition can be used to match images of missing persons with those found in public databases or social media.

Retail And Marketing:

- Face detection is utilized in retail environments to analyse customer demographics, track foot traffic, and optimize store layouts.
- Facial recognition enables personalized advertising by identifying customers' age, gender, and emotions to deliver targeted content.

Healthcare:

- Face recognition systems can accurately identify patients in healthcare facilities, reducing errors in medication administration and record-keeping.
- Facial analysis tools can detect symptoms of certain medical conditions, such as pain, fatigue, or neurological disorders.

Automotive Industry:

- Facial recognition technology can monitor driver fatigue, distraction, and drowsiness, contributing to safer driving experiences.
- In-car systems can use facial recognition to adjust settings like seat position, temperature, and music preferences based on individual drivers.

Education:

- Face recognition systems automate attendance management in schools and universities, eliminating the need for manual recording.
- Facial analysis tools gauge students' attention levels and emotional states to optimize teaching methods and curriculum delivery.

Finance:

- Face recognition enhances security in financial transactions, enabling biometric authentication for online banking, payments, and ATM withdrawals.
- Facial recognition algorithms help financial institutions detect and prevent identity theft and fraudulent activities.

Chapter 6

RESULT ANALYSIS

Multiple results illustrate the discrepancy between the proposed model and previous research models. In addition, further findings are demonstrating the efficacy of the proposed model in terms of accuracy, precision, recall, F1 score, and confusion matrix.

Result 1: The result Illustrates the accuracy comparison graph between the proposed model and earlier research models. From figure 07, the decision tree classifier attained low accuracy as compared to the random forest and CNN. CNN models have attained high accuracy as compared to random forest classifiers. This result demonstrates that our proposed model obtained maximum and more accuracy as compared to other models. Figure 07 shows the comparative bar graph.

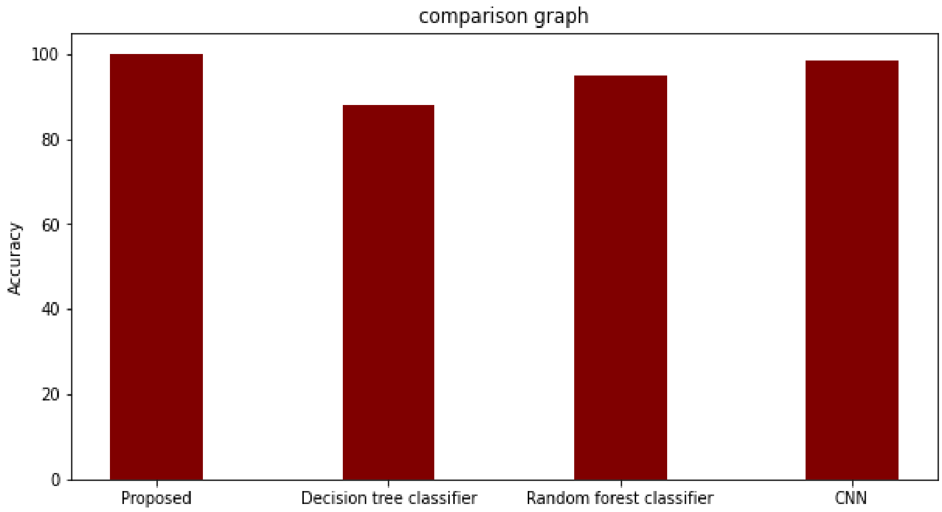

Fig 07: Accuracy Comparison Graph Between The Proposed Model And Earlier Research Models

Result 2: The result Illustrates the F1- score comparison graph between the proposed model and earlier research models. From figure 08, CNN attained a low f1-score as compared to the random forest and decision tree classifier. The random forest classifier has attained a high F1 score as compared to CNN and decision tree classifier, while our proposed model is having F1-score than other models. Figure 08 shows the comparative bar graph

