A
Major Project Report

On

"Infobot- A robot that recognizes objects, faces and extracts information about them"

Submitted in partial fulfillment of the

Requirements for the award of the degree of

Bachelor of Technology

In

Computer Science & Engineering –

Artificial Intelligence & Machine Learning

By

Chigiri Varsha	**– 20R21A6610**
Vinjravath Arun Kumar	**– 20R21A6656**
Matam Jahnavi	**– 20R21A6635**
Surukunti SaiKrishna Reddy	**– 20R21A6648**

Under the guidance of
Vijay Gopal Jagadam
Associate Professor

Department of Computer Science & Engineering – Artificial Intelligence & Machine Learning

MARRI LAXMAN REDDY GROUP OF INSTITUTIONS
MLR INSTITUTE OF TECHNOLOGY
(UGC AUTONOMOUS)
Affiliated to JNTUH, Approved by AICTE
Laxman Reddy Avenue, Dundigal, Hyderabad-500 043, Telangana, India
NBA NATIONAL BOARD OF ACCREDITATION
A NAAC

2024

MARRI LAXMAN REDDY GROUP OF INSTITUTIONS

MLR INSTITUTE OF TECHNOLOGY

(UGC AUTONOMOUS)

Affiliated to JNTUH, Approved by AICTE

Laxman Reddy Avenue, Dundigal, Hyderabad-500 043, Telangana, India

NBA NATIONAL BOARD OF ACCREDITATION

A NAAC

Department of Computer Science & Engineering-

Artificial Intelligence & Machine Learning

CERTIFICATE

This is to certify that the project entitled **"Infobot – A robot that recognizes objects, faces and extracts information about them"** has been submitted by **Chigiri Varsha (20R21A6610), Vinjravath Arun Kumar (20R21A6656), Matam Jahnavi (20R21A6635), Surukunti Sai Krishna Reddy (20R21A6648)** in partial fulfilment of the requirements for the award of degree of Bachelor of Technology in ComputerScience & Engineering – Artificial Intelligence & Machine Learning from Jawaharlal NehruTechnological University, Hyderabad. The results embodied in this project have not beensubmitted to any other University or Institution for the award of any degree or diploma.

Internal Guide **Head of the Department**

Project coordinator **External Examiner**

Department of Computer Science & Engineering-

Artificial Intelligence & Machine Learning

DECLARATION

We hereby declare that the project entitled **"Infobot- A robot that recognizes objects, faces and extracts information about them"** is the work done during the period from **January 2024 to May 2024** andis submitted in partial fulfilment of the requirements for the award of degree of Bachelor of Technology in Computer Science & Engineering – Artificial Intelligence & Machine Learning from Jawaharlal Nehru Technology University, Hyderabad. The results embodied in this project have not been submitted to any other university or Institution for the award of any degree or diploma.

Chigiri Varsha – 20R21A6610

Vinjravath Arun Kumar – 20R21A6656

Matam Jahnavi – 20R21A6635

Surukunti SaiKrishna Reddy – 20R21A6648

Department of Computer Science & Engineering-

Artificial Intelligence & Machine Learning

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of people who made it possible, whose constant guidance and encouragement crowned our efforts with success. It is a pleasant aspect that we now have the opportunity to express our guidance for all of them.

First of all, we would like to express our deep gratitude towards our internal guide **VIJAY GOPAL JAGADAM, Associate Professor, Dept. of CSE-AIML** for his support in the completion of our dissertation. We wish to express our sincere thanks to **Dr.K.SAI PRASAD**, HOD, Dept. of CSE- AIML and principal **Dr. K. SRINIVAS RAO** for providing the facilities to complete the dissertation.

We would like to thank all our faculty and friends for their help and constructive criticism during the project period. Finally, we are very much indebted to our parents for their moral support and encouragement to achieve goals.

Chigiri Varsha – 20R21A6610

Vinjravath Arun Kumar – 20R21A6656

Matam Jahnavi – 20R21A6635

Surukunti SaiKrishna Reddy – 20R21A6648

Department of Computer Science & Engineering-

Artificial Intelligence & Machine Learning

ABSTRACT

The aim of this project is to develop "Infobot," a friendly chatbot robot with advanced capabilities in object and face recognition. The robot's primary functions include recognizing specific objects and faces of certain users, extracting relevant information to enable personalized interactions. Additionally, Infobot serves as an "infobot" for known faces, providing tailored information and assistance. Moreover, it can identify unknown faces of visitors, gathering information about their visit and the person they are intending to meet. Leveraging cutting-edge AI and computer vision technologies, our project seeks to create an intelligent, user-friendly robot for enhanced human-robot interactions. Enhancing human-robot interactions in diverse scenarios by developing an intelligent and adaptive chatbot with advanced object and face recognition capabilities. Implementing cutting-edge AI and computer vision technologies to enable friendly chat interactions, personalized assistance, and efficient face and object recognition.

APPENDIX-1

LIST OF FIGURES

LIST OF FIGURES

APPENDIX-2

LIST OF TABLES

LIST OF TABLES

APPENDIX-3

LIST OF ABBREVIATIONS

ABBREVIATIONS

YOLO	**You Only Look Once**
CNN	**Convolutional Neural Network**

APPENDIX-4

REFERENCES

References

[1] A. Sarda, S. Dixit and A. Bhan, "Object Detection for Autonomous Driving using YOLO [You Only Look Once] algorithm," *2021* Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 2021, pp. 1370-1374, doi: 10.1109/ICICV50876.2021.9388577.

[2] V. Senthilkumar, P. Saranya, B. K. Rani, S. P, R. Kuchipudi and M. A. A. Walid, "Deep Unified Model for Face Recognition based on Convolution Neural Network and Edge Computing," 2023 8th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2023, pp. 10.1109/ICCES57224.2023.10192630.

[3] M. M. Hasan, Lim Hooi Jiun, Ng Wei Cheun and M. S. Shahid, "Smart telephone design-caller identification and answering machine," ICSE'98. 1998 IEEE International Conference on Semiconductor Electronics. Proceedings (Cat. No.98EX187), Bangi, Malaysia, 1998, pp. 217-222, doi: 10.1109/SMELEC.1998.781183.

[4] T. Feng, Y. Yu, L. Wu, Y. Bai, Z. Xiao and Z. Lu, "A Human-Tracking Robot Using Ultra Wideband Technology," in IEEE Access, vol. 6, pp. 42541-42550, 2018, doi: 10.1109/ACCESS.2018.2859754

[5] R. Menaka, N. Archana, R. Dhanagopal and R. Ramesh, "Enhanced Missing Object Detection System using YOLO," *2020* 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020, pp. 1407-1411, doi: 10.1109/ICACCS48705.2020.9074278.

[6] S. Dwijayanti, M. Iqbal and B. Y. Suprapto, "Real-Time Implementation of Face Recognition and Emotion Recognition in a Humanoid Robot Using a Convolutional Neural Network," in IEEE Access, vol. 10, pp. 89876-89886, 2022, doi: 10.1109/ACCESS.2022.3200762.

[7] R. L. Galvez, E. P. Dadios, A. A. Bandala and R. R. P. Vicerra, "YOLO-based Threat Object Detection in X-ray Images," 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Laoag, Philippines, 2019, pp. 1-5, doi: 10.1109/HNICEM48295.2019.9073599.

[8] Y. Ye, H. You and J. Du, "Improved Trust in Human Robot Collaboration with ChatGPT," in IEEE Access, vol. 11, pp.5574855754, 10.1109/ACCESS.2023.3282111

[9] M. Hnewa and H. Radha, "Multiscale Domain Adaptive Yolo for Cross-Domain Object Detection," 2021 IEEE International Conference on Image Processing (ICIP), Anchorage, AK, USA, 2021, pp. 3323-3327, doi: 10.1109/ICIP42928.2021.9506039

[10] W. S. M. Sanjaya, D. Anggraeni, K. Zakaria, A. Juwardi and M. Munawwaroh, "The design of face recognition and tracking for human-robot interaction," 2017 2nd International conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), Yogyakarta, Indonesia, 2017, pp. 315-320, doi: 10.1109/ICITISEE.2017.8285519.

[11] P. Anki, A. Bustamam, H. S. Al-Ash and D. Sarwinda, "High Accuracy Conversational AI Chatbot Using Deep Recurrent Neural Networks Based on BiLSTM Model," 2020 3rd International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2020, pp. 382-387, 10.1109/ICOIACT50329.2020.9332074.

[12] K. Deepika, V. Tilekya, J. Mamatha and T. Subetha, "Jollity Chatbot- A contextual AI Assistant," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 1196 1200, doi: 10.1109/ICSSIT48917.2020.9214076.

[13] N. P. Patel, D. R. Parikh, D. A. Patel and R. R. Patel, "AI and Web-Based Human-Like Interactive University Chatbot (UNIBOT)," 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2019, pp. 148-150, doi: 10.1109/ICECA.2019.8822176.

[14] Shaik, Amjan, S. Lohitha, and J. Ramesh Babu. "HEALTHCARE ASSISTANCE USING AI." *Journal of Engineering Sciences* 14.06 (2023).

[15] S. K. Singh, S. Kumar and P. S. Mehra, "Chat GPT & Google Bard AI: A Review," 2023 International Conference on IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2023, pp. 1-6, doi:10.1109/ICICAT57735.2023.10263706

[16] Shafeeg, A., Shazhaev, I., Mihaylov, D., Tularov, A., & Shazhaev, I. (2023). Voice assistant integrated with chat gpt. Indonesian Journal of Computer Science, 12(1).

[17] S. T. Blue and M. Brindha, "Edge detection-based boundary box construction algorithm for improving the precision of object detection in YOLOv3," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-5, doi: 10.1109/ICCCNT45670.2019.8944852.

[18] S. Mane and S. Mangale, "Moving Object Detection and Tracking Using Convolutional Neural Networks," 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2018, pp. 1809-1813, doi:

10.1109/ICCONS.2018.8662921

[19] Koubaa, A. (2023). ROSGPT: Next-Generation Human Robot Interaction with ChatGPT and ROS Preprints.https://doi.org/10.20944/preprints202304.0827.

[20] Zhao, X. F., Chen, Z. H., Yin, H. F., & Wu, X. J. (2023). Design of intelligent visitor system based on cloud and edge collaborative computing. Journal of Algorithms & Computational Technology, 17, 17483026231169154.

INDEX

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

The robot is designed to engage in friendly conversations with users, while also possessing the ability to recognize specific objects and faces. The system can extract relevant information about certain users, enabling personalized interactions for an enhanced user experience. Additionally, the robot functions as an "infobot" for known faces, providing tailored information and assistance. The innovative aspect lies in its capacity to identify unknown faces, gathering essential information about their visit and the person they are intending to meet. Leveraging modern AI and computer vision technologies, this project addresses the challenges of efficient object and face recognition and strives to transform human-robot interactions in various real-world applications. The system's adaptability, privacy safeguards, and seamless integration of conversational abilities make it an innovative solution for personalized, intelligent, and user-friendly robot interactions.

1.2 PURPOSE OF THE PROJECT

The purpose of this project is to develop a robot named as “Infobot”. It is a conversational robot with advanced speech-based interaction capabilities. This robot will be equipped to engage users in friendly conversations through speech, recognize objects and faces, and provide tailored responses. Additionally, it will serve as an infobot for known faces and gather information about unknown visitors, enhancing human-robot interaction in diverse scenarios.

1.3 MOTIVATION

The inspiration for the project arose from the desire to create an advanced and interactive robotic system capable of engaging in natural and friendly conversations with users through speech. This innovative idea covers a range of aspects, including speech-based interaction and the ability to recognize objects and faces. The idea of creating a robot, that can not only communicate with users

but also understand and reply through speech gives a futuristic picture of human-robot interaction. The idea was to develop a complex system that not only offers enlightening responses but also forges a close bond with the user, much like speaking with a human counterpart. The first idea came from seeing the possibility of combining speech synthesis and recognition technologies with advanced object and face recognition capabilities, which would improve the entire user experience and make the robot a flexible assistance in a variety of circumstances.

CHAPTER 2

LITERATURE SURVEY

An extensive literature survey has been conducted by studying existing systems of Certificate verification and generation. A good number of research papers, journals, and publications have also been referred before formulating this survey.

2.1 EXISTING SYSTEM

The existing system of Chatbots, however, encounters hurdles when it comes to attaining human-like communications owing to the zero contextual comprehension and no emotional Intelligence. The fuzzy and difficult nature of the questions and data, thus, negatively affects the response rates. Prejudice and privacy invitations must be carefully considered before anonymization. Bridging multi-turn communication issues and filling in the language gaps are very significant. The misjudgment, concerns of privacy, and looming of ethics with facial recognition technology like biases are the issues that may be compromising its implementation. Real-time object identification in computer vision remains challenging, with problems of accuracy improvement, speed, and scalability. Visitor monitoring systems may deal with privacy and accuracy problems even with the complexity in the case of their (visitor monitoring systems) Integration. The prospect of universal adoption of biometric systems is, beyond doubt, fraught with the problem of safeguarding security as well as privacy.

1	Deep Unified Model for Face Recognition Based on Convolution Neural Network and Edge Computing	
Reference in APA format	V. Senthilkumar, P. Saranya, B. K. Rani, S. P, R. Kuchipudi and M. A. A. Walid, "Deep Unified Model for Face Recognition based on Convolution Neural Network and Edge Computing," 2023 8th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2023, pp. 1816-1821, doi: 10.1109/ICCES57224.2023.10192630.	
URL of the Reference	**Author's Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8721062	Muhammad Zeeshan Khan(zeeshan.khan@kics.edu.pk)	CNN, face, attendance, RCNN, anchors, RPN, edge computing

	Shahid Mumtaz, Muhammad Imran.	
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution and what is the problem that needs to be solved**	**What are the components of it?**
Face Recognition Based on Convolution Neural Network and Edge Computing	The objective is to improve the efficiency and accuracy of face recognition in devices and accurately recognize individuals in real-time	The author used a deep unified model based on CNN and edge computing for processing the data.

The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantages & Disadvantages of Each Step in This Process

The process helps to identify known and unknown faces with high accuracy, allowing for easy classification of people. This enables robots to interact with recognized individuals.

	Process Steps	Advantage	Disadvantage (Limitation)
1	The input to the proposed system is in the form of an image, which can be uploaded either from the directory or by capturing through the device's camera.	The proposed solution is highly accurate and efficient in recognizing and authenticating individuals in real time.	The system requires a good training dataset to achieve efficient and acceptable results.
2	The captured image is processed to detect the faces in the frame using a deep unified model based on a convolution neural network.	The use of edge computing reduces data latency and increases the real-time response, making the system more efficient.	The system may face challenges in recognizing faces in low-light or noisy environments.
3	The detected faces are then compared with the known faces in the database to recognize and authenticate individuals.	The system is designed to work with a large number of simultaneous images from different smart classrooms, making it suitable for use in educational institutions.	The system may face privacy concerns related to the collection and storage of facial data.
4	The processing of the data is done at the edges of the nodes to reduce the data latency and increase the real-time response.		
5	The recognized individuals are marked as present in the attendance system, and the necessary data is synchronized into the gateway device after the particular time stamp.		

6	The attendance data is passed to the cloud for generating reports of attendance on a specific day.		

Major Impact Factors in this Work

Find all main factors and variables that are related to each solution. Then find the relationship between factors. (Independent variable) causes a change in (Dependent Variable) and it isn't possible that (Dependent Variable) could cause a change in (Independent Variable).

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
The accuracy of face recognition is used as a dependent variable as it measures the performance to evaluate the proposed system.	The algorithm components like CNN and data processing techniques like edge computing are factors acting independently.	The size of the data set provided for training acts as a moderating variable.	The use of convolutional neural networks (CNN) for face detection and recognition acts as a mediating variable.

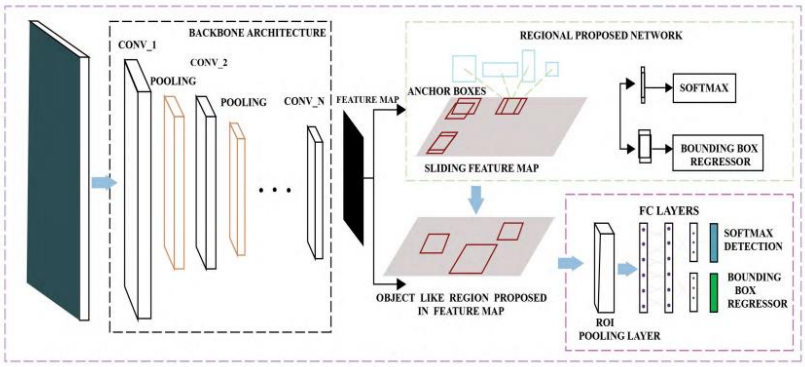
Relationship Among The Above 4 Variables in This article

CNN-based face detection and recognition algorithm, outperforming traditional methods. Utilizing a smart classroom for attendance, achieves 97.9% accuracy, while an IoT-based edge computing architecture enhances data latency and real-time response, acting as mediating and moderating variables.

Input and Output	Feature of This Solution	Contribution & The Value of This Work
Input: image of a face, which can be uploaded from a directory or captured through a device's camera. **Output**: output is the recognition and authentication of the individual in the image, which is marked as present in the attendance system.	The proposed solution excels in face detection and recognition, boasting 94.6% and 85.5% accuracy, respectively. It ensures real-time performance through edge computing, making it efficient for educational settings. Scalable and efficient, it employs a deep unified model for data processing. Attendance data syncs with the cloud for easy management and privacy concerns are addressed in compliance with regulations.	Overall, the proposed solution has the potential to improve attendance management in educational institutions by providing an efficient and accurate system for face recognition-based attendance.

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
The solution simplifies attendance, saves time, scales efficiently, cuts costs, and enhances security with facial recognition, offering major benefits to educational institutions.	Since this is a performance evaluation of various algorithms, not much to project on the negative side as all the things used are defined in advance.

Analyse This Work By Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper

The solution explains an innovative face recognition approach using convolutional neural networks and edge computing to handle IoT data challenges. It details the methodology, practical implementation, and evaluation. Suggested are practical testing and exploration of broader applications.	TensorFlow, pytorch, Flask frame work	I. Abstract II. Introduction III. Literature Review IV. Methodology V. Practical Implementations of the Proposed System VI. Experiments and Results VII. Conclusion
Diagram/Flowchart		
		

2	Real-Time Implementation of Face Recognition and Emotion Recognition in a Humanoid Robot Using a Convolutional Neural Network	
Reference in APA format	S. Dwijayanti, M. Iqbal and B. Y. Suprapto, "Real-Time Implementation of Face Recognition and Emotion Recognition in a Humanoid Robot Using a Convolutional Neural Network," in IEEE Access, vol. 10, pp. 89876-89886, 2022, doi: 10.1109/ACCESS.2022.3200762.	
URL of the Reference	**Authors' Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/abstract/document/9864185	Muhammad Iqbal, Bhakti Yudho Suprapto, Yul Yunazwin Nazaruddin	Accuracy, convolutional neural network, emotion recognition, face recognition, humanoid robot.
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution and what is the problem that needs to be solved**	**What are the components of it?**

A model for real-time implementation of face recognition and emotion recognition in a humanoid robot using a convolutional neural network.	The goal is real-time face and emotion recognition in humanoid robots using a custom CNN. Filling a research gap, it introduces a tailored CNN and distance-based human localization for high accuracy, with applications in robotics, healthcare, and security.	The components in this include a modified CNN architecture, a method for measuring the distance between the object's face and the position of the robot.

The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process

The process helps to identify known and unknown faces with high accuracy, allowing for easy classification of people. This enables robots to interact with recognized individuals.

	Process Steps	Advantage	Disadvantage (Limitat
1	The work involves hardware design, dataset preparation, CNN architecture selection, training, and performance evaluation.	The proposed system can recognize faces and emotions in real time, which can have various applications in fields such as robotics, healthcare, and security	The hardware design step requires careful selection positioning of component which can be time-consu and require technical exp
2	In hardware design, the components are selected and arranged strategically. To ensure the robot can capture face images, move its head to follow the object's face and display a visual appearance of the eyes.	The modified CNN architecture has shown better performance in face recognition and emotion recognition than well-known CNN architectures, i.e., VGG16 and AlexNet.	Dataset preparation dema ample data, often challen to acquire. System accura may be influenced by ligh expressions, and occlusio Training and testing nece substantial computational resources and time, posin system challenges.
3	In the dataset preparation step, face images are captured using a webcam embedded in the robot's eyes. These images are used as a dataset for training the CNN.	One unit combines face recognition and emotion recognition. The recognition system is embedded in the robot, enabling real-time interaction with humans based on their faces and emotions	Although the study show that the distance between recognized object and the position of the robot can measured well, the averag error rate of 2.52% may s considered high in some applications.
4	In the CNN architecture selection step, two well-known CNN architectures, i.e., VGG16 and AlexNet, are compared with the proposed modified architecture. The modified architecture is chosen because it has shown better performance in face recognition and emotion recognition.	Unlike other studies that utilized a previously collected dataset, this study used primary data obtained from male and female students, where some students wore glasses and some female students wore a hijab.	
5	In the training and testing step, the CNN is trained using the face images dataset and tested using a separate dataset. The accuracy of the test is calculated using a formula that considers true positives, true negatives, false	The study confirmed real-time recognition and precise distance measurement by humanoid robots. It underscored distance and illumination as critical recognition factors.	

	positives, and false negatives.		
6	In the performance evaluation step, the proposed system's accuracy is measured as the recognition rate of faces and emotions in real-time. The accuracy value shows the level of effectiveness per class of a classification.		