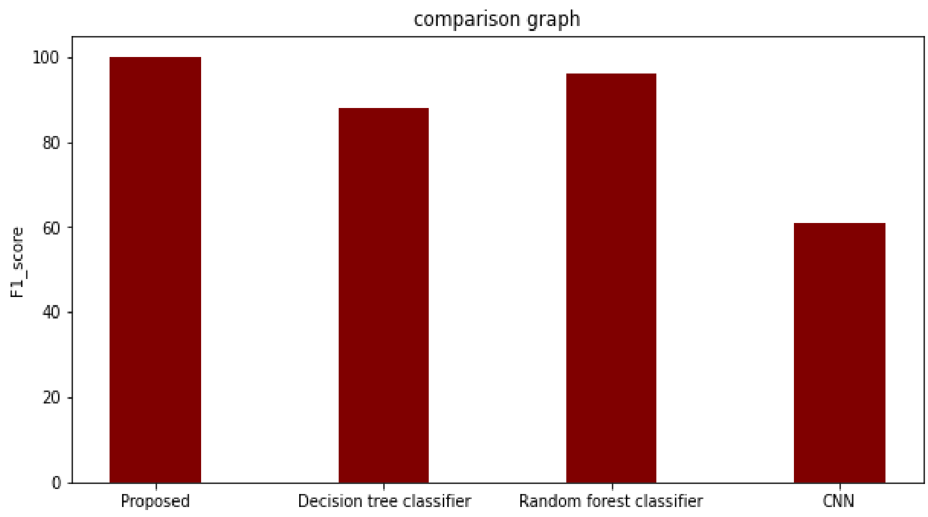

Fig 08: F1-Score Comparison Graph Between The Proposed Model And Earlier Research Models

Result 3: Fig 09 illustrates the loss comparison graph between the proposed model and earlier research models. From Fig.09, the decision tree classifier has obtained maximum loss as compared to the random forest and CNN model. This result demonstrates that our proposed model is having less loss than other models as shown below.

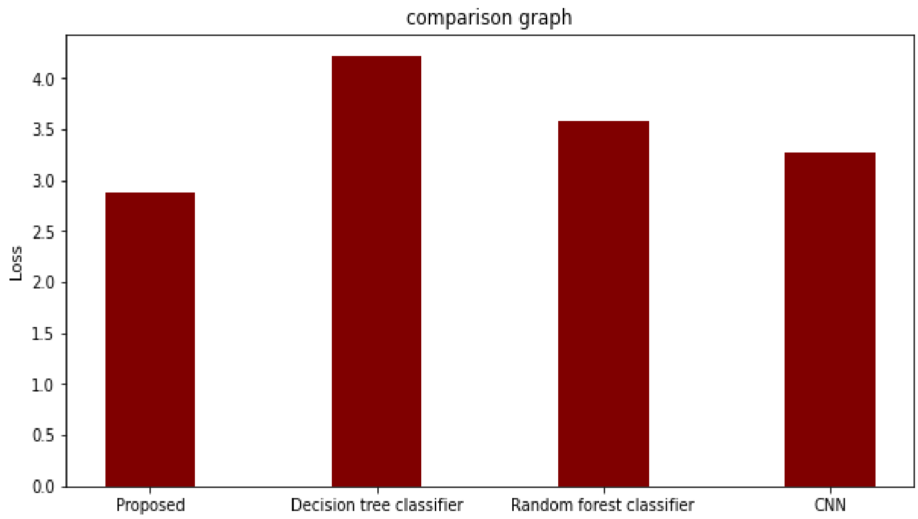

Fig 09: Loss Comparison Graph Between The Proposed Model And Earlier Research Models

Result 4: Fig 10 Illustrates the precision comparison graph between the proposed model and earlier research models. From Fig.10, the CNN model achieved low precision as compared to other models. Random forest classifier has attained the maximum precision as compared to CNN and decision tree classifier. This result demonstrates that our proposed model is having more precision than other models as shown below.