Major Impact Factors in this Work

Find all main factors and variables that are related to each solution. Then find the relationship between factors. (Independent variable) causes a change in (Dependent Variable) and it isn't possible that (Dependent Variable) could cause a change in (Independent Variable).

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) varia
Recognition accuracy of faces and emotions in real-time by the humanoid robot.	The proposed modified CNN architecture is used to train the robot to recognize faces and emotions.	The real-time implementation of the systems moderates the relationships between the independent variables and the dependent variables. It implements the overall system performance.	The distance betwee recognized object an position of the affects the accurac recognition.

Relationship Among The Above 4 Variables in This article

Neural network architectures and data quality influence face and emotion recognition accuracy and dist measurement error. Real-time implementation moderates these relationships, while distance measure mediates overall system performance.

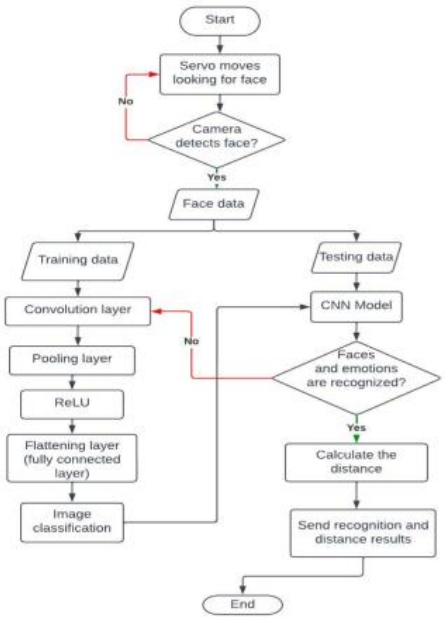
Input and Output		Feature of This Solution	Contribution & The Value of This Work
Input	**Output**	The proposed system utilizes a customized CNN architecture for real-time face and emotion recognition, employing a camera module linked to a laptop to capture visual data. It processes point coordinates of recognized faces, achieving impressive accuracy (87% for faces, 67% for emotions). Furthermore, the system adeptly calculates distances and performs well in different lighting conditions, while also facilitating precise servo and dot matrix control.	The proposed system is successfully implementing face and emotion recognition in a CNN-based humanoid robot using a modified CNN architecture. The proposed model attains high accuracy in face and emotion recognition, with 87% and 67% accuracy, respectively. Moreover, the system effectively calculates distances and exhibits robust performance in various lighting conditions. Its value extends to applications in healthcare, security, and entertainment, enhancing human-robot interactions.
A camera module connected to a laptop captures the visual appearance of a person's face and emotions. The captured data is then processed by the modified	Output is the recognition accuracy of faces and emotions in real-time by the humanoid robot, which is displayed on a dot matrix that shows the visual appearance		

CNN architecture, which is trained to recognize faces and emotions.	of the eyes of the humanoid robot.		

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
The solution's positive impact is substantial, and applicable in healthcare, security, and entertainment, enhancing human-robot interactions through face and emotion recognition. It benefits patient monitoring, security, and user engagement across domains.	Any technology can have potential negative impacts, such as privacy concerns or unintended consequences. It is important to consider these factors and address them appropriately when implementing such systems.

Analyse This Work By Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
This study outlines a well-structured implementation of face and emotion recognition in a humanoid robot, utilizing a modified CNN architecture for high accuracy. It offers strength in diverse data sources and architectural comparisons but overlooks ethical considerations, such as privacy and socio-economic impact. Overall, it holds promise for enhancing human-robot interactions across domains, provided ethical concerns are duly addressed.	Hardware components: 1. Webcam 2. JX Servo 60KG 3. Arduino 4. Raspberry Pi 5. Dot matrix	I. Abstract II. Introduction III. Methods IV. Results and Discussions V. Conclusion

Diagram/Flowchart



3	MULTISCALE DOMAIN ADAPTIVE YOLO FOR CROSS-DOMAIN OBJECT DETECTION	
Reference in APA format	M. Hnewa and H. Radha, "Multiscale Domain Adaptive Yolo For Cross-Domain Object Detection," 2021 IEEE International Conference on Image Processing (ICIP), Anchorage, AK, USA, 2021, pp. 3323-3327, doi: 10.1109/ICIP42928.2021.9506039.	
URL of the Reference	**Authors' Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/document/9506039	Mazin Hnewa, Hayder Radha	Object detection, Domain adaptation, Adversarial training, Domain shift.
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution and what is the problem that needs to be solved**	**What are the components of it?**
MultiScale Domain Adaptive YOLO (MS-DAYOLO) framework for domain adaptation in object detection using multiple adaptation paths in YOLOv4.	The goal of the proposed solution MS-DAYOLO proposes a solution to the domain shift problem faced by object detection applications. It generates domain-invariant features using multiple domain adaptation paths and classifiers at different scales of the YOLOv4 detector. This improves the detector's performance when trained with the proposed solution and tested on target data representing challenging weather conditions for autonomous driving.	The backbone, the neck, and the head. The backbone is responsible for extracting multiple layers of features at different scales. The neck collects these features and feeds them to the head. The head predicts bounding boxes
The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process		
The proposed MS-DAYOLO framework improves the robustness and accuracy of object detection in		

cross-domain scenarios, making it a promising solution for real-world applications.

	Process Steps	Advantage	Disadvantage (Limitation)
1	Extract features at multiple scales using YOLOv4's backbone network applied domain adaptation separately to each of the three scales of features, as well as to different combinations of two scales at a time, and finally applied domain adaptation to all three feature scales simultaneously using our proposed MS-DAYOLO.	The proposed framework achieves significant improvements in object detection performance on the target domain compared to the original YOLOv4 approach.	The proposed MS-DAYOLO framework is specifically designed for the YOLOv4 object detector and may not be directly applicable to other object detection architectures.
2	Train the object detector using both labeled source and unlabeled target data and train domain classifiers for each scale.	The framework generates domain-invariant features that reduce the impact of domain shift, making it more robust to changes in the target domain.	The proposed framework addresses the domain shift problem encountered in cross-domain object detection scenarios. It may not be suitable for other types of domain adaptation problems.
3	Evaluate the performance of the object detector in the target domain and compare results of domain adaptation to MS-DAYOLO at different scales.	The proposed method does not require any annotation of the target domain, making it more practical and cost-effective for real-world applications.	It is evaluated on a specific set of objects commonly found in autonomous driving scenarios. It may not be effective for other types of objects or scenes
4	Analyze the results and conclude that applying domain adaptation to all three feature scales improves the detection performance on the target domain and achieves the best result.	The YOLOv4 architecture is already known for its fast and efficient object detection capabilities, and the proposed framework builds on this by adapting it to cross-domain scenarios.	
5	Show that the proposed MS-DAYOLO outperforms the original YOLOv4 approach by a significant margin and almost reaches the performance of the ideal (oracle) scenario, especially for some object classes in terms of average precision and overall mAP.	The proposed framework can be applied to a wide range of applications that involve cross-domain object detection, such as autonomous driving, surveillance, and robotics.	
6	Provide examples of detection results of the proposed method as compared to the original YOLOv4.		

Major Impact Factors in this Work

<Find all main factors and variables that are related to each solution. Then find the relationship between factors. (Independent variable) causes a change in (Dependent Variable) and it isn't possible that

(Dependent Variable) could cause a change in (Independent Variable).

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
The evaluation metrics such as Average Precision (AP), mean average precision (mAP), and class-specific metrics (P: Person, R: Rider, C: Car, M: Motorcycle, B: Bicycle, V: Vehicle, TS: Traffic Sign, TL: Traffic Light, SLS: Speed Limit Sign) represent the dependent variables	The system is tested under different weather conditions so conditions like foggy, rainy sunny may act as the independent variables.	The lambda parameter controls the impact of the Domain Adaptive Network (DAN) on the backbone. It balances the trade-off between the detection loss and domain classification loss.	Domain-invariant features extracted from different scales of the YOLOv4 mediate the impact of domain shifts on object detection.

Relationship Among The Above 4 Variables in This article

The relationship among these variables involves how weather conditions and domain adaptation methods influence the object detection performance, with Lambda moderating the adaptation process and domain-invariant features acting as a mediator to mitigate the effects of domain shifts. The goal is to enhance the system's robustness and accuracy across diverse environmental conditions.

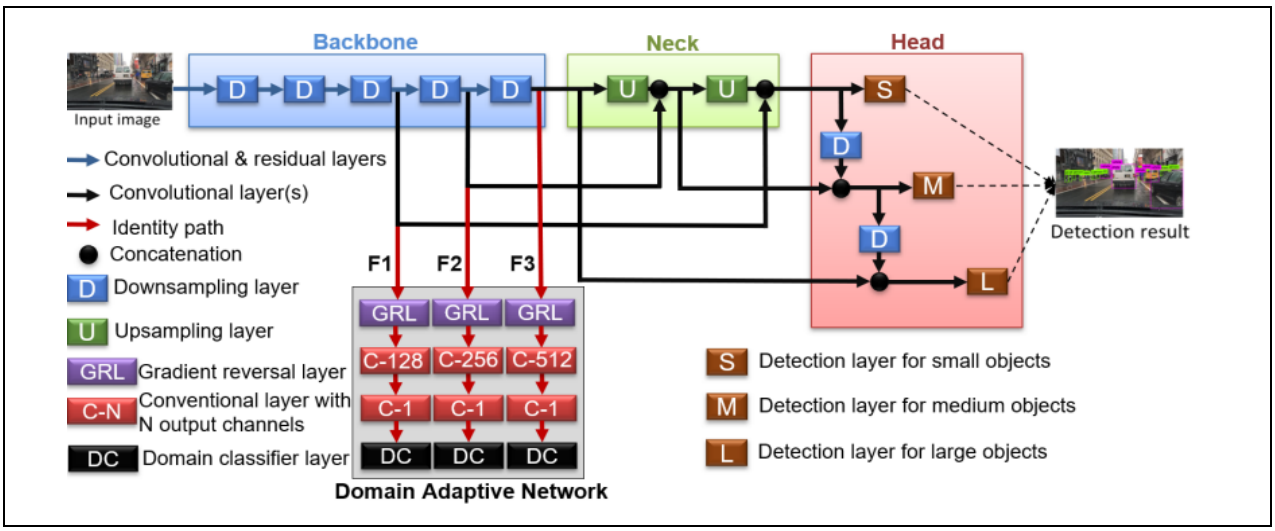
Input and Output		Feature of This Solution	Contribution & The Value of This Work
Input	**Output**	The framework introduces multi-scale domain adaptation for YOLOv4 object detection, enhancing domain-invariant feature extraction. It employs end-to-end training and operates in real-time, making it valuable for time-critical applications. A significant advantage is that it doesn't require target domain data annotation. Furthermore, it outperforms YOLOv4 in diverse autonomous driving scenarios, improving object detection performance.	In this work, we propose a novel multi-scale domain adaptation framework for the widely-used real-time object detector YOLOv4. The main contribution of our framework is the improvement of detection performance in the target domain thanks to robust domain-invariant features that reduce the impact of domain shift. An important advantage of our framework is that it can successfully adapt YOLOv4 to target domains without the need for any annotation, which is particularly useful in scenarios where labeling is time-consuming or expensive.
Input is a set of labeled data from a source domain and a set of unlabeled data from a target domain. The labeled data is used to train the YOLOv4 object detector, while the unlabeled data is used to adapt the detector to the target domain	The output is a framework that produces a feature representation that works across different domains. This helps to enhance the accuracy of object detection in the target domain by detecting objects and providing their class labels and bounding boxes.		

		Additionally, our proposed MS-DAYOLO outperforms state-of-the-art YOLOv4 in various testing scenarios for autonomous driving applications. The value of our work lies in its potential to enhance the performance of object detection systems in real-world scenarios where domain shift is prevalent.

Positive Impact of this Solution in This Project Domain	**Negative Impact of this Solution in This Project Domain**
The solution holds promise in enhancing object detection systems for autonomous driving by mitigating domain shifts without data annotation. Autonomous driving heavily relies on object detection, and this adaptation can boost system reliability and safety, crucial for this domain.	The suggested framework's success hinges on the quality and diversity of data in the target domain. Substantial differences and limited or biased data can still affect the object detector's performance due to domain shift. Additionally, compared to the original YOLOv4, the proposed framework may require more computational resources and training time.

Analyse This Work By Critical Thinking	**The Tools That Assessed this Work**	**What is the Structure of this Paper**
The analysis reveals noteworthy aspects of the work. The framework's novelty lies in its multi-path domain adaptation approach, distinct from prior methods. The evaluation demonstrates its superiority over YOLOv4 but lacks comparison with other domain adaptation methods. Acknowledged limitations include increased resource requirements and data quality dependence. Future work should consider different threshold values, alternative detection architectures, and applications beyond the current scope, providing potential avenues for further research and improvement.	TensorFlow, Open CV, Dataset, Matplotlib	I. abstract II. Introduction III. Related Work IV. Experiments V. Conclusion

Diagram/Flowchart



4	Smart Telephone Design - Caller Identification and Answering Machine	
Reference in APA format	M. M. Hasan, Lim Hooi Jiun, Ng Wei Cheun, and M. S. Shahid, "Smart telephone design-caller identification and answering machine," ICSE'98. 1998 IEEE International Conference on Semiconductor Electronics. Proceedings (Cat. No.98EX187), Bangi, Malaysia, 1998, pp. 217-222, doi: 10.1109/SMELEC.1998.781183.	
URL of the Reference	**Author's Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/abstract/document/781183	Md. Mahmud Hasan, mahmud@eng.upm.edu.my Lim Hooi Jiun, Ng Wei Chuen	Smart Telephone Design, Caller Identification, Answering Machine, microcontrollers, digital telephone networks, Frequency Shift Keying (FSK), modem tones, ring detector circuit, and outgoing message sender circuit.
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution and what is the problem that needs to be solved**	**What are the components of it?**
Caller Identification and Answering Machine functions in telephones are based on the potential use of low-cost microcontrollers.	The solution aims to explore the potential use of low-cost microcontrollers to implement advanced facilities and services related to everyday telephone use, such as Caller Identification and Answering Machine functions. The problem that needs to be solved is making these features accessible to a wider audience by using affordable technology. The paper also highlights the power and connectivity of the telephone system, which can be leveraged for other business purposes beyond traditional communication.	PIC16C84 microcontroller, Caller Identification (Caller ID) circuit, Answering Machine circuit, Tele-security system circuit, Ring detector circuit, Ring detector circuit, Outgoing message sender circuit, LCD display unit, EEPROM memory

The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process

The work presents a promising solution to the problem of making Caller Identification and Answering Machine functions accessible to a wider audience by using affordable technology.

	Process Steps	Advantage	Disadvantage (Limitation)
1	The problem that needs to be solved is how to make Caller Identification and Answering Machine functions accessible to a wider audience by using affordable technology.	The use of low-cost microcontrollers and digital telephone networks makes Caller Identification and Answering Machine functions accessible to a wider audience.	The design and implementation of the Smart Telephone Device require technical expertise and may be too complex for some users.
2	The researchers investigated the potential use of low-cost microcontrollers and digital telephone networks to incorporate advanced facilities and services associated with daily telephone use.	The Tele-security system circuit monitors the telephone line for any unauthorized access or tampering, providing enhanced security.	The Smart Telephone Device is designed to perform specific functions, including Caller Identification, Answering Machine, and Tele-security system, and may not meet the needs of all users.
3	The design of the Smart Telephone Device consists of several components, including the PIC16C84 microcontroller, Caller Identification circuit, Answering Machine circuit, Tele-security system circuit, Ring detector circuit, Outgoing message sender circuit, LCD display unit, and EEPROM memory.	The Answering Machine circuit records and plays back outgoing messages, and detects incoming messages, providing convenience to users.	It is evaluated on a specific set of objects commonly found in autonomous driving scenarios. It may not be effective for other types of objects or scenes
4	The implementation of the Smart Telephone Device by programming the PIC16C84 microcontroller in Intel 8052 8-bit microcontroller assembly language and assembling the hardware components.		
5	The testing of the Smart Telephone Device to ensure that it performs the intended functions, including Caller Identification, Answering Machine, and Tele-security system.		

Major Impact Factors in this Work

<Find all main factors and variables that are related to each solution. Then find the relationship between factors. (Independent variable) causes a change in (Dependent Variable) and it isn't possible that (Dependent Variable) could cause a change in (Independent Variable).

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
The performance of the Caller Identification system and the Answering Machine functionality can be considered as the dependent variable. It includes the accuracy and efficiency of identifying callers and recording messages.	The hardware components, tele-security components, and answering machine components act as the independent variables	The microcontroller, acting as the brain of the system, can be considered a moderating variable. It controls various logical operations and influences the interaction between independent variables, affecting the overall system performance.	The software program, specifically the assembly language program stored in the EEPROM of the microcontroller, serves as a mediating variable. It mediates between the hardware components and the overall system functionality, influencing how the system processes and responds to different inputs.

Relationship Among The Above 4 Variables in This article
The relationships among the variables are interconnected and cooperative. Hardware components collectivel contribute to system functionality, signaling mechanisms influence Caller ID performance, answerin machine components affect the performance of the answering machine, tele-security system component impact security system operation, and the microcontroller, along with its program, controls and mediate these interactions to achieve the desired smart telephone system functionality.

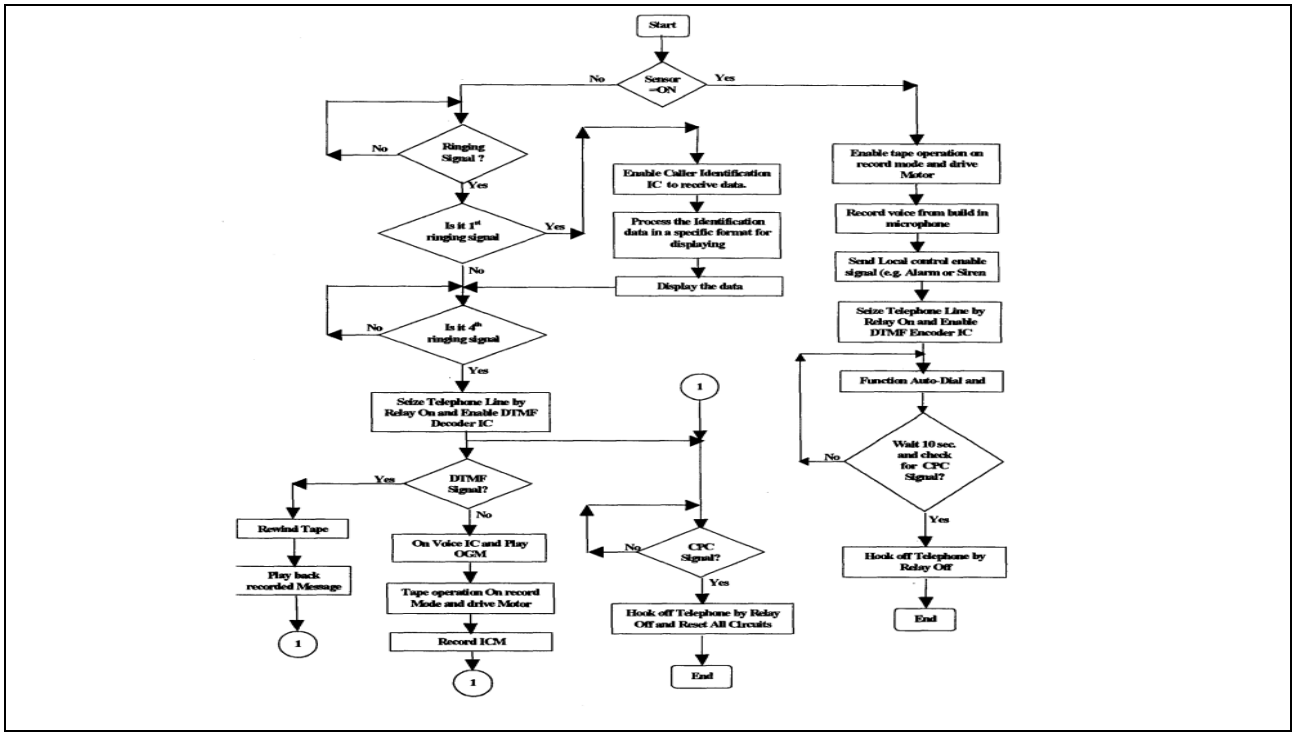
Input and Output	Feature of This Solution	Contribution & The Value of This Work
Input: The input is the components consisting of the telephone line and the ring detector circuit. The telephone line serves as the source of incoming calls, while the ring detector circuit detects the ringing signal of an incoming call and transmits a signal to the PIC16C84 **Output**: The output is the LCD display unit displays the Caller ID information decoded by the Caller ID circuit. The outgoing message sender circuit sends an outgoing message to the caller when the answering machine is activated. The tele-security system circuit	The Smart Telephone Device offers a range of features, including Caller Identification through a dedicated circuit displaying information on an LCD unit. It incorporates an answering machine to record messages during user unavailability and functions as a tele-security system, activating alarms against unauthorized access or tampering. Utilizing low-cost microcontrollers ensures affordability. Designed for compatibility with prevalent digital telephone networks, the device explores the application of Signaling System No. 7 (SS 7) for enhanced Caller ID in the realm of digital telephony networks.	The contribution is to create affordable and efficient telephone systems that can be used daily. The Smart Telephone Device is a comprehensive solution for household telephone use as it provides Caller ID, Answering Machine, and tele-security system functionalities. This has been made possible by the use of low-cost microcontrollers and digital telephone networks, making the device accessible to a wider audience. The paper also delves into the use of SS 7 for Caller ID, which is a signaling protocol used in digital telephone networks, providing insights into the technical aspects of the system. Overall, this study offers practical solutions for everyday telephone use and

microcontroller, which triggers the Caller ID and Answering Machine circuits.	activates an alarm if any unauthorized access or tampering is detected on the telephone line.		contributes significantly to the development of telephone technology.

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
This solution offers a cost-effective and comprehensive answer for daily telephone needs in the project domain. The Smart Telephone Device, with Caller ID, an answering machine, and a tele-security system, proves practical for household use. Its use of low-cost microcontrollers and compatibility with digital networks enhances accessibility, while the added security layer contributes value to the overall project.	The negative impact could be the reliance on digital telephone networks and SS 7 for Caller ID. If these systems experience technical difficulties or outages, the Smart Telephone Device may not function properly. Additionally, the tele-security system circuit may not be sufficient for high-security applications, as it is designed for household use.

Analyse This Work By Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
The work presents a cost-effective and easily accessible telephone solution that emphasizes affordability. The Smart Telephone Device comes with a Caller ID, an answering machine, and a tele-security system that adds an extra layer of security. However, the tele-security system may not be suitable for household use, and there may be limitations to the device's functionality in areas with poor network coverage. Despite these limitations, the paper presents a valuable and practical solution for everyday telephone use.	Smart Telephone System, Caller ID decode and display, Answering machine, Sensor, DTMF decoder IC (MC145436A), ISD1420 Single-chip voice recording playback IC, Opto-isolator, CPC circuit, Protection circuit, Motor, Microcontroller	I. abstract II. Introduction III. Related Work IV. Hardware Design V. Software Design VI. Conclusion

Diagram/Flowchart



5	Improved Trust in Human-Robot Collaboration With ChatGPT	
Reference in APA format	Y. Ye, H. You and J. Du, "Improved Trust in Human-Robot Collaboration With ChatGPT," in IEEE Access, vol. 11, pp. 55748-55754, 2023, doi: 10.1109/ACCESS.2023.3282111.	
URL of the Reference	**Author's Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/abstract/document/10141597	Yang ye, Hengxu you, Jing du, eric.du@essie.ufl.edu	ChatGPT, human factors, human-robot interaction, large language model, trust, LLMs (Language and Learning Models), HRC (Human-Robot Collaboration), Digital Object Identifier (DOI)
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution and what is the problem that needs to be solved**	**What are the components of it?**
The current solution presented in this paper is "Robo GPT". A robot control system utilizes ChatGPT to enhance communication and understanding between humans and robots in a human-robot collaboration assembly task.	The solution aims to introduce the Robo GPT solution, aiming to enhance human-robot collaboration (HRC) in assembly tasks by addressing communication and trust issues. The primary challenge is the lack of effective communication between humans and robots, leading to errors and safety concerns. The proposed solution involves integrating ChatGPT, an	ChatGPT, Robotic control modules, AI assistant,Human operator.

	AI-based natural language processing tool, with robotic control modules to create an intelligent AI robot control assistant. The goal is to assess the impact of ChatGPT on human operators and evaluate the overall effectiveness of the Robo GPT system in improving HRC performance.	

The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process

RoboGPT integrates language models with robotic controls, employing GPT3.5 for decision-making. It enables bidirectional communication, increasing transparency, and triggers ROS-based robotic control functions for assembly tasks.

	Process Steps	Advantage	Disadvantage (Limitation)
1	The Robo GPT workflow integrates LLMs with robotic control modules to build an intelligent AI robot control assistant. The workflow first transforms human operators' spoken language into textual input for the AI assistant to process.	Improved communication and collaboration between humans and robots	The use of AI and natural language processing tools may require additional training and expertise
2	The decision-making core of the AI assistant utilizes GPT3.5 to understand the information and respond. By considering the contextual information and evaluating the ambiguity of information, GPT3.5 generates natural responses to either further clarify the information with the human operators via conversations or control the robot.	Increased transparency and reduced ambiguity in the decision-making process	The system may not be suitable for all types of assembly tasks or environments
3	The Robo GPT AI assistant generates prompts, presents the prompts to human operators, and waits for further instructions. Such bidirectional communication clarifies the intention of both human operators and the robotic system, which could increase transparency and reduce ambiguity.	Enhanced trust in the robotic system	The cost of implementing the system may be high

4	AI assistant considers the information adequate for decision-making, the responses will be sent to a decoder which further processes the commands into Robotic Operating System (ROS) topics and triggers robotic control functions to perform tasks correspondingly. In the context of HRC in assembly, the major robotic control functions are grab, move, and drop.	Better task performance and efficiency	

Major Impact Factors in this Work

<Find all main factors and variables that are related to each solution. Then find the relationship between factors. (Independent variable) causes a change in (Dependent Variable) and it isn't possible that (Dependent Variable) could cause a change in (Independent Variable).

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
Trust in Human-Robot Collaboration the participants' level if trust in the collaboration between humans and robots.	Use of ChatGPT in Robotic Control, it represents the presence or absence of ChatGPT as an intelligent assistant in controlling the robot.	Task performance might influence the strength	Cognitive Load might help explain the process through which the independent variable (ChatGPT) influences the dependent variable (Trust).

Relationship Among The Above 4 Variables in This article

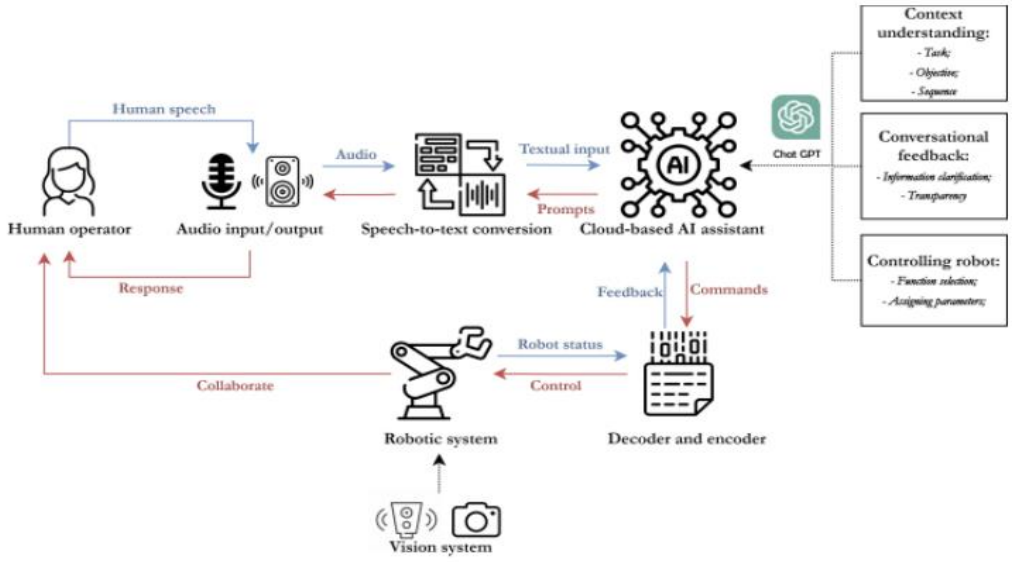
The use of ChatGPT in robotic control has direct effects on trust, potentially mediated by reduced cognitive load and moderated by task performance. Task performance itself might also have a direct influence on trust.

Input and Output		Feature of This Solution	Contribution & The Value of This Work
Input	**Output**	The feature of this solution is the integration of ChatGPT in Robo GPT facilitates bidirectional communication, enhancing transparency and efficiency in decision-making between humans and robots. ChatGPT's natural language processing capabilities simplify human-robot interaction, enabling autonomous decision-making and memory retention for improved collaboration. Despite the text-based I/O, the system shows promise in fostering better partner relationships and	The introduction of ChatGPT within the Robo GPT system contributes significantly to human-robot collaboration by enhancing transparency and decision-making efficiency. This improvement builds trust in the robotic system, particularly crucial in industrial and manufacturing environments where optimized task performance is essential. The integration of ChatGPT showcases advancements in natural language processing,
The input of this solution is the use of ChatGPT in the Robo GPT system, which enables bidirectional communication between humans and robots and improves the	The output of this solution is improved trust in the robotic system and better task performance. The use of ChatGPT in the Robo GPT system enables bidirectional communicatio		

transparency and efficiency of the decision-making process. It also includes the implementation of robotic control modules and an AI assistant that continuously monitors and controls the robot's actions throughout the assembly task.	n between humans and robots, which improves the transparency and efficiency of the decision-making process.	elevating overall task performance in human-robot collaboration scenarios.	emphasizing its potential to elevate communication and collaboration between humans and robots. The human-subject experiment conducted in the study serves as critical validation for the efficacy of the HRC design, utilizing language models and deepening our understanding of human trust in the collaborative process.

Positive Impact of this Solution in This Project Domain	**Negative Impact of this Solution in This Project Domain**
The integration of ChatGPT in Robo GPT enhances transparency and efficiency of decision-making, resulting in better task performance and productivity. Bi-directional communication through ChatGPT enables natural human-robot interaction, essential in industrial and manufacturing environments, where LLMs' use as intelligent robot assistants improves decision-making efficiency and transparency.	Robo GPT faces technical limitations such as reliance on high-quality data and computational resources. Implementation costs and user acceptance challenges, coupled with ethical considerations regarding privacy and security, need careful attention.

Analyse This Work By Critical Thinking	**The Tools That Assessed this Work**	**What is the Structure of this Paper**
This study evaluates the impact of ChatGPT on trust in human-robot collaboration. The use of Large Language Models like ChatGPT improves semantic understanding in communication. The experiment using Robo GPT shows increased trust. However, potential miscommunication calls for robust and adaptable AI systems. The study examines ChatGPT's influence on trust, considering opportunities and limitations.	Human-subject experiment, NASA TLX questionnaire, Statistical Analysis, Performance Metrics	I. abstract II. Introduction III. Related Work IV. Methodology V. Results VI. Discussions VII. Conclusion
Diagram/Flowchart		



6	YOLO-based Threat Object Detection in X-ray Images	
Reference in APA format	R. L. Galvez, E. P. Dadios, A. A. Bandala and R. R. P. Vicerra, "YOLO-based Threat Object Detection in X-ray Images," 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Laoag, Philippines, 2019, pp. 1-5, doi: 10.1109/HNICEM48295.2019.9073599.	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/document/9073599	Reagan L. Galvez, reagan_galvez@dlsu.edu.ph Elmer P. Dadios, Argel A. Bandala, Ryan Rhay P. Vicerra	automated detection, convolutional neural networks, threat object, transfer learning, X-ray image, YOLO
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution & What is the problem that needs to be solved**	**What are the components of it?**
YOLO-based object detection algorithm to detect threat objects in X-ray images.	The goal of the solution is to automate the detection of threat objects in X-ray images, which is a tedious and time-consuming task for human inspectors. The problem that needs to be solved is the possibility of missed detections during peak hours due to the limited time available to scan and analyze the baggage. It can be solved by using a fast object detector as decision support for threat object detection.	YOLOv3 Architecture, Darknet-53 Feature Extractor
The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process		

The algorithm uses CNN and a fast object detector to detect threat objects in X-ray images. It is trained on a dataset of threat and non-threat objects and can quickly detect multiple objects in one image

	Process Steps	Advantage	Disadvantage (Limitation)
1	A dataset of X-ray images containing threat objects and non-threat objects is collected.	The YOLO-based object detection algorithm is fast and can detect objects in real-time.	The algorithm may not be as accurate as other object detection algorithms that use more complex architectures.
2	The X-ray images are preprocessed to enhance the quality of the images and to normalize the pixel values.	The algorithm can detect multiple objects in a single image	The algorithm may not be able to detect small or occluded objects.
3	The YOLOv3 architecture is trained on the preprocessed dataset using both transfer learning and training from scratch. The performance of the two methods is compared, and it is found that training from scratch achieves better results.	The algorithm can be trained on a new dataset to detect different types of objects	The algorithm may require a large amount of training data to achieve good performance.
4	The trained model is evaluated on a test set of X-ray images to measure its performance in detecting threat objects.		
5	The trained model is deployed as a fast object detector to assist human inspectors in detecting threat objects in X-ray images.		

Major Impact Factors in this Work

<Find all main factors and variables that are related to each solution. Then find the relationship between factors. (Independent variable) causes a change in (Dependent Variable) and it isn't possible that (Dependent Variable) could cause a change in (Independent Variable).

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable

Mean Average Precision (mAP) represents the accuracy of threat object detection in X-ray images. It is the measure used to evaluate the performance of the YOLO-based object detector.	transfer learning and training from scratch and input images acted as the independent variables as we used the YOLO model for training.	Task complexity, particularly the thickness of threat objects, could moderate the performance of the YOLO-based detector.	Inference Speed, the time taken for the model to make predictions per image, Average Precision (AP) for different threat object classes, and the performance on individual classes can mediate the overall mAP, providing insights into which types of threat objects the model is better at detecting.

Relationship Among The Above 4 Variables in This article

The dependent variable is the mean average precision, while independent variables include the training approach and input image size. The moderating variable is task complexity, and mediating variables include time per inference and the performance on specific threat object classes.

Input and Output		Feature of This Solution	Contribution & The Value of This Work
Input	**Output**	The proposed solution is a YOLO-based object detection algorithm used in this solution is fast and efficient, and it achieves a mean average precision (mAP) of up to 52.40% in a multi-scale image. The algorithm uses a feature extractor called Darknet-53, which has 53 convolutional layers trained on the ImageNet dataset. The YOLOv3 architecture is used to detect threat objects in X-ray images, and it can accurately classify objects over handcrafted features used by other object detection algorithms.	The contribution of this work is the development of a fast and efficient YOLO-based object detection algorithm for automated detection of threat objects in X-ray images. The algorithm achieves high accuracy in detecting threat objects, and it outperforms other object detection algorithms that use handcrafted features. The value of this work lies in its potential to improve security in public places such as airports, train stations, and commercial establishments by providing a fast and effective way to detect threat objects in X-ray images.
The input is X-ray images of baggage or other items that need to be scanned for potential threat objects.	Output is the list of detected objects in the input X-ray image, along with their corresponding class labels and bounding boxes. The algorithm can detect multiple objects in a single image and is fast		

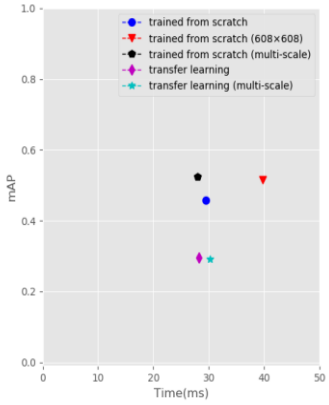
Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
The positive impact is it provides a fast and efficient way to detect threat objects in X-ray images, which can improve security in public places such as airports, train stations, and commercial establishments. By automating the detection process, the workload of security personnel can be reduced, and the public's overall safety can be	The use of YOLO-based object detection algorithms may require additional hardware and software resources, which could be expensive for some organizations.

improved.	

Analyse This Work By Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
The solution addresses X-ray image threat object detection, emphasizing public security enhancement. YOLO architecture is detailed, and outcomes are well-presented, but limited by a small dataset. The solution significantly advances threat object detection for heightened security.	You Only Look Once (YOLO) architecture, a dataset of scanned X-ray images of an improvised explosive device (IED), TensorFlow or PyTorch	I. Abstract II. Introduction III. Methodology IV. Results and Discussions V. Conclusion

Diagram/Flowchart

model	mAP	$AP_{battery}$	AP_{mortar}	AP_{wires}	time(ms)
Trained from scratch	0.4589	0.3733	0.8466	0.1569	29.41
Trained from scratch (608×608)	0.5148	0.3986	0.9401	0.2058	39.72
Trained from scratch (multi-scale)	**0.5240**	0.4627	0.9209	0.1885	**27.99**
Transfer learning	0.2954	0.0383	0.8115	0.0363	28.22
Transfer learning (multi-scale)	0.2917	0.0216	0.8096	0.0439	30.19



7	Object Detection for Autonomous Driving using YOLO [You Only Look Once] algorithm	
Reference in APA format	A. Sarda, S. Dixit and A. Bhan, "Object Detection for Autonomous Driving using YOLO [You Only Look Once] algorithm," *2021* Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 2021, pp. 1370-1374, doi: 10.1109/ICICV50876.2021.9388577.	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/document/9388577	Abhishek Sarda, Dr. Shubhra Dixit, Dr. Anupama Bhan	Deep learning, Object detection, Classification algorithms, Autonomous vehicles, computer vision, learning (artificial intelligence), state-of-the-art algorithm YOLO, and autonomous driving.
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ...	**The Goal (Objective) of this Solution & What is the problem that need to be solved**	**What are the components of it?**

etc)		
Object detection using YOLOV4 algorithm	The goal of this solution is to enhance safety and efficiency by precisely identifying and categorizing objects on the road. This addresses the issue of accidents stemming from human error, reducing them through autonomous driving technology.	The solution uses YOLOv4's CSPDarknet53 for image feature extraction, a SPP (Spatial pyramid pooling) and PAN (Path Aggregation Network) layers for processing of images, a YOLOv4 detection model for drawing bounding box and class probability predictions, its data augmentation techniques include Mosaic augmentation and DropBlock regularization, it uses Mish and Leaky ReLU activation functions, the Darknet training framework, it was trained on google colab using GPU.

The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process

The YOLO algorithm for object detection in autonomous driving involves YOLOV4 model where the input image is processed through a model's neural network to detect and classify objects in real-time.

	Process Steps	Advantage	Disadvantage (Limitation)
1	The YOLO model trains on the custom image data we provide to identify objects on images in google colab environment using yolo's training algorithm.	Training the YOLO algorithm on custom classes allows for the detection of specific objects.	The accuracy of the algorithm depends on the quality and quantity of the training data.
2	The new unseen input image is processed through a neural network to detect and classify objects in real-time.	YOLO is a single-shot detector, which means it can detect and classify objects in real-time.	It may not perform as well as multi-stage detectors in detecting small objects.
3	The objects on the image are identified and classified into different objects by drawing the bounding boxes around them.	The Identified objects are separated by bounding boxes.	

Major Impact Factors in this Work

The major impact factors in this work include addressing utilizing the advanced YOLO algorithm, customizing training for relevant classes, rigorous performance evaluation, and the potential for real-world safety improvements in autonomous driving.

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
Object detection accuracy	Deep learning algorithms (YOLO)		
Speed of object detection	Training dataset		
Ability to detect specific objects	Object detection algorithms		
	Pre-processing techniques		

Relationship Among the Above 4 Variables in This article

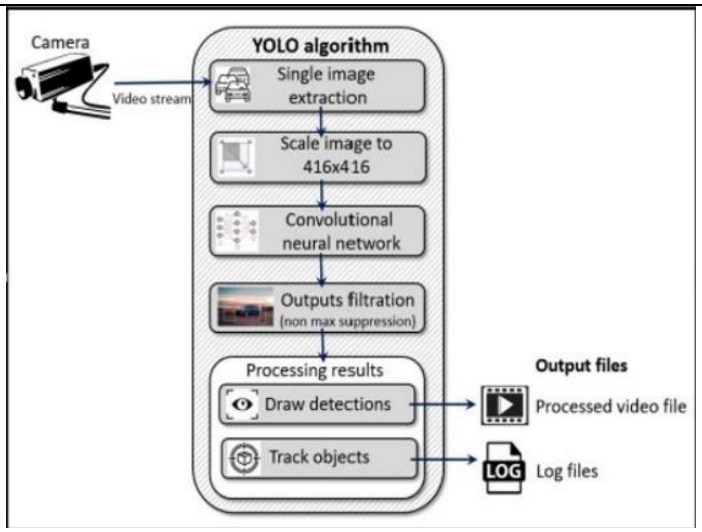
The model uses yolo model for object detection through which object detection speed and accuracy are improved. Through the custom training model can detect specific objects on the image.