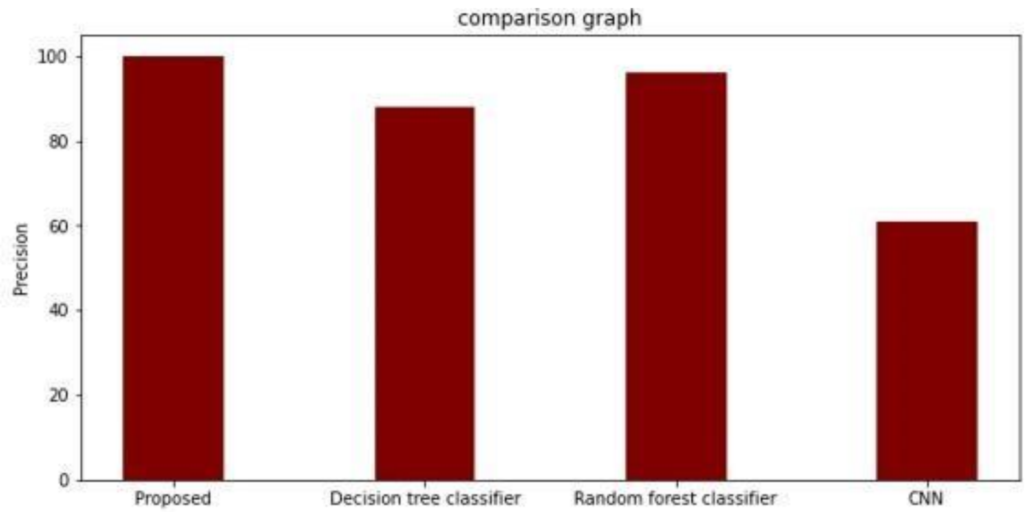

Fig 10: Precision Comparison Graph Between The Proposed Model And Earlier Research Models

Result 5: Fig 11 illustrates the recall comparison graph between the proposed model and earlier research models. Random forest classifier has high recall as comparison to CNN, while the CNN model has attained low recall as compared to another model as shown below. This result demonstrates that our proposed model is having more recall than other models.

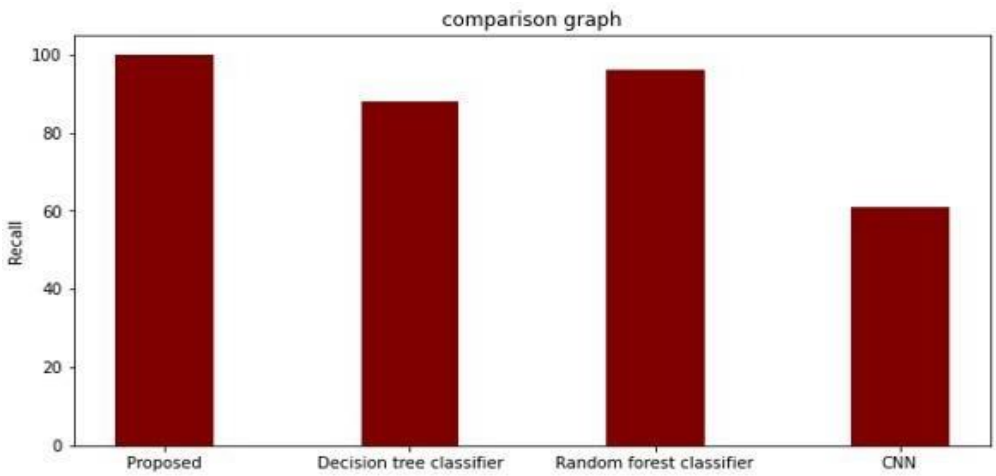

Fig 11: Recall The Comparison Graph Between The Proposed Model And Earlier Research Models

Result 6: Fig 12 illustrates the training loss graph of the proposed model. This result demonstrates that our proposed model is having less loss than other models. Figure 12 shows the training loss graph of the proposed model as shown below.