Input and Output	Feature of This Solution	Contribution & The Value of This Work
<table><tr><th>Input</th><th>Output</th></tr><tr><td>Image or video feed from a camera sensor</td><td>Labelled image or video feed that shows the detected objects and their correspondin g bounding boxes</td></tr></table>	This algorithm is a single-shot detector that allows for real-time detection and classification of objects, making it suitable for use in autonomous driving.	This work offers a custom object detection model for autonomous driving achieving an impressive 74.6% mAP (mean average precision).

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
The solution of using YOLO algorithm can significantly reduce accidents and improve safety.	The YOLO algorithm may have false positives and false negatives, which can lead to incorrect object detection and potential accidents.

Analyse This Work by Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
The authors provided a detailed explanation of the algorithm and its training process, and presented impressive results on a custom dataset. However, a more thorough discussion of limitations and comparisons with other algorithms would have enhanced the analysis.	The tools used to assess this work include the YOLO algorithm for object detection, the Darknet repository for training the YOLO model. Additionally, Google Colab for setting up the training environment and analyzing the results.	1. Abstract 2. Introduction 3. Literature Review 4. YOLO Architecture 5. Methodology 6. Data 7. Initial Result 8. Custom Training 9. Result 10. Conclusion 11. References

Diagram/Flowchart



8	Edge detection-based boundary box construction algorithm for improving the precision of object detection in YOLOv3	
Reference in APA format	S. T. Blue and M. Brindha, "Edge detection-based boundary box construction algorithm for improving the precision of object detection in YOLOv3," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-5, doi: 10.1109/ICCCNT45670.2019.8944852	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/document/8944852	Shaji Thorn Blue, M. Brindha	object detection, boundary boxes, precision, YOLOv3, COCO dataset, edge detection, pixel values, IOU comparison.
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution & What is the problem that need to be solved**	**What are the components of it?**
Edge detection-based boundary box construction algorithm for improving the precision of object detection using YOLOv3	The aim of this work is improving the precision of boundary boxes that will be been drawn after objects has been detected on image in order to improve the accuracy of object localization and identification within images.	The proposed work includes the YOLOv3 algorithm, which is used to obtain the coordinates of boundary boxes and the class of the object. After YOLOv3 outputs an image, pre-processing is applied to obtain an image with the edges of the detected object. The proposed work also includes an edge detection algorithm, image partitioning into four parts, and the construction of a new boundary box based on threshold values.
The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage &		

Disadvantage of Each Step in This Process

The algorithm evaluates the accuracy of object localization and identification within images using YOLOv3 algorithm through which edges can be detected so that boundary boxes can be drawn accurately.

	Process Steps	Advantage	Disadvantage (Limitation)
1	The model applies the algorithm only on the part of the image where YOLOv3 has already detected an object.	Reduces computation and improves algorithm speed.	May miss objects outside the detected area.
2	The image is blurred to make colour transitions smooth and remove noise and outlier pixels.	Smooths colour transitions and removes noise.	May blur important details.
3	The proposed work uses the Canny edge detection algorithm to detect edges in the blurred image.	Accurately detects edges.	May miss some edges or produce false positives.
4	The image is divided into four parts to focus on the areas where boundary box edges are more likely to be found.	Focuses on areas where boundary box edges are more likely to be found.	May miss boundary box edges in the middle of the image.
5	Each divided image is analysed, and a threshold value is calculated based on the density of white pixels. A new boundary box is drawn based on the calculated threshold value.	Draws more precise boundary boxes.	May struggle with sharp objects and noisy images.

Major Impact Factors in this Work

Key impact factors of the proposed work include precision enhancement in boundary boxes, building upon the YOLOv3 object detection system, leveraging a pretrained COCO dataset, edge detection algorithm.

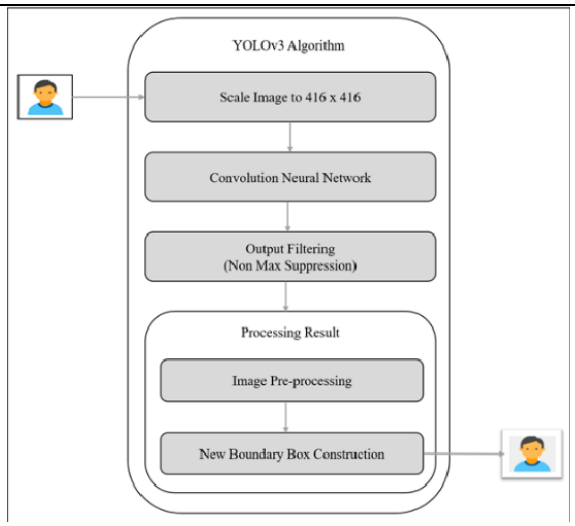
Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening variable
Edge detection	Pretrained dataset		
Precision of Boundary Boxes	Threshold value		
	Image preprocessing techniques		

Relationship Among the Above 4 Variables in This article

Through edge detection of the objects, we can find the coordinates of the detected objects and a threshol value is calculated by which we can draw new accurate boundary boxes.

Input and Output	Feature of This Solution	Contribution & The Value of This Work

<table>
<tr>
<td>
<table>
<tr><th>Input</th><th>Output</th></tr>
<tr><td>Image or video feed from a camera sensor</td><td>more precise boundary box compared to YOLOv3.</td></tr>
</table>
</td>
<td>This solution prioritizes precision by enhancing the accuracy of boundary boxes around detected objects, outperforming YOLOv3 through edge detection and pixel values. Leveraging YOLOv3 as a foundation, it utilizes pre-trained COCO data, determines threshold values for boundary sides, and excels in producing accurate boundary boxes.</td>
<td>This work contributes to refining object detection precision by leveraging YOLOv3 and the pre-trained COCO dataset. It utilizes edge detection and pixel analysis to create more accurate boundary boxes around objects, improving localization.</td>
</tr>
<tr><th colspan="2">Positive Impact of this Solution in This Project Domain</th><th>Negative Impact of this Solution in This Project Domain</th></tr>
<tr><td colspan="2">The solution enhances object detection precision with edge detection and pixel analysis.</td><td>The solution's drawbacks include reduced real-time performance due to added complexity, and the potential trade-off between precision and speed.</td></tr>
<tr><th>Analyse This Work by Critical Thinking</th><th>The Tools That Assessed this Work</th><th>What is the Structure of this Paper</th></tr>
<tr><td>The analysis highlights the integration of YOLOv3 for object detection, followed by image pre-processing which includes Gaussian blur and Canny edge detection. New boundary boxes are constructed based on white pixel density analysis. Results indicate improved accuracy over YOLOv3, particularly in terms of IOU (intersection over union) values, though challenges exist in handling sharp objects and noisy images.</td><td>Evaluation relies on a comparison graph, demonstrating the work's superior accuracy in boundary box predictions over YOLOv3.</td><td>1. Introduction
2. Proposed Work
3. YOLOv3 Object Detection Algorithm
4. Image Pre-Processing
5. Boundary-Box Construction
6. Result and Analysis
7. Conclusion
8. References</td></tr>
<tr><th colspan="3">Diagram/Flowchart</th></tr>
<tr><td colspan="3"></td></tr>
</table>



9	Moving Object Detection and Tracking Using Convolutional Neural Networks	
Reference in APA format	S. Mane and S. Mangale, "Moving Object Detection and Tracking Using Convolutional Neural Networks," 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2018, pp. 1809-1813, doi: 10.1109/ICCONS.2018.8662921	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/document/8662921	Shraddha Mane, Prof. Supriya Mangale	Object detection, Object tracking, Convolutional neural networks, Computational modelling, cellular neural nets, convolutional neural nets, image sequences, convolutional neural networks, computer vision algorithm, TensorFlow object detection API, object tracking algorithm, moving-object detection.
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution & What is the problem that need to be solved**	**What are the components of it?**
Moving object detection and tracking using CNN and TensorFlow object detection API.	The goal of this solution is to accurately detect and track moving objects in real-time video streams. The problem that needs to be solved is the difficulty of detecting and tracking objects accurately and robustly in real-world scenarios.	Moving object detection using TensorFlow object detection API, Location of detected object passed to object tracking algorithm, CNN-based object tracking algorithm for robust object detection, Offline training to learn spatial and temporal classes for human tracking, shift variant architecture to combine global features and local characteristics
The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage &		

Disadvantage of Each Step in This Process

Our approach uses TensorFlow object detection API for moving object detection and a CNN-based tracking algorithm for robust object detection. Offline training and shift-variant architecture are also used for improved accuracy.

	Process Steps	Advantage	Disadvantage (Limitation)
1	We use the TensorFlow object detection API to detect moving objects in real-time video streams.	This API provides a simple and efficient way to construct, train, and detect models for object detection.	This may require some knowledge of TensorFlow and deep learning methods.
2	Once the object is detected, we use a CNN-based tracking algorithm to track the object's location over time.	This algorithm is designed to be robust to variations in illumination and occlusion.	This may require a large number of computational resources.
3	We use offline training to learn spatial and temporal classes for human tracking.	This improves the accuracy of object detection and tracking.	Offline training may require a large amount of data and time to train the model.
4	We use a shift-variant architecture to combine global features and local characteristics.	This improves the accuracy of object detection and tracking.	This may require some knowledge of computer vision and image processing techniques.

Major Impact Factors in this Work

Key impact factors of this work include a remarkable 90.88% accuracy, robustness against illumination and occlusion, computational efficiency through TensorFlow, adaptability to diverse scenarios, and the novelty of fusing TensorFlow object detection API and CNN-based tracking.

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
Sensitivity	True Positive (TP) values		
Specificity	True Negative (TN) values		
Accuracy	False Positive (FP) values		
	False Negative (FN) values		

Relationship Among the Above 4 Variables in This article

We can calculate the sensitivity, specificity and accuracy of the model's performance by using TP, TN, FP, FN values which are the probabilistic values of the object detection scenarios.

Input and Output	Feature of This Solution	Contribution & The Value of This Work
Input / Output	The key features of this solution include a novel Tensor flow-based object detection and CNN object tracking system for robust object detection and tracking in complex scenes. Evaluation	The contribution of this work is the development of an algorithm for object detection and tracking using a combination of object detection with Tensor Flow and object tracking with CNN. The algorithm is tested on

Image or video feed from a camera sensor	Qualitative and quantitative analysis of the object detection and tracking results.	involves sensitivity, specificity, and accuracy metrics, derived from TP, TN, FP, and FN values, assessing the system's performance in identifying moving and stationary objects.	various video sequences and achieves good performance in terms of sensitivity, specificity, and accuracy.
Positive Impact of this Solution in This Project Domain		**Negative Impact of this Solution in This Project Domain**	
The solution offers a novel, robust approach to object detection and tracking in complex scenarios which enhances accuracy and exhibits efficient performance in Python-based testing.		Issues in the project domain may arise from dynamic backgrounds, clutter leading to tracking false alarms, potential misclassification of objects, and the pivotal selection of an appropriate classifier influencing algorithm performance.	
Analyse This Work by Critical Thinking		**The Tools That Assessed this Work**	**What is the Structure of this Paper**
This solution focuses on object detection and tracking through a unique approach involving TensorFlow's object detection API and CNN algorithms. Evaluation across diverse video sequences reveals impressive sensitivity, specificity, and accuracy metrics. The use of a shift variant architecture enhances CNN performance. The approach seems promising in terms of accuracy and robustness.		The tools used to assess this work include sensitivity, specificity, and accuracy parameters. These parameters are calculated using True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).	1. Introduction 2. Literature Survey 3. Methodology 4. Results 5. Conclusion 6. References
Diagram/Flowchart			

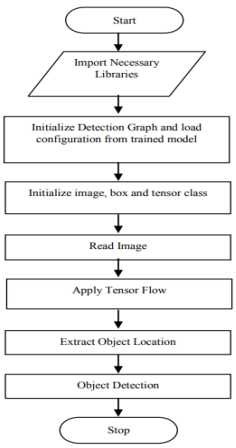


Fig. 2. TensorFlow Based Object detection flowchart

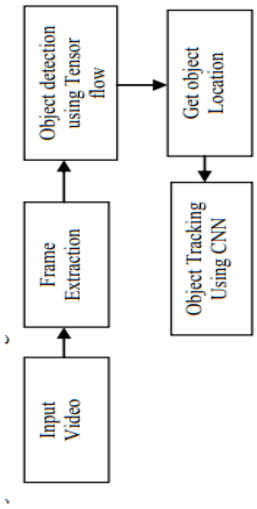


Fig. 1. Block Diagramof proposed system

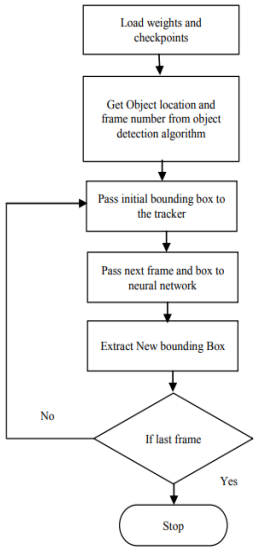


Fig. 3. Flowchart for object detection

10	Enhanced Missing Object Detection System using YOLO	
Reference in APA format	R. Menaka, N. Archana, R. Dhanagopal and R. Ramesh, "Enhanced Missing Object Detection System using YOLO," *2020* 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020, pp. 1407-1411, doi: 10.1109/ICACCS48705.2020.9074278	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/document/9074278	Menaka, Archana, Dhana Gopal, Ramesh	YOLO, object identification, continuous object detection, class probabilities, CNN, regression algorithms, image segmentation, object recognition, intersection over union (IoU), non-max suppression, anchor box, object co-detection, anomaly detection, CNN architecture, frame analysis.
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution & What is the problem that need to be solved**	**What are the components of it?**
Enhanced Missing Object Detection System using YOLO	The aim of this solution is to develop an Enhanced Missing Object Detection System using the YOLO approach. The problem that this system aims to solve is the identification of missing objects in real-time by	Object Identification Techniques (YOLO), Bounding Box Prediction, Anchor Box Technique, Comparison with Other Real-Time Systems, Frame Analysis.

	continuously analysing frames and alerting the user.	

The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process

The solution involves dividing the image into grids, predicting bounding boxes and class probabilities for objects in each grid cell, comparing objects in current and previous frames to identify missing objects, and generating alerts. It utilizes the YOLO algorithm for real-time object detection and classification.

	Process Steps	Advantage	Disadvantage (Limitation)
1	The YOLO algorithm is used to identify objects in each frame of the input image or video. This step involves predicting bounding boxes and class probabilities for objects using a single neural network.	The YOLO algorithm helps in detecting objects directly from full image in a single pass, making it faster compared to other detection algorithms.	The YOLO algorithm may make some localization errors but predicts fewer false positives in the background.
2	The current frame is compared with the previous frame to check if any objects that were present in the previous frame are missing in the current frame.	Comparing frames helps in continuous monitoring of the scene without the need for manual observation.	This step assumes that the objects of interest remain consistent in appearance and position between frames, which may not always be the case in real-world scenarios with complex scenes.
3	If any objects are detected as missing, an alert is generated to notify the user about the absence of those objects.	The alert system ensures that the user is immediately informed about any missing objects, allowing for timely action.	The effectiveness of the alert system depends on the reliability of the communication channels and the responsiveness of the user to the alerts.

Major Impact Factors in this Work

The major impact factors of this work include real-time object detection, efficiency, and speed, because of YOLO's single-network architecture. The potential for object co-detection further enhances the system's performance and applicability in surveillance and monitoring applications.

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
Mean Average Precision (mAP)	Training and validation datasets		
Detection accuracy	Learning rate		
Object localization accuracy	Batch size		
Speed and efficiency	Loss function		

False positive and false negative rates	Data augmentation		

Relationship Among the Above 4 Variables in This article

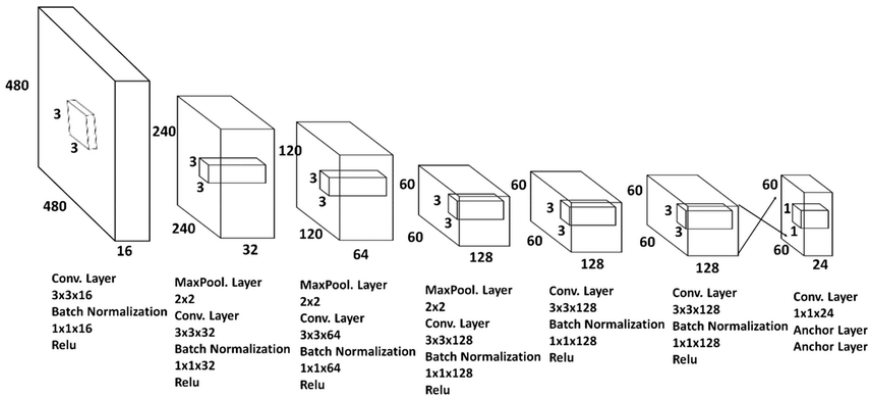
Because of the variables like learning rate, batch size, loss function the mAP, object detection and localization, speed and efficiency of the model changes. And through these variables model's performance is measured.

Input and Output	Feature of This Solution	Contribution & The Value of This Work
Input: Picture that is divided into 3x3 grids **Output**: Class labels and bounding boxes for the objects	This solution features the use of convolutional layers trained on the ImageNet dataset for object detection. It leverages the efficient YOLO algorithm for direct prediction of bounding boxes and class probabilities. A frame analysis algorithm stores objects across frames, enabling missing object detection. Training involves YOLO's pretrained model and fine-tuning.	This work's contribution lies in developing a YOLO-based detection system for accurate object recognition and locating missing objects.

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
The YOLO algorithm's efficiency, real-time capabilities, and adaptability across domains, including surveillance and face recognition, provide a valuable solution for accurate object detection and missing object alerts.	Challenges in the project domain include potential data limitations for YOLO's training, high computational demands during training, and concerns regarding overfitting prevention strategies.

Analyse This Work by Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
This work outlines a system that employs the YOLO algorithm for real-time object detection, primarily for surveillance. The methodology includes convolutional layers, frame analysis, and data augmentation techniques. The system's potential applications span surveillance, object detection, with prospects for further enhancements like object-co detection and GSM integration for user alerts.	The tools/metrics that are used to assess this work are mean average precision, object localization accuracy, false positive and false negative rates of the model.	1. Abstract 2. Introduction 3. Comparison with other systems 4. Design and Architecture 5. Future scope 6. Conclusion 7. References

Diagram/Flowchart



11	A Human-Tracking Robot Using Ultra Wideband Technology	
Reference in APA format	T. Feng, Y. Yu, L. Wu, Y. Bai, Z. Xiao and Z. Lu, "A Human-Tracking Robot Using Ultra Wideband Technology," in IEEE Access, vol. 6, pp. 42541-42550, 2018, doi: 10.1109/ACCESS.2018.2859754.	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/document/8419698	Tao Feng, Yao Yu, Lin Wu, Yanru Bai, Ziang Xiao, Zhen Lu.	Integrated navigation, neural networks-based control, wireless sensor networks, ultra-wideband technology, human-tracking, measurement errors, hyperbolic positioning algorithm, virtual spring model, autonomous movement controlling algorithms, vision-based techniques, deep network-based models, and object detection.
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution & What is the problem that need to be solved**	**What are the components of it?**
Human-tracking in a robot using ultra-wideband technology.	The goal of the proposed solution is to implement a human-tracking robot using ultra-wideband (UWB) technology. The problem that needs to be solved is the detection and calculation of the position of the target person, which is essential for the implementation of the human-tracking function.	The proposed solution encompasses a range of hardware components to support its functionalities, including Ultra-wideband (UWB) technology for human tracking, ultrasound modules to estimate distances and enable obstacle avoidance, a Bluetooth module for data communication, an A33 SOC responsible for UWB data reception, calibration, filtering, and executing the human tracking algorithm.
The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process		

The proposed solution uses UWB technology. The UWB data is calibrated and filtered in the A33 SOC, while the STM32F103 MCU controls the robot's movement, data communication, and distance estimation. The human-tracking algorithm uses a modified virtual spring model to track the target person smoothly.

	Process Steps	Advantage	Disadvantage (Limitation)
1	The solution is based on UWB technology, which provides accurate tracking of the target person.	Robustness against environmental conditions.	
2	A modified hyperbolic positioning algorithm is used to overcome measurement errors in determining the position and orientation of the target person.		
3	A modified virtual spring model is implemented in the robot to track the target person, allowing for effective following and tracking.	Enables effective tracking and following of the target person.	
4	The target person's position and orientation are obtained with minimum errors using a moving-average filter algorithm.	Provides target person position and orientation data accurately.	
5	The obtained data from the filter algorithm are used for path planning and control of the robot.	Enables efficient movement and tracking of the target person.	

Major Impact Factors in this Work

The solution utilizes Ultra-wideband (UWB) technology for precise target person tracking, even in challenging environments. It employs a modified hyperbolic positioning algorithm to mitigate measurement errors and integrates a virtual spring model for effective tracking. A moving-average filter enhances accuracy, enabling efficient path planning and control for tracking the target person.

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
Position and orientation of the target person	Ultra-wideband (UWB) technology		
Tracking performance	Modified hyperbolic positioning algorithm		
Path planning and control	Modified virtual spring model		

Relationship Among the Above 4 Variables in This article

The current model uses a UWB technology, a hyperbolic positioning algorithm, and a virtual spring model in order to find a target person's position and orientation, track the target person and control the robot's movements respectively.

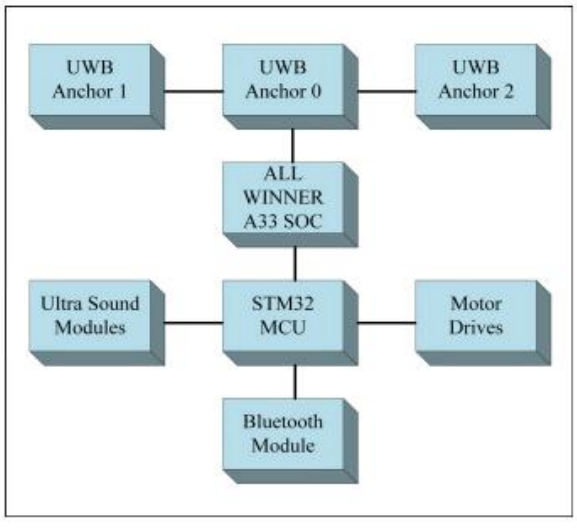
Input and Output		Feature of This Solution	Contribution & The Value of This Work
Input	**Output**	The robot features leveraging a UWB technology, it excels in precision and anti-interference	This work presents a human-tracking robot using UWB technology, featuring a modified

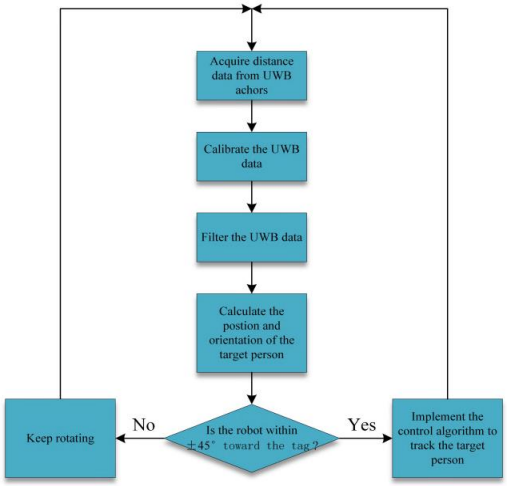
UWB Data	Target's Tracking Position	abilities providing a reliable human tracking model. The human-tracking algorithm, employing a modified virtual spring model, ensures robust performance in various environments. Experimental results validate its effectiveness in distance detection and overcoming measurement errors.	hyperbolic positioning algorithm and a virtual spring model for precise tracking.

Positive Impact of this Solution in This Project Domain	**Negative Impact of this Solution in This Project Domain**
UWB technology positively impacts the human-tracking robot by offering precise positioning, robustness in diverse environments, low system complexity, high location precision, and potential mainstream adoption.	Challenges include limitations in outdoor environments, higher hardware costs, potential measurement errors, tracking capability restrictions, and nonholonomic constraints, impacting the human-tracking robot's application.

Analyse This Work by Critical Thinking	**The Tools That Assessed this Work**	**What is the Structure of this Paper**
It is evident that the research paper focuses on the development of a human-tracking robot using UWB technology. The authors provide details about the robot's hardware architecture, dynamic model, human tracking algorithm, and experimental results.	The tools used to assess this work were the hyperbolic positioning algorithm, ultrasound sensors, UWB anchors, and the A33 SOC and STM32F103 MCU controllers. Additionally, the moving-average filter and human tracking algorithm were used for evaluation.	1. Introduction 2. System Architecture 3. Robot Hardware Architecture 4. UWB Data Calibration and Tag Position Calculation 5. Moving-Average Filter 6. Dynamic Model of the Robot 7. Human Tracking Algorithm 8. Experimental Results and Discussion 9. Conclusion

Diagram/Flowchart





12	ROSGPT: Next-Generation Human-Robot Interaction with ChatGPT and ROS	
Reference in APA format	Koubaa, A. (2023). ROSGPT: Next-Generation Human Robot Interaction with ChatGPT and ROS Preprints: https://doi.org/10.20944/preprints202304.0827.	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://www.preprints.org/manuscript/202304.0827/v3	Anis Koubaa	large language models (LLMs), ChatGPT, human-robot interaction, robotics automation, suggestion mode, natural language processing, deep learning techniques, transformer models, fine-tuning, BERT, ontology, prompt engineering, ROS Parser, ROS nodes, structured data, ROSGPT framework, ROS integration, elicitation ability, adaptability capabilities, zero-shot learning, few-shot learning.
The Name of the Current Solution (Technique/ Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution & What is the problem that need to be solved**	**What are the components of it?**
ROSGPT	The goal of the ROSGPT solution is to enhance human-robot interaction by leveraging large language models (LLMs) like ChatGPT and integrating them with the Robot Operating System (ROS). The specific problem that ROSGPT aims to solve is the challenge of seamless communication between humans and robots.	Large language models (LLMs) such as ChatGPT, GPT-3, and GPT-4, the Robot Operating System (ROS), the ROS Parser module, prompt engineering techniques, ontology-based approaches, the integration of ChatGPT with ROS.
The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process		

The paper introduces ROSGPT, which combines large language models with robotic systems for human-robot interaction. ROSGPT integrates ChatGPT with ROS2 to convert natural language instructions into structured robotic commands using zero-shot and few-shot learning.

	Process Steps	Advantage	Disadvantage (Limitation)
1	ROSGPT without ontology: In this step, the ChatGPT model is evaluated for generating accurate and relevant responses in human-robot interaction without the guidance of a predefined ontology. Various test scenarios and prompts are used to assess the accuracy, relevance, and comprehensibility of the generated JSON commands.	The model demonstrates the ability to make appropriate calculations and generate accurate JSON structures based on given information and can adapt to new tasks with only a few training examples, making human-robot interaction more flexible and adaptable.	Without the guidance of an ontology, the model may struggle to understand complex relationships between entities and concepts, leading to potential inaccuracies or misunderstandings in the generated commands.
2	ROSGPT using ontology: In this step, the ROSGPT system is integrated with an ontology-based approach to enhance comprehension and precision in generating structured commands for robotic systems. The ontology serves as a formal representation of knowledge in a specific domain, providing a consistent and unambiguous understanding of relationships between entities and concepts.	Improved comprehension: By leveraging structured knowledge from the ontology, the model can more effectively process natural language commands and translate them into accurate and executable actions reducing the chances of errors or misunderstandings for robots.	The effectiveness of the system heavily relies on the quality and completeness of the ontology. Inaccuracies or limitations in the ontology can impact the performance of the model.

Major Impact Factors in this Work

This paper underscores the impact of Large Language Models (LLMs) like ChatGPT. It explores the revolutionary role of self-attention and transformer models, highlighting LLMs' adaptability through zero-shot and few-shot learning.

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
Large Language Models (LLMs)	Training Data		
Robot Operating System (ROS)	Ontology Design		
Ontology	Prompting Strategies		
Prompt Engineering			

Relationship Among the Above 4 Variables in This article

Because of the variables like learning rate, batch size, loss function the mAP, object detection and localization, speed and efficiency of the model changes. And through these variables model's performance is measured.

Input and Output	Feature of This Solution	Contribution & The Value of This Work

Input	Output
Unstructured textual command (natural human speech)	JSON command that represents the structured action to be performed by the robot

ROSGPT innovatively integrates Large Language Models (LLMs) with the Robot Operating System (ROS), fostering intuitive human-robot interaction. Guided by ontology, it translates unstructured commands into task-specific instructions.

This paper evaluates ROSGPT in human-robot interaction, omitting a predefined ontology. Assessing ChatGPT's performance, it examines JSON command accuracy, relevance, and comprehensibility. This exploration informs the development of more efficient communication systems between humans and robots.

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
ROSGPT enhances human-robot interaction, boosts efficiency through LLMs, enables real-time command execution, and bridges ROS and ChatGPT for seamless communication.	ROSGPT, relying on LLMs, faces challenges in scalability, context understanding, prompt sensitivity, domain knowledge limitations, and ethical considerations. Ongoing refinement is crucial for responsible human-robot interactions.

Analyse This Work by Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
The paper critically analyses ROSGPT, focusing on ChatGPT's performance in human-robot interaction without ontology-based prompting. It commends ChatGPT's strong elicitation and adaptability, noting accurate JSON command generation but highlights limitations, including deviation from ontology affecting responses.	Evaluation of the ROSGPT solution can employ empirical tests, measuring JSON command accuracy, relevance, and response comprehensibility. User studies, surveys, and comparative analyses against existing systems offer insights into real-world usability and performance. An additional focus on ontology-based evaluation compares the impact of ontologies on ChatGPT's precision and comprehension.	1. Abstract 2. Introduction 3. Conceptual Architecture od ROSGPT 4. Proof of Concept 5. Conclusion 6. References 7. Appendices

Diagram/Flowchart

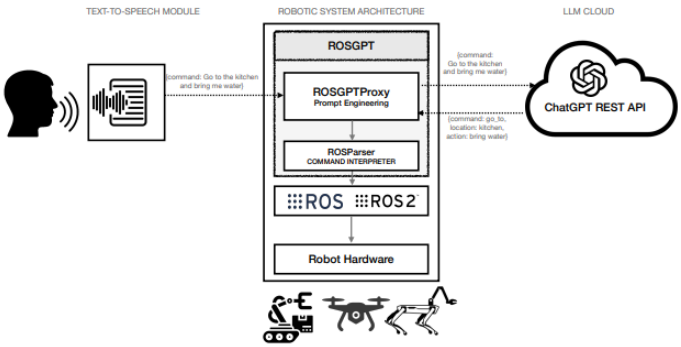


Fig. 1: ROSGPT Architecture for Human-Robot Interaction

13	High Accuracy Conversational AI Chatbot Using Deep Recurrent Neural Networks Based on BiLSTM Model	
Reference in APA format	P. Anki, A. Bustamam, H. S. Al-Ash and D. Sarwinda, "High Accuracy Conversational AI Chatbot Using Deep Recurrent Neural Networks Based on BiLSTM Model," 2020 3rd International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2020, pp. 382-387, 10.1109/ICOIACT50329.2020.9332074	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/document/9332074	• Mahardika Darmawan (darmawan.tif418@gmail.com) • Bustamamis (mcmutable@yahoo.co.id) • Heri Ahmadi (heri.ahmadi@tif.uad.ac.id)	• Chatbots • Python program • User input • Output data • Model design • Text and voice • Continuous responses
The Name of the Current Solution(Technique/Method/ Scheme/ Algorithm/ Model/Tool/Framework/ ... etc)	**The Goal (Objective) of this Solution & What is the problem that need to be solved**	**What are the components of it?**
BiLSTM Model	Aim is to create an accurate chatbot that can respond to user questions quickly based on a predefined dataset.	• Data collection • Identify data characteristics • Choose the parameters to be processed • Processing data

The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process

Selecting a model that matches the characteristics of the data is crucial for program performance.
The BILSTM model is chosen as the model to be applied to the program.

	Process Steps	Advantage	Disadvantage (Limitation)
1	Data Identity: •The data source is the Cornell Movie Dialog Corpus, which contains fictional conversations extracted from film scripts. •The dataset has a large number of conversation exchanges.	1.Integration of Multiple Models: The approach combines various elements such as BILSTM, Greedy decoding, and Seq2seq, allowing for a versatile and robust chatbot system.	1.Lack of Multilingual Support: If the model is not trained on multilingual data, it may struggle with handling queries in languages other than the one it was trained on.
2	2.Data Input: •Dialog sentences from the film are used as input for the program.	2.Large Dataset: The dataset mentioned contains a substantial number of conversation exchanges, which can contribute to the chatbot's ability to generate diverse and contextually relevant responses.	2.Limited Context Understanding: Despite the use of BILSTM, the chatbot may still struggle with long and complex conversations or nuanced language.
	3.Chatbot Program Development: •Consideration of various translation choices, such as sequence-to-sequence. •Selection of the BILSTM model to improve chatbot	3. Efficient Data Processing: The Greedy method chosen for data processing is known for its speed and simplicity. It is a practical choice for real-time or near-real-time applications, where rapid responses are	3.Data Quantity and Quality: The approach doesn't discuss data preprocessing or data cleaning. Poor data quality can lead to incorrect training, and a limited dataset might not capture the full diversity of

	response accuracy.	essential.	language and user queries.
	4.Output Evaluation: •The final step involves evaluating whether the model provides accurate results in comparison to the user's input dialog.		

Major Impact Factors in this Work

Major impact factors in the research on implementing chatbots using the BiLSTM model include Performance comparision, accuracy improvements and model interpretability.

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
Model architecture	Discrete levels	Quantitative	Knowledge gained from training program
Hyperparameter value	Continuum	Complexity of the user input	Overfitting

Relationship Among The Above 4 Variables in This arti

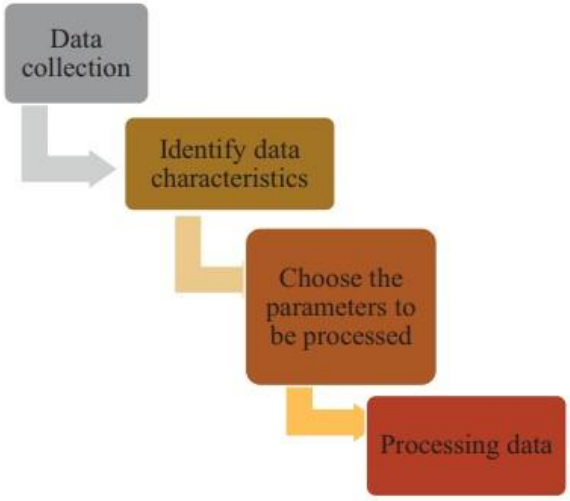
Optimizing this relationship is the primary objective based on the information provided. No clear moderato mediators are incorporated into this optimization experiment.

Input and Output	Feature of This Solution	Contribution & The Value of This Work
Input / **Output** Commands in the form of dialog sentences / Generated by the chatbot program	Utilizes a Bidirectional Long Short Memory model to capture context effectively and enhance response accuracy.Natural Language Understanding: The chatbot can understand and respond to user queries in natural language, making it user-friendly and accessible.	The use of the BILSTM model allows the chatbot to better understand and respond to user queries with enhanced context awareness. This contributes to more accurate and contextually relevant responses, improving the user experience. The commitment to output evaluation ensures that the chatbot provides accurate and relevant responses, enhancing the overall quality of user interactions.

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
Information retrieval, and user engagement, while offering adaptability. I enhance productivity, fosters user satisfaction, and addresses ethical considerations.	The solution includes biases in responses, user dissatisfaction with chatbot interactions, limitations in handling specialized queries.

Analyse This Work By Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
This work presents a methodical approach to developing a chatbot program using the BILSTM model drawing from a dataset. While the model selection and user feedback mechanisms show promise for enhancing user engagement and performance in conversational contexts. However, careful consideration of resource requirements, ethical implications, and adaptability to the project domain is essential for the work's successful integration and long-term value.	The assessment of this work likely employed a combination of essential tools, including Python for model development, deep learning frameworks like TensorFlow or PyTorch for neural network implementation, data analysis tools like Pandas, and version control systems like Git may have been used to measure and refine the chatbot's performance. The integration of these tools facilitated a comprehensive assessment of the chatbot's development and effectiveness in the project.	I. Abstract II. Introduction III. Materials and Methods IV. Data Implementation in the Python Program V. Applying Data Implementation in the Python Program VI. Program Making and Discussion of Test Results VII. Conclusion and Future Work VIII. Acknowledgements IX. References

Diagram/Flowchart



14	Jollity Chatbot- A contextual AI Assistant	
Reference in APA format	K. Deepika, V. Tilekya, J. Mamatha and T. Subetha, "Jollity Chatbot- A contextual AI Assistant," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 1196 1200, doi: 10.1109/ICSSIT48917.2020.9214076	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/document/9214076	• Kanakamedala Deepika 17wh1a1209@bvrithyderabad.edu.in • Subetha T subetha.t@bvrithyderabad.edu.in • Veeranki Tilekya 17wh1a1215@bvrithyderabad.edu.in	• Chatbot • Conversational agent • Contextual AI assistant • Intent recognition • Dialog management • Generative model • Rasa • Rasa NLU • Rasa Core

	• Jatroth Mamatha 17wh1a1208@bvrithyderabad.edu.in	
The Name of the Current Solution (Technique/Method/ Scheme/ Algorithm/ Model/ Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution & What is the problem that need to be solved**	**What are the components of it?**
Convolutional Neural Network (CNN), Hierarchical Attention Network (HAN), Recurrent Neural Network (RNN):	The problem being addressed by this solution is likely to improve the performance and robustness of chatbots in understanding user intents, responding appropriately, and maintaining coherent conversations.	• Intents • Responses • Training Data • Session Management • Session Configuration

The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process

The first step in developing these chatbots is understanding the specific user needs. This includes defining its purpose, interaction flow, and features.

	Process Steps	Advantage	Disadvantage (Limitation)
1	Define the Purpose and Goals: Clearly define the purpose and objectives of the chatbot.	Quick Response Times: Chatbots can provide instant responses, reducing user wait times and improving customer satisfaction.	Lack of Emotional Understanding: Chatbots, especially rule-based ones, often struggle to understand and respond to complex human emotions. They may provide scripted, impersonal responses in emotionally sensitive situations.
2	Collect and Prepare Data: If using machine learning, gather and prepare training data, which may include intent examples, entity recognition data, and dialogue samples.	Reduced Human Error: Chatbots operate with a high degree of accuracy, reducing the risk of human errors in tasks like data entry and information retrieval.	Limited Problem-Solving: Chatbots are typically designed for specific tasks and may not have the ability to solve complex or unique problems that require human creativity, critical thinking, or empathy.
3	Natural Language Understanding (NLU): Implement natural language understanding (NLU) components to analyze and interpret user messages, identifying intents and entities	Privacy and Security: Implement security measures to protect user data and ensure that the chatbot complies with privacy regulations.	Dependency on Training Data: Machine learning-based chatbots depend on high-quality training data. If the training data is limited or biased, the chatbot's performance may be suboptimal.

Major Impact Factors in this Work

The chatbot is specifically designed to be uplifting, motivating, and encouraging for users. This user-centric approach could lead to greater benefits. Using Rasa to create a generative chatbot that understands context should allow more natural, engaging conversations.

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
User mood/mental state	Use of the chatbot	Technology affinity	Increased social
Frequency of positive/negative	Chatbot interaction time	Duration of use	Interaction
Emotions	Chatbot response types	Chatbot personality	Change in thoughts

	Chatbot mood support feature		Increased activity

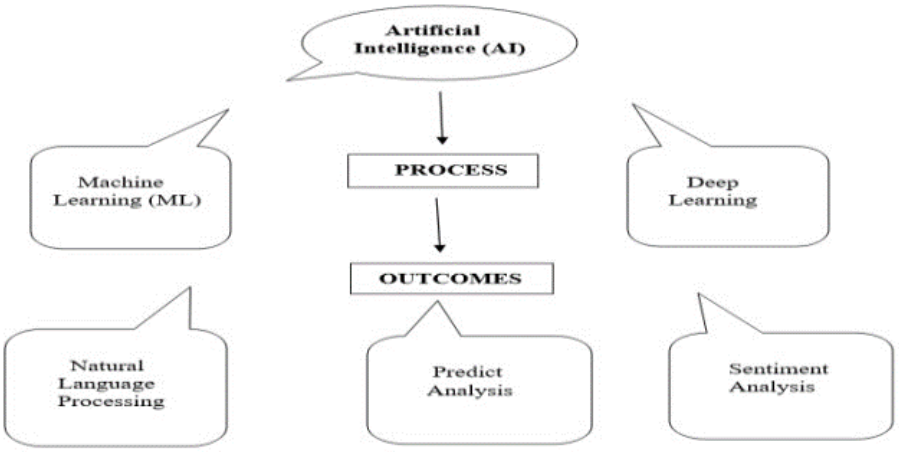
Relationship Among The Above 4 Variables in This article
Testing these variables and their interrelationships will provide greater understanding of how, for whom, and under what conditions the Jollity Chatbot achieves benefits as an accessible, low-cost mental health intervention.

Input and Output	Feature of This Solution	Contribution & The Value of This Work
Input / **Output** User queries and statements. / Responses and information.	AI and NLP Technology: The chatbots utilize AI and NLP technologies for understanding and responding to user input. Machine Learning and Classification: Machine learning classifiers are used to assess and support users' mental well-being. Continuous Improvement: The solution is committed to ongoing development and enhancement of chatbot capabilities.	The work outlined in the provided information makes significant contributions in improving accessibility and support through chatbots in major anticipating a future where these AI-driven systems play a central role in human interactions.

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
The system's dialog management and context awareness allow it to tailor responses to each user's needs and situation. The chatbot interaction provides anonymity.	The chatbot's conversations inherently lack the empathy, compassion, and emotional support provided by human interactions. the chatbot provides benefits in terms of accessibility, it cannot fully replace human counselors.