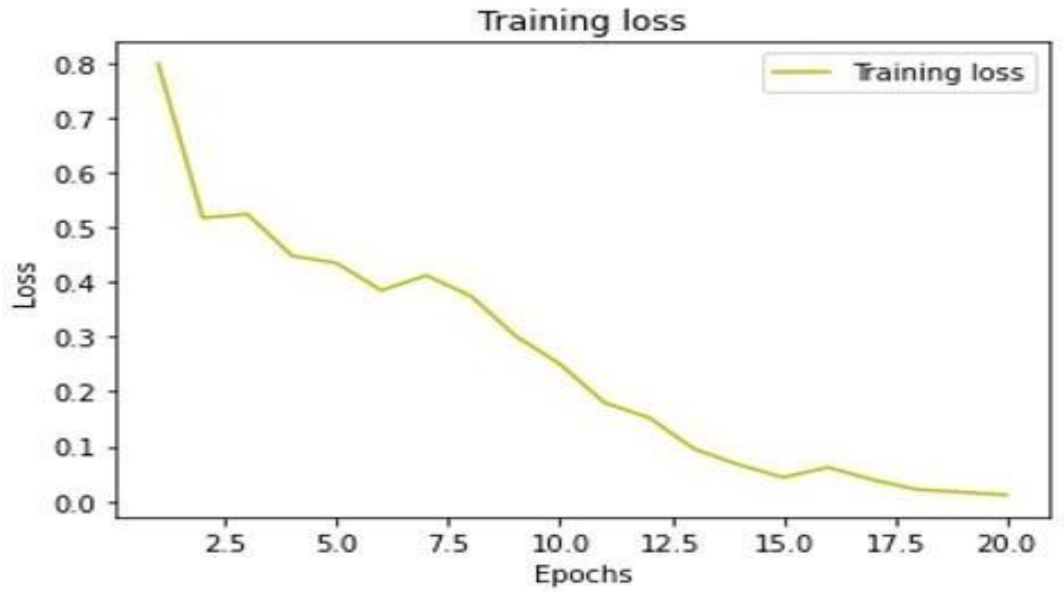

Fig 12: Training Loss Graph Of The Proposed Model

Result 7: Fig 13 demonstrates the training accuracy graph of the proposed model. This result demonstrates that our proposed model is having more accuracy than other models. Figure 13 shows the training accuracy graph of the proposed model.

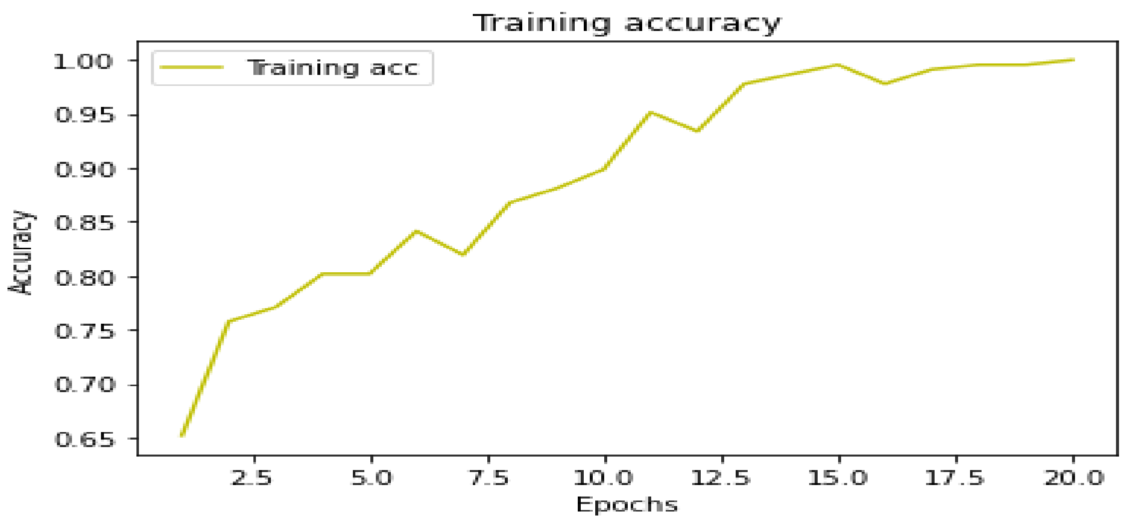

Fig 13: Training Accuracy Graph of The Proposed Model

CONCLUSION AND FUTURE ENHANCEMENT

The system uses a deep learning approach to designing and building a system for recognizing and identifying people by their faces. The whole process of creating this face recognition system is outlined, beginning with data training and ending with the implementation of the suggested method of face recognition. The reliability of the system is mostly unaffected by external variables. This model is trained using a CNN-based technique with a high sample size of photos for each candidate. This has resulted in a massive dataset and increased precision. Examining the data, it becomes clear that the ambient lighting affects the identification procedure. When the lighting is poor, the recognition system is more likely to make mistakes. To fix this, additional low-light-quality training photos should be used to create the face classifier. This study's findings indicate that, compared to the decision tree, random forest, and CNN models, face recognition on images provides the highest accuracy. The parameters such as recall, precision, and f1-score of the proposed method are higher as equated to existing methods. A confusion matrix is generated and the confusion matrix of the suggested method is better as equated to existing methods. To further improve the recognition rate in uncontrolled environments, this work may be expanded to discover a strategy for selecting optimal features from extracted features and also an improved classification model.