Analyse This Work By Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
Evaluating the system through intent recognition, story accuracy, and confusion matrices provides quantifiable performance data. Expand testing and enhancement of conversational capabilities through more comprehensive dialog training and Turing tests.	Risk analysis - Systematically identify risks like improper advice or privacy breaches. Expert-bot dialogs - Observe conversational sessions between the chatbot and domain experts. Confusion matrices - Quantify types of classification errors to identify weaknesses. Turing Tests - Measure how well humans judge natural language conversations with the bot.	I. Abstract II. Introduction III. Related Works IV. Proposed Jollity Chatbot V. Experimental Results VI. Conclusion VII. References

Diagram/Flowchart



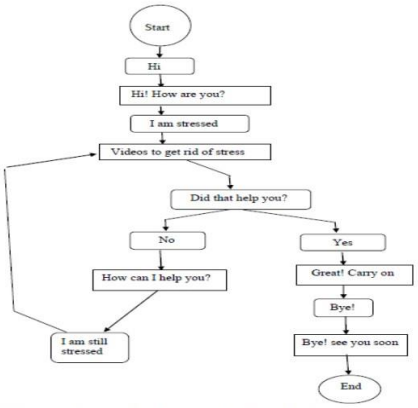


Figure 3: Sample flow of Jollity chatbot

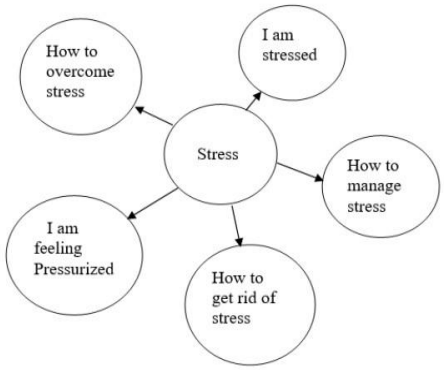


Figure 3: Stress intent

15	AI and Web-Based Human-Like Interactive University Chatbot (UNIBOT)	
Reference in APA format	N. P. Patel, D. R. Parikh, D. A. Patel and R. R. Patel, "AI and Web-Based Human-Like Interactive University Chatbot (UNIBOT)," 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2019, pp. 148-150, doi: 10.1109/ICECA.2019.8822176.	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/document/8822176	• Neel kumar P. Patel neelpatel710@gmail.com • Devangi R. Parikh dparikh74@gmail.com	• Artificial Intelligence • Chatbot • Human-like interactive • Machine Learning • University Chatbot
The Name of the Current Solution (Technique/Method/Scheme/Algorithm/Model/Tool/Framework/ ... etc)	**The Goal (Objective) of this Solution & What is the problem that need to be solved**	**What are the components of it?**
An AI and web-based human-like interactive university chatbot system called UNIBOT.	The goal is to chatbot processes user queries and retrieves the most relevant information from a database, eliminating the need for users to navigate through	• User Interface • Natural Language Processing • Database

	websites to find information.	• Security Features • Response Generation Module

The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process

The steps in this process include preprocessing, keyword matching, spell checking, option display, learning new questions, etc.

	Process Steps	Advantage	Disadvantage (Limitation)
1	User Input: Users input questions related to university information through the graphical user interface (GUI), similar to a messaging application.	Efficiency: The algorithm is designed to be efficient, requiring less memory and minimizing database hits.	Limited to Keywords: The system relies on keywords rather than natural language questions, potentially limiting the range of user inputs.
2	Message Processing: The algorithm checks and corrects spelling. Messages are split into words. SQL queries, using Regular Expression, check if words are present in the database.	User-Friendly: The GUI, provides a familiar and user-friendly experience.	Sensitivity to Spelling: The algorithm may be sensitive to spelling errors, requiring user correction.
3	Database Interaction: Relevant words are stored in an array called "important words". Another SQL query uses the array to retrieve information from the database.	Adaptability: The algorithm's adaptability allows it to handle various queries and contexts.	Dependency on Database: The system's effectiveness is contingent on the completeness and accuracy of the database
4	Response Generation: If a single result is obtained, display the answer. If multiple results, provide options based on the "title" column. If no result, check keywords in "Answer" columns: If a match, store for future reference. Otherwise, display a sorry message.		
5	User Interaction Loop: Users can choose options or enter another message, restarting the process.		

Major Impact Factors in this Work

The algorithm incorporating NLP techniques like preprocessing, spell check, keyword matching. Regularly enhancing the chatbot with new features, technologies like AI/ML, and user feedback will make it more capable over time.

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
User satisfaction	Preprocessing	User demographics	Chatbot response time

Response accuracy	Chatbot system	Question complexity	User perceptions
Perceived intelligence	Architecture and implementation	Natural language processing Capabilities	Response relevancy

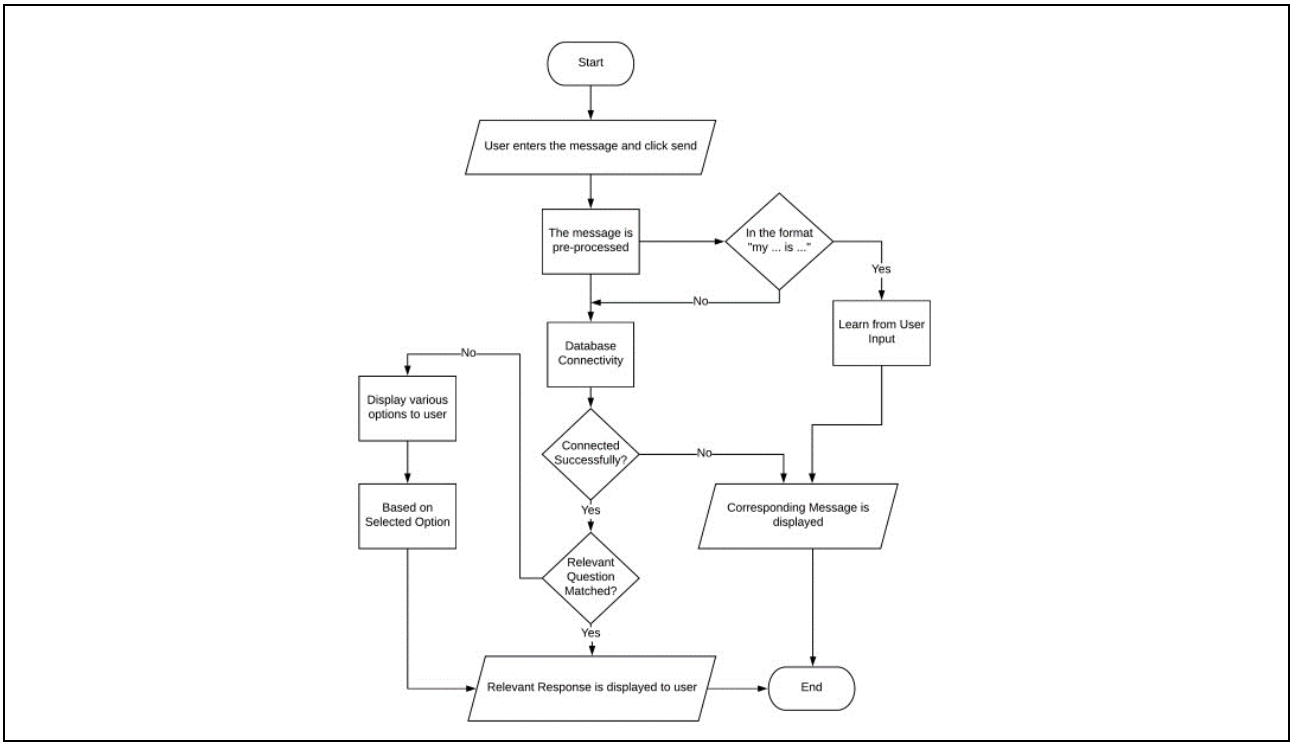
Relationship Among The Above 4 Variables in This artic
The chatbot system (IV) directly influenced the relevancy of responses (DV) based on its algorithm, interfa personalization capabilities.

Input and Output	Feature of This Solution	Contribution & The Value of This Work
Input: User submits a message in the text field. **Output**: The message is instantly displayed in the chat window.	Voice Input Support: Capability: Users can submit messages in voice format, enhancing accessibility and user convenience. Real-Time Display: Instant Messaging: Messages are instantly displayed in the chat window.	The inclusion of voice input support, optimized resource usage, and interactive response presentation further enhance the system's value.

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
Positively impacts the university information domain by providing a quick and user-friendly solution for accessing relevant details.	Potential negative impacts include nuanced query understanding, sensitivity to spelling errors, privacy concerns regarding user data Storage.

Analyse This Work By Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
This appears effective in streamlining university information retrieval with its user-friendly interface and dynamic algorithm. However, potential drawbacks include sensitivity to spelling errors, and privacy concerns. The system's impact depends on its ability to balance these aspects, providing effective responses while addressing user needs and concerns.	AI assessment involves testing frameworks, analytics tools, and possibly machine learning libraries. (e.g., TensorFlow, PyTorch), natural language processing libraries, and metrics for assessing the model's accuracy and performance. Specific AI tools depend on the algorithms and methodologies implemented in UNIBOT.	1. Abstract 2. Introduction 3. Proposed System 4. Design a. Database Design b. Algorithm Implemented 5. Sample Results 6. Conclusion 7. Future Scope 8. References

Diagram/Flowchart



16	AI- BASED CHATBOT WITH GPT-3	
Reference in APA format	Shaik, Amjan, S. Lohitha, and J. Ramesh Babu. "HEALTHCARE ASSISTANCE USING AI." *Journal of Engineering Sciences* 14.06 (2023).	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://jespublication.com/uploads/2023-14I8081.pdf	• M. Swathi • K. Gopal reddy	• Chatbot • OPENAI • Artificial Intelligence • Neural Networks • Machine Learning • GPT-3
The Name of the Current Solution (Technique/Method/Scheme/Algorithm/Model/Tool/Framework/... etc)	**The Goal (Objective) of this Solution & What is the problem that need to be solved**	**What are the components of it?**
GPT-3 (Generative Pre-trained Transformer 3): This is an autoregressive language model developed by OpenAI that uses deep learning to generate human-like text**.**	The goal of this project is to develop an AI-powered conversational agent that can effectively communicate with users. The problem it aims to solve is the difficulty of creating chatbots that understand natural language, engage in coherent dialogues, and provide satisfying responses. By leveraging the advanced natural language capabilities of GPT-3, this chatbot seeks to overcome limitations of previous chatbots and deliver more human-like, contextual conversations that meet users' needs.	• GPT-3 API • Training data • Bot framework • User Interface • Logging/tracking
The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process		
The chatbot is created by accessing the GPT-3 API, training the model on large datasets, fine-tuning on domain-specific data, integrating into a bot framework, creating a user interface, logging conversations for		

further training, evaluating performance metrics, hosting the bot, and maintaining/updating it over time.

	Process Steps	Advantage	Disadvantage (Limitation)
1	GPT-3 API - This provides access to the GPT-3 language model to generate responses.	Access GPT-3 API: Advanced AI capabilities	Access GPT-3 API: Can be costly
2	Training data - A large dataset of text conversations is used to pretrain GPT-3 to understand natural language.	Train on large dataset: Improves language understanding	Train on large dataset: Potential for data biases
3	Fine-tuning data - Domain-specific data to adapt the model to your chatbot's purpose.	Fine-tune on domain data: Increases response accuracy	Fine-tune on domain data: Limits general knowledge
4	Bot framework - The underlying framework and code that integrates GPT-3 into the chatbot system.	Integrate into bot framework: Enables natural conversations	Integrate into bot framework: Complex implementation
5	User interface - The front-end interface where users interact with the chatbot.	Create user interface: Makes chatbot usable	Create user interface: Design impacts experience
6	Logging/tracking - Tools to record conversations for further training and analysis.	Log conversations: Allows further training	Log conversations: Requires data processing
7	Evaluation metrics - Ways to test the chatbot's performance in areas like response quality, accuracy, and user satisfaction.	Evaluate metrics: Measures improvements	Evaluate metrics: Hard to evaluate qualitatively
8	Maintenance/updating - Ongoing work to improve the chatbot with new data.	Maintain/update: Improves chatbot over time	Maintain/update: Labor intensive

Impact Factors in this Work

The use of machine learning chatbots, particularly the ChatGPT model, which has revolutionized natural language processing and human-AI interaction.

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
Response quality	GPT-3 model architecture	User expectations	Conversational context modelling.
Accuracy	Evaluation metrics	Individual differences	Conversational strategy
	User profiles		

Relationship Among the Above 4 Variables in This article

The chatbot's performance on metrics like coherence, accuracy, and user satisfaction, also change the nature of the Relationship.

Input and Output	Feature of This Solution	Contribution & The Value of This Work

Input	Output
User queries	Relevant chatbot replies

Uses advanced AI (GPT-3) to understand language and generate human-like responses.
Trained on massive amounts of conversation data to chat naturally on diverse topics.

- Creates a conversational agent that can engage users and answer questions more effectively.
-Advances chatbot technology by leveraging the latest AI language model (GPT-3).

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
Personalization - GPT-3's ability to learn and adapt over time could allow for more personalized and tailored conversations catered to individual users. Faster service - Chatbots can provide instantaneous responses at all hours, improving resolution time compared to waiting for human agents.	Job losses - Automating customer service with chatbots could lead to fewer employment opportunities for human agents

Analyse This Work by Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
- Leverages powerful AI to improve chatbot conversational ability. -Training on large datasets enables more natural conversations. -However, potential for inheriting biases from training data.	Combining automated testing tools to gather metrics with human-in-the-loop techniques like user studies and transcript reviews is important for a comprehensive evaluation.	1. Abstract 2. Introduction 3. Related works 4. Proposed System 5. Architecture 6. Discussion 7. Output 8. Future Scope and Conclusion 9. References

Diagram/Flowchart

17	Chat GPT & Google Bard AI: A Review	
Reference in APA format	S. K. Singh, S. Kumar and P. S. Mehra, "Chat GPT & Google Bard AI: A Review," 2023 International Conference on IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2023, pp. 1-6, doi:10.1109/ICICAT57735.2023.10263706	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
https://ieeexplore.ieee.org/abstract/document/10263706	• Shashi Kant Singh shashikantsingh_2k19co356@dtu.ac.in • Shubham Kumar shubhamkumar_2k19co371@dtu.ac.in • Pawan Singh Mehra pawansinghmehra_2k19co357@dtu.ac.in	• Chat GPT • Google Bard AI • Artificial Intelligence • Chatbots • NLP • Neural Network
The Name of the Current	**The Goal (Objective) of this Solution & What is the problem that need to be**	**What are the components of it?**

Solution(Technique/Method/Scheme/Algorithm/ Model/Tool/ Framework/etc)	solved	
Chat GPT and Google Bard AI	The goal of the solutions, Chat GPT and Google Bard AI, is to provide advanced chatbot capabilities using Artificial Intelligence and Machine Learning.	Natural Language Processing (NLP) Sentiment Analysis Data storage for information retrieval Integration with other technologies

The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process

The process involves defining the aim, creating, testing, publishing, and evaluating the chatbot, aiming to provide tailored, accurate responses while requiring technical expertise and ongoing refinement.

	Process Steps	Advantage	Disadvantage (Limitation)
1	Defining the Aim: Identifying the purpose and domain of the chatbot, as well as the user requirements.	Enhanced cybersecurity: Chat GPT can assist in identifying and mitigating cybersecurity threats.	Security concerns: Chat GPT handles a large amount of data, which raises privacy and security concerns.
2	Creation: Developing the chatbot using advanced AI and NLP techniques.	Automated incident response: Chat GPT can assist in automating incident response processes.	Potential for misuse: Due to its vast knowledge in various fields, Chat GPT can be misused by terrorist organizations, criminal groups, and hostile countries.
3	Testing: Evaluating the chatbot's performance and identifying errors.	Improved user awareness: By answering user queries and providing explanations in a user-friendly manner.	Effect on human creativity: Chat GPT's ability to Enhance artificial creativity may have a negative impact on human creativity.
4	Publishing and Integration: Making the chatbot available online and integrating it into applications.	Continuous learning and adaptation: Chat GPT can continuously learn from new cybersecurity trends, emerging threats, and evolving attack techniques.	Impact on students: While Chat GPT can be helpful for students in many ways, it can also hinder their academic growth.
5	Evaluation: Continuously assessing the chatbot's performance and identifying areas for improvement.		

Major Impact Factors in this Work

Major impact factors include the analysis of Chat GPT and Google Bard AI, the exploration of applications in various fields, acknowledgment of limitations and biases.

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
Response quality	GPT-3 model architecture	User expectations	Conversational context modelling.
Accuracy	Evaluation metrics	Individual differences	Conversational strategy
	User profiles		

Relationship Among the Above 4 Variables in This article
The chatbot's performance on metrics like coherence, accuracy, and user satisfaction, also change the nature of the Relationship.

Input and Output	Feature of This Solution	Contribution & The Value of This Work
Input: User queries; **Output**: Relevant chatbot replies	Ability to show a combination of creativity, power, and intelligence. It also has a feedback system where it can take feedback from users and improve accordingly.	It explores the contribution of Chat GPT in different fields such as questioning-answering, coding and debugging, writing, math tutoring, language translation, text classification, and even in geophysics problem-solving.

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
Chat GPT can assist with various types of questions asked by users. Chat GPT can also be used to create new Chatbots that can do conversations with us.	Chat GPT may not always provide accurate solutions.

Analyse This Work by Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
Evaluates the information presented, questioning assumptions, examining evidence, considering alternative perspectives. Evaluate the methodology and examine the credibility of the sources.	The tools used to assess this work include survey analysis, specific works such as "Google Bard and Open ai's chat gpt Head-to-Head and "Chat GPT vs Google Bard: Generative AI comparison" for comparison and analysis.	1. Abstract 2. Introduction 3. Literature Review 4. Introduction of Chatbots 5. ChatGPT 6. Google Bard AI 7. A Brief Comparison between ChatGPT and Google Bard AI 8. Conclusion and Future Scope 9. References

Diagram/Flowchart

18	Voice Assistant Integrated with Chat GPT	
Reference in APA format	Shafeeg, A., Shazhaev, I., Mihaylov, D., Tularov, A., & Shazhaev, I. (2023). Voice assistant integrated with chat gpt. Indonesian Journal of Computer Science, 12(1).	
URL of the Reference	**Authors Names and Emails**	**Keywords in this Reference**
http://3.8.6.95/ijcs/index.php/ijcs/article/view/3146	• Ilman Shazhaev - ilman@farcana.com • Arbi Tularov - arbi@farcana.com • Dmitry Mikhaylov - dr.d.mikhaylov@nus.edu.sg • Islam Shazhaev - islam@farcana.com • Abdulla Shafeeg - abdulla.shafeeg@farcana.com	• Artificial Intelligence • Chat GPT • Digital community • Voice assistant
The Name of the Current Solution(Technique/Method/Scheme/Algorithm/Model /Tool/ Framework/ ... etc)	**The Goal (Objective) of this Solution & What is the problem that need to be solved**	**What are the components of it?**
ChatGPT, which is a combination of the GPT language model and a voice assistant.	The goal of the ChatGPT solution is to provide advanced AI-powered virtual assistance for various tasks.	• Neural Networks • OpenAI's GPT Language Model • Memory • Filters

The Process (Mechanism) of this Work; Means How the Problem has Solved & Advantage & Disadvantage of Each Step in This Process

The process or mechanism used in this work is the implementation of an AI chatbot called ChatGPT. It is trained on a large dataset of text from various sources to generate human-like responses.

	Process Steps	Advantage	Disadvantage (Limitation)
1	Machine Learning: ChatGPT utilizes machine learning algorithms to learn from large amounts of data.	High-quality responses: ChatGPT provides responses that are often considered to be of higher quality than those obtained from traditional search engines like Google.	Code quality: ChatGPT has been found to generate poor quality code.
2	NLP Annotations: Natural Language Processing (NLP) annotations and data tagging are used to make conversations easier for the chatbot to understand.	Code assistance: ChatGPT can help with coding tasks by providing guidance.	Lack of contextual understanding: While ChatGPT can remember details of the conversation, it may still struggle with understanding the context accurately.
3	Voice Assistant Integration: ChatGPT can be integrated with voice assistants to provide a unique gaming experience.	Language generation: ChatGPT can generate various types of written content.	Wordiness: ChatGPT tends to use excessive phrases and repetitions, resulting in wordy responses that may not be concise or efficient. Struggle with understanding the context accurately.

4	Client Authentication: The chatbot can handle the process of client authentication by asking a series of questions to verify the customer's identity.	Context understanding: ChatGPT, particularly the GPT-3.5 version, is designed to handle context better.	
5	Request Classification: The chatbot uses natural language processing techniques to categorize and understand the requests made by customers.	Time-saving and efficiency: ChatGPT can save time for employees by providing comprehensive and up-to-date information.	
6	Continuous Improvement: The chatbot can continuously learn and improve its responses by analysing past interactions and incorporating user feedback.	Sentiment analysis: ChatGPT can analyse the sentiment of the user based on response length and word usage.	

Major Impact Factors in this Work

Major impact factors in this work include AI which helps in the development and integration of AI technologies, Automation and Efficiency and Productivity.

Dependent Variable	Independent Variable	Moderating variable	Mediating (Intervening) variable
User engagement	ChatGPT integration	Trust in AI	User trust
Player performance	Knowledge gain	Competitiveness	Knowledge gain
User experience	Training data		Time savings
Cost savings	Game attributes		

Relationship Among the Above 4 Variables in This article

This relationship may be moderated by factors like the player's trust in AI and need for cognition. This relationship could be moderated by game genre complexity. The impact of ChatGPT integration on user engagement may be mediated through knowledge gain from the assistant's explanations.

Input and Output		Feature of This Solution	Contribution & The Value of This Work
Input	**Output**	The Chat GPT model used in the voice assistant is capable of providing high-quality responses. The digital voice assistant uses natural language processing techniques to categorize and understand the requests of users.	The integration of a voice assistant with Chat GPT in this work contributes to the development of AI in the digital society, by combining the functionality of a chatbot and a voice assistant. The value of this work lies in providing users with a more immersive and interactive experience.
ChatGPT integration, Training data, Conversational context, User attributes, Game attributes, Restrictiveness of filters.	User engagement, Player performance, Query resolution rate, Account security, Cost savings, Customer support load, User		

	experience.		

Positive Impact of this Solution in This Project Domain	Negative Impact of this Solution in This Project Domain
Quick verbal account security and authentication could reduce account theft and fraud, enhancing trust and safety for users.	Substantial development, testing, maintenance and continuous training costs may be required to make the AI assistant perform well across contexts.

Analyse This Work by Critical Thinking	The Tools That Assessed this Work	What is the Structure of this Paper
The authors provide a good overview of ChatGPT's capabilities and limitations, citing recent research on its language generation abilities. This provides helpful context.	Tools like Turnitin or PlagScan could check for improper reuse of content from other sources without attribution.	X. Abstract XI. Introduction XII. Research Method XIII. Literature Review XIV. Case Development XV. Conclusion XVI. References

Diagram/Flowchart

2.2 COMPARISION TABLE:

Author	Year	Approach	Description
Abhishek Sarda, Dr. Shubhra Dixit, Dr. Anupama Bhan	2021	Training, Data Collection, Non-Max Suppression, Calculation, Augmentation, Techniques	Training the YOLO algorithm on custom classes using the open images dataset, implementing non-max suppression, and applying specific techniques for object detection in the context of autonomous driving
M. Z. Khan, Shahid Mumtaz, Muhammad Khurram Khan	2019	Using Machine learning, edge detection techniques, and CNN, the faces are recognized	The suggested solution uses CNN, a cutting-deep learning technique in computer vision. Even with several faces in the shot, the system can still identify persons. Due to the system's lack of sensitivity to light, it can also recognize individuals in a variety of situations and lighting scenarios.
Md. Mahmud Hasan, Lim	1998	tele-security system, hardware design,	The development and deployment of a smart telephone device equipped with Caller ID, an

Hooi Jiun, Ng Wei Chuen		Signaling System, Caller ID system	answering machine, and a tele-security system. It details the hardware design, including using a photo coupler for ring signal detection and a multifunctional IC for FSK signal reception and decoding. It elucidates the transmission formats, operational aspects of the smart telephone device and indoor sensors, and the Caller ID system within digital telephone networks, supported by Signaling System No. 7 (SS 7).
T. Feng, Y. Yao, L. Wu, Y. Bai, Z. Xiao, Z. Lu	2018	Ultrawideband (UWB) Technology, hyperbolic algorithm, virtual spring model in the robot to track target person.	The process for tracking the target person is based on Ultrawideband (UWB) technology, a modified hyperbolic positioning algorithm, and a virtual spring model implemented in the robot to track the person. The use of a moving-average filter to reduce the noise in the UWB data. The potential use of saliency methods for object identity agnostic models.
Menaka R Archana N Dhana gopal Ramesh R	2020	YOLO algorithm, Intersection over Union (IoU), Non-max suppression, anchor boxes	It uses the convolutional layers to reduce the feature space and uses frame analysis and anchor boxes for object detection. The data augmentation and dropout are used to avoid overfitting. The YOLO algorithm is pre-trained using the related layer and then fine-tuned for object detection.
Muhammad Iqbal, Bhakti YudhoSuprapto, Suci Dwijayanti	2022	Convolutional neural network (CNN), architectures like Alex Net VGG16, accuracy measurements	Real-time face emotion recognition in a humanoid robot is achieved by the use of convolutional neural network designs such as VGG16 and Alex Net, in this technique. The performance is assessed based on accuracy in real-time interactions, and the system is built to detect faces and emotions simultaneously.
Reagan L. Galvez, Elmer P. Dadios, Argel A. Bandala, Ryan Rhay P. Vicerra	2019	Transfer learning involves using pre-trained models and fine-tuning them on a specific dataset, training from scratch with pre-existing weights, mean average precision(mAP) used for specifications, IED X-ray dataset for X-rays images	The authors show that training YOLO from scratch works better than transfer learning in the rapid identification of threat objects by comparing the performance of the two methods using an IED X-ray dataset. The work also covers training and evaluation using the IED X-ray dataset and the Darknet-53 architecture.
Y. Ye, Y. Zhang, Y. Li, Y. Liu	2023	ChatGPT to enhance communication, designing RoboGPT, NLP, accuracy or performance scores and cognitive load measures	The method used in this involves designing the RoboGPT system, which utilized ChatGPT to enable natural language communication between humans and robots. A human-subject experiment was conducted to validate the effectiveness of solution. The performing a simple assembly task with the robot arm under conditions and was assessed using performance scores, self-evaluated

			performance, and cognitive load measures.
Mazin Hnewa, Hayder Radha	2021	development of a Multi Scale Domain Adaptive YOLO, Domain Adaptive Network (DAN) with multiscale feature inputs, performance of YOLOv4 for cross-domain object detection by producing robust domain-invariant features that reduce the impact of domain shift.	The method outlined in this study entails creating a Multi-scale Domain adaptive YOLO framework, which combines the YOLOv4 object detector within a domain adaptive network (DAN). When learning domain invariant features from the multiscale feature inputs, the DAN is connected to YOLOv4 architecture. To predict the domain class, the DAN combines many domain classifiers with neural layers. This method seeks to achieve domain invariance during domain adaption-based training and resealing features against domain shifts at various sizes.
W.S. Mada Sanjaya, Dyah Anggraeni, Kiki Zakaria, Atip Juwardi, MadinatulMunawwaroh V. Amala Rani and S. Lalitha Kumari	2017	Development of a Social Robot named SyPEHUL, Cascade Classification and LBPH (Local Binary Pattern Histogram), Real-time face detection and recognition, with the system being capable of mentioning the person's name upon identification	The creation of a social robot called SyPEHUL, combines cutting-edge image processing techniques to follow and identify human faces. For precise face tracking and recognition, the system makes use of LBPH and cascade classification techniques which are based on the OpenCV library and python2.7. The hardware configuration comprises 12 Degree of Freedom(DoF) motor servos to move the robot's head and face. It is based on an Arduino microcontroller. A database of photos of faces is used to train the system, and its ability to accurately track and recognize faces is evaluated. It is possible to do real-time face detection and recognition, and the system can mention an individual's name after successful identification.
Prasnurzaki Anki, Alhadi Bustamam, Herley Shaori Al-Ash, DevviSarwinda	2020	Chatbot, program, BiLSTM, accuracy, input, output	The research provides a practical demonstration of how chatbots can be implemented in Python using deep learning models like BiLSTM. This can help developers and researchers build real-world chatbot applications. The BiLSTM model architecture provides more interpretability compared to “black box” models like neural networks. The use of the Cornell Movie Dialog corpus provides a diverse, real-world dataset for training the models. This enhances applicability.
Kanakamedala Deepika, VeerankiTilekya, Jatroth Mamatha,	2020	Chatbot, Contextual AI Assistant, Rasa, Rasa NLU, Rasa Core	The chatbot is designed to converse with users, suggest uplifting content like positive quotes, motivational videos, etc. based on the user’s mood, and act as a virtual friend. It uses Rasa, an open-source conversational AI

Subetha T			framework, with custom intents, sample conversations, and responses defined to understand user needs and maintain context.
Neelkumar P. Patel, Devangi R. Parikh, Darshan A. Patel, Ronak R. Patel	2019	Artificial Intelligence, Chatbot, Human-like interactive, Machine Learning, University Chatbot	The overall chatbot system architecture and interface for interacting with users. The custom algorithm was developed for processing user queries and retrieving responses from the database. This algorithm involves steps like preprocessing, keyword matching, spell checking, option display, learning new Q&A pairs, etc. Expanding the chatbot's knowledge base, deploying it on servers to handle multiple users, and maintaining performance is key for broader implementation.
M. Swathi, K.Gopal Reddy	2023	Chatbot, OPENAI, Artificial Intelligence, Agents, GPT-3	The chatbot, known as ChatGPT, is based on a language model that excels in natural language processing tasks is capable of generating human-like text responses. The model underwent extensive training on a large corpus of text data, enabling it to understand language nuances, generate accurate responses, and handle multiple inputs and languages. The destiny scope of chatbots includes interacting with multiple users and expanding inside the numerous applications inclusive of studies, schooling, and creative writing.
Shashi Kant Singh, Shubham Kumar, Pawan Singh Mehra	2023	Artificial Intelligence, Chatbots, Chat GPT, Google Bard AI	It especially specializes in the evaluation of Chat GPT, it is a first-rate innovation within the location of Artificial Intelligence. It recognizes the restrictions of language fashions like GPT-2 and GPT-3, collectively with biases and restrained accuracy. It highlights the want for interactive language fashions like Chat GPT to cope with those limitations and decorate easy normal performance. Enables the chatbot to recognize and reply to complicated queries. It permits making the chatbot available online and integrating it into programs.
IlmanShazhaev, Arbi Tularov, Dmitry Mikhaylov, Islam Shazhaev, Abdulla Shafeeg	2023	Artificial Intelligence, ChatGPT, digital community, voice assistant	Neural Networks allow the chatbot and voice assistant to perform tasks that include answering questions, writing code, and composing essays. OpenAI's GPT language the model has been trained to generate meaningful responses to questions and can perform various tasks, including finding bugs in code and composing poetry. The system has filters to keep away from debatable subjects and save you from growing texts

			about illegal or immoral sports. Farcana has integrated the functionality of the GPT chatbot and voice assistant to enhance player experience, familiarize players with game mechanics, and manage gaming accounts.
Shaji Thorn Blue, M. Brindha	2019	Boundary Box prediction, Edge Detection, Object Detection, YOLO, YOLO9000, YOLOv3	The proposed algorithm demonstrates improved precision in drawing boundary boxes compared to YOLOv3. Results show the proposed method outperforms YOLOv3 in terms of accuracy, as indicated by the Intersection over Union (IOU) comparison. The proposed work exhibits potential for object detection with precise boundary boxes, although limitations exist, particularly in scenarios involving sharp objects and high image noise.
Shraddha Mane, Supriya Mangale	2018	CNN, Object detection, Tensorflow, Tracking	The model uses Tensorflow object detection API which is robust in detecting objects in complex scenes and background conditions. Detects object location (x,y,w,h) which is then passed to the tracking algorithm. Uses a CNN-based tracking algorithm instead of traditional computer vision techniques. Loads weights of a pre-trained model capable of incorporating temporal information. Able to track objects at 150 FPS and handle occlusion scenarios. The main innovation is in using deep learning models to improve generalization across challenging uncontrolled video settings.
Anis Koubaa	2023	Human-Robot Interactions, ROS, ROS2, ChatGPT, Large Language Model, Robot Operating System	The system utilizes prompt engineering to convert unstructured human language commands into structured robotic instructions, enabling seamless communication between humans and robots. The implementation includes a proof-of-concept for robot navigation and demonstrates the use of an ontology to guide ChatGPT in generating context-specific robotic commands. It also emphasizes the importance of prompt engineering and ontology development, as well as the limitations and potential biases of using LLMs for human-robot interaction.
Xiao-Feng Zhao, Zi-Heng Chen, He-Feng Yin, Xiao-Jun Wu	2023	Artificial Intelligence cloud, edge computing, face recognition, speech recognition	The system addresses the management requirements for visitors in enterprises and government departments, aiming to improve efficiency and accuracy while reducing labor costs. By utilizing AI cloud and edge computing, the system incorporates face recognition and speech recognition technologies to enable functions such as

			visitor verification, and voice self-registration. The system is composed of a cloud service layer and an edge computing layer, with hardware components including RK3399 motherboards, microphone array, camera, and IR temperature sensor. The system's cloud-edge collaboration enhances computing capacity, improves real-time performance, and ensures high accuracy. The testing results showed high recognition accuracy and short recognition delays.

2.3 WORK EVALUATION TABLE:

	Work Goal	System' s Components	System' s Mechanism	Features/Characteristics	Performance	Advantages	Limitations/Disadvantages	Results
Pranurzaki Anki, Alhadi Bustamam, Herley Shaori Al-Ash, Devvi Sarwinda 2020	The main goal of this work is to create an accurate conversational AI chatbot that can respond quickly to user questions based on a predefined dataset. The specific objective is to develop a chatbot program using the Bidirectional Long Short-Term	The BiLSTM (Bidirectional Long Short-Term Memory) model is chosen as the core model for the chatbot program.The approach combines elements like BiLSTM, Greedy decoding, and sequence-to-sequence (seq2seq) models.	Selecting a suitable dataset (Cornell Movie Dialog Corpus) that provides a large number of conversational exchanges. Using the dialog sentences from the dataset as input for the chatbot program. Implementing the BiLSTM model, along with other techniques like Greedy decoding and seq2seq	The input is in the form of natural language text. The chatbot generates responses in the form of natural language text. The responses are intended to be contextually relevant and accurate based on the user's input dialog.	The use of the BiLSTM model is expected to improve the chatbot's response accuracy by effectively capturing context from the conversational data.The integration of multiple models (BiLSTM, Greedy decoding, seq2seq) is stated to contribute to a "versatile and robust chatbot system," suggesting improved	The approach combines various elements such as BiLSTM, Greedy decoding, and seq2seq, allowing for a versatile and robust chatbot system.The Greedy method chosen for data processing is known for its speed and simplicity, making it suitable for real-time or near-real-time	If the model is not trained on multilingual data, it may struggle with handling queries in languages other than the one it was trained on.Despite the use of BiLSTM, the chatbot may still struggle with long and complex conversations or nuanced language.	The proposed approach should lead to improved response accuracy and context understanding, but it does not present numerical results or comparisons with other models or baselines.

	Memory (BiLSTM) model, which aims to improve the accuracy of the chatbot's responses by effectively capturing the context of conversations.		models, to develop the chatbot program.Evaluating the output responses generated by the chatbot against the user's input dialog to assess the accuracy and performance of the system.		overall performance. The choice of the Greedy method for data processing is mentioned as an advantage due to its speed and simplicity, making it suitable for real-time or near-real-time applications where rapid responses are essential.	applications where rapid responses are essential.		
Kanakamedala Deepika, Veeranki Tilekya, Jatroth Mamatha, Subetha T 2020	The main goal of this work is to design a jollity chatbot that can converse with human users, entertain them, provide suggestions, and offer motivation during tough times.	The chatbot is implemented using Rasa, an open-source conversational AI framework. It utilizes natural language understanding (NLU) for intent recognition and entity extraction. The system incorporates dialogue management capabilities using Rasa Core.	Rasa NLU processes the user input to identify the intent and extract relevant entities. Based on the identified intent and the conversational context (defined in stories.md), Rasa Core determines the appropriate response or action.	The system can handle spelling mistakes and variations in user input through its natural languageunderstanding capabilities. The chatbot is built using the Rasa framework, which includes components for natural language understanding (NLU) and dialogue management (Core). Rasa Core handles dialogue management and	The accuracy of the system in correctly identifying the user's intents. The accuracy of the system in correctly identifying the flow of interactions.	The use of Rasa, an open-source conversational AI framework, allows for customization and integration of different components. The incorporation of NLU capabilities enables the chatbot to understand user intents and extract relevant entities, improving the accuracy of responses.	The chatbot's primary focus is on providing motivation and suggestions, which may limit its applicability in other domains or scenarios.	Reports an accuracy of 90% in recognizing the user's intents correctly. The system is reported to perform well in identifying the correct flow of interactions (stories), although specific metrics are not provided.

				determines the appropriate responses based on the identified intents and conversational context.				
Neelkumar P. Patel, Devangi R. Parikh, Prof. Darshan A. Patel, Prof. Ronak R. Patel 2019	The goal of this work is to develop a chatbot named "UNIBOT" (University Chatbot) that can interact with users and provide various information. The chatbot aims to overcome the tedious and time-consuming process of visiting universities or colleges to gather information.	The chatbot has a graphical user interface (GUI) similar to a messaging application, developed using HTML, CSS, jQuery, and Ajax.The chatbot uses PHP as the back-end language for processing user queries and generating responses.	The user's input message is preprocessed, including spell checking, word splitting, and keyword extraction using regular expressions. An SQL query is executed to search for relevant information in the database based on the extracted keywords.	The chatbot checks for misspelled words and provides suggestions to the user. If multiple relevant answers are found, the chatbot presents options to the user.Responses are generated dynamically based on the user's input and the database content.	Provides efficient and relevant responses to the user. Has minimal response time. Requires less memory and fewer database hits. Most of the time, the provided replies satisfy the user's requirements.	The GUI is designed to be similar to messaging applications, which users are already familiar with, providing a friendly environment for interaction. Users can enter questions or queries in natural language, making the interaction more intuitive and human-like.	Does not provide specific quantitative metrics or comparisons to evaluate the chatbot's performance objectively.	The chatbot provides efficient and relevant responses to the user in most cases. The response time is minimal. The memory requirements and database hits are less compared to other approaches. The majority of the provided replies satisfy the user's requirements.
Ms. M. SWATHI, K.GOPAL REDDY 2023	The goal is to design a chatbot using OpenAI's GPT-3 (Generative Pre-trained Transformer 3) model,	GPT-3 language model from OpenAI, Programming language (e.g., Python), Libraries and APIs for	Create an account on OpenAI and obtain API keys. Choose a programming language and set up the development	Natural language input and output Context-aware response generation. Ability to handle complex and nuanced	GPT-3 can understand and respond to complex and nuanced user inputs, allowing for more natural and effective communication compared to simpler chatbot	Improved communication and user experience Enhanced support for customer service agents. Increased customer	Potential biases in the training data. Limited understanding of context and real-world knowledge. Difficulty with sarcasm, irony, and non-	GPT-3 was trained on a massive corpus of text data, which can potentially improve its ability to generate high-quality

	which can generate human-like text responses.	integrating with GPT-3	environment. Install required libraries and dependencies. Build a function to call the GPT-3 API and generate responses based on user input. Integrate the GPT-3 response generation into a chatbot interface.	queries. Multi-lingual support. Fine-tuning on specific domains for specialized responses.	models.	satisfaction. Ability to handle complex queries and generate coherent responses. Multilingual support.	standard language. Inability to generate truly creative or abstract responses.Limited reasoning and logical capabilities.	responses across a wide range of topics and domains.
Shashi Kant Singh, Shubham Kumar, Pawan Singh Mehra 2023	Analyze and review Chat GPT and Google Bard AI, two prominent language models/chatbots developed by OpenAI and Google respectively.	Chat GPT is based on the GPT-3 language model architecture developed by OpenAI. Google Bard AI is powered by LaMDA (Language Model for Dialogue Applications).	Chat GPT uses natural language processing to generate human-like text responses based on its training data. Google Bard AI is designed for conversational interactions, reacting quickly to its environment.	Chat GPT has 175 billion parameters and a vast vocabulary, allowing it to operate in many contexts. Google Bard AI has 1.37 billion parameters, 1.5 trillion-word vocabulary, and refrains from making presumptuous judgments.	Chat GPT can solve math problems, write code, generate essays/articles, and assist in various tasks across domains. Google Bard AI aims to provide more detailed and easily understandable responses powered by the latest data.	Aid in education, creative writing, analysis, coding, and knowledge sharing across fields. Google Bard AI overcomes some biases and inaccuracies present in Chat GPT.	Potential biases, inaccuracies, limited common sense, and security/privacy concerns. Chat GPT's data is limited until 2021; Google Bard AI has limited access currently.	Both systems have wide applications but require improvements in accuracy, bias reduction, and expanded knowledge bases. Ongoing competition to develop more advanced conversational AI assistants.