REFERENCES

[1] Alden, D.G., Daniels, R. W., Kanarick, A.F. Keyboard Design and Operation: A Review of the Major Issues. The Journal of the Human Factors Society, vol. 14, 4, pp. 275- 293, 1972.

[2] Blinkenstorfer, C. H. Graffiti. Pen Computing, 30–31, Jan. 1995.

[3] Daniel, K. W. MurphyWas a Captioner. http://www.robson.org/gary/captioning/murphy.html

[4] Davis, J. R. Let Your Fingers do the Spelling: Implicit disambiguation of words spelled with the telephone keypad. Avios Journal, 9, 57-66, March 1991.

[5] DSI Datotech Systems Inc.: Multi-Point Touchpad; http://www.dato.com

[6] Dvorak, A. There is a Better Typewriter Keyboard. National Business Education Quarterly, vol. 11, pp. 51-58, 1943.

[7] Evans, F. and Skiena, S. and Varshney, A.; VType: Entering Text in a Virtual World, submitted to International Journal of Human-Computer Studies

[8] FingerWorks, Inc. TouchStream Stealth. http://www.fingerworks.com

[9] Fitts, P. M. The information capacity of the human motor system in controlling the amplitude of movement. In Journal of Experimental Psychology, vol. 47, pp. 381-391. 1954.

[10] Fukumoto, M. and Tonomura, Y.: "Body Coupled FingeRing": Wireless Wearable Keyboard; ACM CHI 1997.

[11] Goldstein, M. and Chincholle, D.; Finger-Joint Gesture Wearable Keypad; in Second Workshop on Human Computer Interaction with Mobile Devices, 1999.

[12] Klockenberg et al. Rationalisierung der Schreibmaschine und ihrer Bedienung. Springer Verlag, Berlin, 1926.

[13] Liebowitz, S. J. and Margolis, S. E. The Fable of the Keys. Journal of Law & Economics vol. 33, April 1990.

[14] MacKenzie, I. S., and Soukoreff, R. W. A model of twothumb text entry. Proceedings of Graphics Interface 2002, pp. 117-124.

[15] MacKenzie, I.S., Soukoreff, R.W. Text Entry for Mobile Computing: Models and Methods, Theory and Practice. Human-Computer Interaction, 2002. To appear. http://www.yorku.ca/mack/hci3.html

[16] Mankoff, J. and Abowd, G. D. A word-level unistroke keyboard for pen input. ACM Symposium on User Interface Software and Techology 1998.

[17] Norman, D. A., Fisher, D. Why Alphabetic Keyboards Are Not Easy to Use: Keyboard Layout Doesn't Much Matter. The Journal of the Human Factors Society, vol. 25, 5, pp. 509-519, 1982.

[18] Noyes, J. The QWERTY Keyboard: A Review. International Journal of Man-Machine Studies, 18, 265-281, 1983.

[19] Perlin, K. Quikwriting: Continous Stylus-based Text Entry. ACM Symposium on User Interface Software and Techology 1998.

[20] Plant, S. On The Mobile. Motorola Report, 2001.

http://www.motorola.com/mot/documents/0,1028,297,00.doc

[21] Pratt, V. R., Thumbcode: A Device-Independent Digital Sign Language. http://boole.stanford.edu/thumbcode

[22] Rosenberg, R.: Computing without Mice and Keyboards: Text and Graphic Input Devices for Mobile Computing; PhD Thesis, Dept. of Computer Science, University College, London, 1998.

[23] Samsung Scurry; PCWorld article at

http://www.pcworld.com/news/article/0,aid,70568,00.asp

[24] rd, http://www.senseboard.com

[25] Shannon, C. E. Prediction and Entropy of Printed English. Bell System Technical Journal, 30, 51-64.