Ilman Shazhaev, Arbi Tularov, Dmitry Mikhaylov, Islam Shazhaev, Abdulla Shafeeg 2023	Combine the functionality of the GPT chatbot and a voice assistant to provide a unique gaming experience and helps new players learn game mechanics in Farcana's new play-to-hash shooter game.	Chat GPT language model based on GPT-3.5 architecture developed by OpenAI. Voice assistant algorithms for speech recognition and natural language processing.	Chat GPT generates human-like text responses based on its training data.Voice assistant converts speech to text, which is fed into Chat GPT to generate relevant responses.Responses from Chat GPT are converted back to speech via text-to-speech synthesis.	Remembers context and can build on previous conversation.Avoids controversial topics and has filters for inappropriate content.Can handle a variety of tasks like coding, writing, analysis, etc.	Assists new players in understanding game mechanics, strategies, settings, and account management. Authenticates users, classifies requests, manages accounts, blocks accounts if needed.Conducts outgoing calls to inform gaming community about new services.	Provides personalized assistance and immersive experience for players. Automates routine tasks, allowing human staff to focus on complex issues. Leverages latest advancements in AI, NLP, and machine learning.	Can generate nonsensical or biased responses at times.Limited knowledge base focused on pre-2021 information.Potential for misuse or spreading misinformation if not carefully controlled.	Innovative solution combining voice AI and Chat GPT capabilities.Aims to enhance gaming experience, security, and community engagement.Ongoing development to improve accuracy, reduce biases, and expand capabilities.
Md. Mahmud Hasan, Lim Hooi Jiun, Ng Wei Chuen, Md. Shahjahan Shahid 1998	Design an advanced telephone system that integrates caller identification, answering machine, and tele-security system functionalities.	Caller ID decoding circuit with Motorola IC MC14LC5447 Answering machine circuits (ring detector, OGM sender, tape driver, DTMF encoder, CPC circuit) Tele-security system with sensors and auto-dialer. PIC16C84 8-bit	Caller ID information is decoded from FSK modem tones received between the first and second ring. Answering machine answers calls after set number of rings, plays OGM, records incoming messages on tape. Tele-security system detects sensor activation, hooks the line, and auto-dials a	Displays caller ID information on LCD.Outgoing message playback and incoming message recording on tape.Remote control and message retrieval via DTMF tones. Sensor-based security monitoring and auto-dialing capability.	Provides advanced telephone services like caller ID display, automatic call answering, and remote security monitoring.	Integrates multiple useful features into a single device. Low-cost design using an 8-bit microcontroller.Programmable functionality for customization.	Limited by capabilities of the 8-bit microcontroller.Tape-based answering machine may be less reliable than digital recorders.Security system has limited range based on sensor placement.	Successful design and implementation of a multi-functional "smart" telephone system.Combines caller ID, answering machine, and tele-security capabilities in a single low-cost device.Leverages existing telephone network for communication and remote monitoring.

		microcontroller as the central controller. LCD display, isolation transformer, relays, and other supporting circuits.	pre-programmed number. Microcontroller coordinates all operations based on programmed logic.					
Muhammad Zeeshan Khan, Shahid Mumtaz, Muhammad Imran 2019	The objective is to improve the efficiency and accuracy of face recognition in devices and accurately recognize individuals in real-time.	The author used a deep unified model based on CNN and edge computing for processing the date.	The input to the proposed system is in the form of an image, which can be uploaded either from the directory or by capturing through the device's camera.The captured image is processed to detect the faces in the frame using a deep unified model based on a convolution neural network.The detected faces are then compared with the known faces in the database to recognize and authenticate individuals. The processing of the data is done at	The proposed solution excels in face detection and recognition, boasting 94.6% and 85.5% accuracy, respectively. It ensures real-time performance through edge computing, making it efficient for educational settings. Scalable and efficient, it employs a deep unified model for data processing. Attendance data syncs with the cloud for easy management and privacy concerns are addressed in compliance with regulations.	The solution simplifies attendance, saves time, scales efficiently, cuts costs, and enhances security with facial recognition, offering major benefits to educational institutions.	The proposed solution is highly accurate and efficient in recognizing and authenticating individuals in real-time.The use of edge computing reduces data latency and increases the real-time response, making the system more efficient.The system is designed to work with a large number of simultaneous images from different smart classrooms, making it suitable for use in educational institutions.	The system requires a good training dataset to achieve efficient and acceptable results.The system may face challenges in recognizing faces in low-light or noisy environments. The system may face privacy concerns related to the collection and storage of facial data.	The solution explains an innovative face recognition approach using convolutional neural networks and edge computing to handle IoT data challenges. It details the methodology, practical implementation, and evaluation. Suggested are practical testing and exploration of broader applications.

			the edges of the nodes to reduce the data.					
Muhammad Iqbal, Bhakti Yudho Suprapto, Yul Yunazwin Nazaruddin 2022	The goal is real-time face and emotion recognition in humanoid robots using a custom CNN. Filling a research gap, it introduces a tailored CNN and distance-based human localization for high accuracy, with applications in robotics, healthcare, and security.	The components in this include a modified CNN architecture, a method for measuring the distance between the object's face and the position of the robot.	In the CNN architecture selection step, two well-known CNN architectures, i.e., VGG16 and AlexNet, are compared with the proposed modified architecture. The modified architecture is chosen because it has shown better performance in face recognition and emotion recognition. In the performance evaluation step, the proposed system's accuracy is measured as the	The proposed system utilizes a customized CNN architecture for real-time face and emotion recognition, employing a camera module linked to a laptop to capture visual data. It processes point coordinates of recognized faces, achieving impressive accuracy (87% for faces, 67% for emotions). Furthermore, the system adeptly calculates distances and performs well in different	The solution's positive impact is substantial, and applicable in healthcare, security, and entertainment, enhancing human-robot interactions through face and emotion recognition. It benefits patient monitoring, security, and user engagement across domains.	The modified CNN architecture has shown better performance in face recognition and emotion recognition than well-known CNN architectures, i.e., VGG16 and AlexNet. The study confirmed real-time recognition and precise distance measurement by humanoid robots. It underscored distance and illumination as critical recognition factors.	Dataset preparation demands ample data, often challenging to acquire. System accuracy may be influenced by lighting, expressions, and occlusions. Training and testing necessitate substantial computational resources and time, posing system challenges.	Overall, it holds promise for enhancing human-robot interactions across domains, provided ethical concerns are duly addressed.

			recognition rate of faces and emotions in realtime. The accuracy value shows the level of effectiveness per class of a classification.	lighting conditions, while also facilitating precise servo and dot matrix control.				
Mazin Hnewa, Hayder Radha 2021	The goal of the proposed solution MS-DAYOLO proposes a solution to the domain shift problem faced by object detection applications. It generates domain-invariant features using multiple domain adaptation paths and classifiers at different scales of the YOLOv4 detector.	The backbone, the neck, and the head. The backbone is responsible for extracting multiple layers of features at different scales. The neck collects these features and feeds them to the head. The head predicts bounding boxes.	Evaluate the performance of the object detector in the target domain and compare results of domain adaptation to MS-DAYOLO at different scales. Analyze the results and conclude that applying domain adaptation to all three feature scales improves the detection performance on the target domain and	The framework introduces multi-scale domain adaptation for YOLOv4 object detection, enhancing domain-invariant feature extraction. It employs end-to-end training and operates in realtime, making it valuable for time-critical applications. A significant advantage is that it doesn't require target domain data annotation. Furthermore, it	The analysis reveals noteworthy aspects of the work. The framework's novelty lies in its multipath domain adaptation approach, distinct from prior methods. The evaluation demonstrates its superiority over YOLOv4 but lacks comparison with other domain adaptation methods. Acknowledged limitations include increased resource requirements and data quality	The framework generates domain-invariant features that reduce the impact of domain shift, making it more robust to changes in the target domain. The proposed framework can be applied to a wide range of applications that involve crossdomain object detection, such as autonomous driving, surveillance, and	It is evaluated on a specific set of objects commonly found in autonomous driving scenarios. It may not be effective for other types of objects or scenes.	The solution holds promise in enhancing object detection systems for autonomous driving by mitigating domain shifts without data annotation. Autonomous driving heavily relies on object detection, and this adaptation can boost system reliability and safety, crucial for this domain.

	This improves the detector's performance when trained with the proposed solution and tested on target data for autonomous driving.		achieves the best result. Show that the proposed MS-DAYOLO outperforms the original YOLOv4 approach by a significant margin and almost reaches the performance of the ideal (oracle).	outperforms YOLOv4 in diverse autonomous driving scenarios, improving object detection performance.	dependence. Future work should consider different threshold values, alternative detection architectures, and applications beyond the current scope, providing potential avenues for further research and improvement.	robotics		
Md. Mahmud Hasan, Lim Hooi Jiun, Ng Wei Chuen 1998	The solution aims to explore the potential use of low-cost microcontrollers to implement advanced facilities and services related to everyday telephone use, such as Caller Identification and Answering	PIC16C84 microcontroller, Caller Identification (Caller ID) circuit, Answering Machine circuit, Tele-security system circuit, Ring detector circuit, Ring detector circuit, Outgoing message sender circuit, LCD display unit,	The implementation of the Smart Telephone Device by programming the PIC16C84 microcontroller in Intel 8052 8-bit microcontroller assembly language and assembling the hardware components. 5 The testing of the Smart	The Smart Telephone Device offers a range of features, including Caller Identification through a dedicated circuit displaying information on an LCD unit. It incorporates an answering machine to record messages during user unavailability and functions as a tele-	The work presents a cost-effective and easily accessible telephone solution that emphasizes affordability. The Smart Telephone Device comes with a Caller ID, an answering machine, and a telesecurity system that adds an extra layer of security. However, the tele-security system may	The use of low-cost microcontrollers and digital telephone networks makes Caller Identification and Answering Machine functions accessible to a wider audience. The Tele-security system circuit monitors the telephone line for any unauthoriz	The Smart Telephone Device is designed to perform specific functions, including Caller Identification, Answering Machine, and Tele-security system, and may not meet the needs of all users.	This solution offers a cost-effective and comprehensive answer for daily telephone needs in the project domain. The Smart Telephone Device, with Caller ID, an answering machine, and a tele-security system, proves

	g Machine functions. The problem that needs to be solved is making these features accessible to a wider audience by using affordable technology.	EEPROM memory	Telephone Device to ensure that it performs the intended functions, including Caller Identification, Answering Machine, and Tele-security system.	security system, activating alarms against unauthorized access or tampering. Utilizing low-cost microcontrollers ensures affordability. Designed for compatibility with prevalent digital telephone networks, the device explores the application of Signaling System No. 7 (SS 7).	not be suitable for household use, and there may be limitations to the device's functionality in areas with poor network coverage. Despite these limitations, the paper presents a valuable and practical solution for everyday telephone use.	ed access or tampering, providing enhanced security.The Answering Machine circuit records and plays back outgoing messages, and detects incoming messages, providing convenience to users.		practical for household use. Its use of low-cost microcontrollers and compatibility with digital networks enhances accessibility, while the added security layer contributes value to the overall project.
Yang ye, Hengxu you, Jingdu 2023	Aims to introduce the Robo GPT solution, aiming to enhance human-robot collaboration (HRC) in assembly tasks by addressing communication and trust issues.	ChatGPT, Robotic control modules, AI assistant, Human operator.	The Robo GPT AI assistant generates prompts, presents the prompts to human operators, and waits for further instructions. Such bidirectional communication clarifies the intention of both	The feature of this solution is the integration of ChatGPT in Robo GPT facilitates bidirectional communication, enhancing transparency and efficiency in decision-making between humansand robots.	The integration of ChatGPT in Robo GPT enhances transparency and efficiency of decision-making, resulting in better task performance and productivity. Bi-directional communication through ChatGPT enables	Improved communication and collaboration between humans and robotsIncreased transparency and reduced ambiguity. in the decision-making process.Enhanced trust in the robotic system.	The use of AI and natural language processing tools may require additional training and expertiseThe system may not be suitable for all types of assembly tasks or environments. The cost of implementing the system may be high.	Evaluates the impact of ChatGPT on trust in human-robot collaboration. The use of Large Language Models like ChatGPT improves semantic understanding in

	The primary challenge is the lack of effective communication between humans and robots, leading to errors and safety concerns. The proposed solution involves integrating ChatGPT, an AI-based natural language processing tool, with robotic control modules.		human operators and the robotic system, which could increase transparency and reduce ambiguity. AI assistant considers the information adequate for decision-making, the responses will be sent to a decoder which further processes the commands into Robotic Operating System (ROS) topics and triggers robotic control functions.	ChatGPT's natural language processing capabilities simplify human-robot interaction, enabling autonomous decision making and memory retention. Despite the text-based I/O, the system shows promise in fostering better partner relationships and elevating overall task performance in human-robot collaboration scenarios.	natural human-robot interaction, essential in industrial and manufacturing environments, where LLMs' use as intelligent robot assistants improve decision-making efficiency and transparency.			communication. The experiment using Robo GPT shows increased trust. However, potential miscommunication calls for robust and adaptable AI systems. The study examines ChatGPT's influence on trust, considering opportunities and limitations.
Abhishek Sarda, Dr. Shubhra Dixit, Dr. Anupama Bhan 2021	The goal of this solution is to enhance safety and efficiency by precisely identifying and categorizing objects	The solution uses YOLOv4's CSPDarknet53 for image feature extraction, a SPP (Spatial pyramid pooling) and PAN (Path	The YOLO model trains on the custom image data we provide to identify objects on images in google colab environment using yolo's	This algorithm is a single-shot detector that allows for real-time detection and classification of objects, making it suitable for use in	The authors provided a detailed explanation of the algorithm and its training process, and presented impressive results on a custom dataset. However, a	Training the YOLO algorithm on custom classes allows for the detection of specific objects. YOLO is a single-shot detector, which means it	The accuracy of the algorithm depends on the quality and quantity of the training data. It may not perform as well as multi-stage detectors in	The solution of using YOLO algorithm can significantly reduce accidents and improve safety.

	on the road. This addresses the issue of accidents stemming from human error, reducing them through autonomous driving technology.	Aggregation Network) layers for processing of images, a YOLOv4 detection model for drawing bounding box and class probability predictions, its data augmentation techniques include Mosaic augmentation and DropBlock regularization, it uses Mish and Leaky ReLU activation functions, the Darknet training framework, it was trained on google colab using GPU.	training algorithm. The new unseen input image is processed through a neural network to detect and classify objects in real-time.The objects on the image are identified and classified into different objects.	autonomous driving.	more thorough discussion of limitations and comparisons with other algorithms would have enhanced the analysis.	can detect and classify objects in real-time. The Identified objects are separated by bounding boxes.	detecting small objects.	
Shaji Thorn Blue, M. Brindha 2019	The aim of this work is improving the precision of boundary boxes that will be been drawn	The proposed work includes the YOLOv3 algorithm, which is used to obtain the coordinates of boundary	The algorithm evaluates the accuracy of object localization and identification within images using YOLOv3	This solution prioritizes precision by enhancing the accuracy of boundary boxes around detected objects, outperforming YOLOv3	The analysis highlights the integration of YOLOv3 for object detection, followed by image pre-processing which includes Gaussian	Reduces computation and improves algorithm speed. Smooths colour transitions and	May miss objects outside the detected area. May blur important details.	The solution enhances object detection precision with edge detection and pixel analysis.

	after objects has been detected on image in order to improve the accuracy of object localization and identification within images.	boxes and the class of the object. After YOLOv3 outputs an image, pre-processing is applied to obtain an image with the edges of the detected object. The proposed work also includes an edge detection algorithm, image partitioning into four parts, and the construction of a new boundary box based on threshold values.	algorithm through which edges can be detected so that boundary boxes can be drawn accurately.	through edge detection and pixel values. Leveraging YOLOv3 as a foundation, it utilizes pre-trained COCO data, determines threshold values for boundary sides, and excels in producing accurate boundary boxes.	blur and Canny edge detection. New boundary boxes are constructed based on white pixel density analysis. Results indicate improved accuracy over YOLOv3, particularly in terms of IOU (intersection over union) values, though challenges exist in handling sharp objects and noisy images.	removes noise. Accurately detects edges. Focuses on areas where boundary box edges are more likely to be found. Draws more precise boundary boxes.	May miss some edges or produce false positives. May miss boundary box edges in the middle of the image. May struggle with sharp objects and noisy images.	
Shraddha Mane, Prof. Supriya Mangale 2018	The goal of this solution is to accurately detect and track moving objects in real-time video streams. The problem that needs to be solved is the	Moving object detection using TensorFlow object detection API, Location of detected object passed to object tracking algorithm, CNN-based	Our approach uses TensorFlow object detection API for moving object detection and a CNN-based tracking algorithm for robust object detection. Offline	The key features of this solution include a novel Tensor flow-based object detection and CNN-object tracking system for robust object detection and tracking in complex	The solution offers a novel, robust approach to object detection and tracking in complex scenarios which enhances accuracy and exhibits efficient performance	This API provides a simple and efficient way to construct, train, and detect models for object detection. This algorithm is designed to be robust to	This may require some knowledge of TensorFlow and deep learning methods. This may require a large number of computational resources. Offline training may require a large amount	This solution focuses on object detection and tracking through a unique approach involving TensorFlow's object detection API and CNN algorithms.

	difficulty of detecting and tracking objects accurately and robustly in real-world scenarios.	object tracking algorithm for robust object detection, Offline training to learn spatial and temporal classes for human tracking, shift variant architecture to combine global features and local characteristics.	training and shift-variant architecture are also used for improved accuracy.	scenes. Evaluation involves sensitivity, specificity, and accuracy metrics, derived from TP, TN, FP, and FN values,assessing the system's performance in identifying moving and stationary objects.	in Python-based testing.	variations in illumination and occlusion. This improves the accuracy of object detection and tracking. This improves the accuracy of object detection and tracking.	of data and time to train the model. This may require some knowledge of computer vision and image processing techniques.	Evaluation across diverse video sequences reveals impressive sensitivity, specificity, and accuracy metrics. The use of a shift variant architecture enhances CNN performance. The approach seems promising in terms of accuracy and robustness.
Menaka, Archana, Dhana Gopal, Ramesh 2020	The aim of this solution is to develop an Enhanced Missing Object Detection System using the YOLO approach. The problem that this system aims to solve is the identification of missing objects in real-time by continuously	Object Identification Techniques (YOLO), Bounding Box Prediction, Anchor Box Techniques, Comparison with Other Real-Time Systems, Frame Analysis.	The solution involves dividing the image into grids, predicting bounding boxes and class probabilities for objects in each grid cell, comparing objects in current and previous frames to identify missing objects, and generating alerts. It utilizes the YOLO algorithm for real-time object	This solution features the use of convolutional layers trained on the ImageNet dataset for object detection. It leverages the efficient YOLO algorithm for direct prediction of bounding boxes and class probabilities. A frame analysis algorithm stores objects across frames,	This work outlines a system that employs the YOLO algorithm for real-time object detection, primarily for surveillance. The methodology includes convolutional layers, frame analysis, and data augmentation techniques. The system's potential applications span surveillance, object detection, with	The YOLO algorithm helps in detecting objects directly from full image in a single pass, making it faster compared to other detection algorithms. Comparing frames helps in continuous monitoring of the scene without the need for manual	The YOLO algorithm may make some localization errors but predicts fewer false positives in the background. The effectiveness of the alert system depends on the reliability of the communication channels and the responsiveness of the user to the alerts.	The YOLO algorithm's efficiency, real-time capabilities, and adaptability across domains, including surveillance and face recognition, provide a valuable solution for accurate object detection and missing object alerts.

	analysing frames and alerting the user.		detection and classification.	enabling missing object detection. Training involves YOLO's pretrained model and fine-tuning.	prospects for further enhancements like object-co detection and GSM integration for user alerts.	observation. The alert system ensures that the user is immediately informed about any missing objects, allowing for timely action.		
Tao Feng, Yao Yu, Lin Wu, Yanru Bai, Ziang Xiao, Zhen Lu. 2018	The goal of the proposed solution is to implement a human-tracking robot using ultra-wideband (UWB) technology. The problem that needs to be solved is the detection and calculation of the position of the target person, which is essential for the implementation of the human-	The proposed solution encompasses a range of hardware components to support its functionalities, including Ultra-wideband (UWB) technology for human tracking, ultrasound modules to estimate distances and enable obstacle avoidance, a Bluetooth module for data communication, an A33 SOC responsible for UWB data reception, calibration,	The proposed solution uses UWB technology. The UWB data is calibrated and filtered in the A33 SOC, while the STM32F103 MCU controls the robot's movement, data communication, and distance estimation. The human-tracking algorithm uses a modified virtual spring model to track the target person smoothly.	The robot features leveraging a UWB technology, it excels in precision and anti-interference abilities providing a reliable human tracking model. The human-tracking algorithm, employing a modified virtual spring model, ensures robust performance in various environments. Experimental results validate its effectiveness in distance detection and overcoming	The current model uses a UWB technology, a hyperbolic positioning algorithm, and a virtual spring model in order to find a target person's position and orientation, track the target person and control the robot's movements respectively.	Robustness against environmental conditions. Enables effective tracking and following of the target person. Provides target person position and orientation data accurately. Enables efficient movement and tracking of the target person.	-	UWB technology positively impacts the human-tracking robot by offering precise positioning, robustness in diverse environments, low system complexity, high location precision, and potential mainstream adoption.

	tracking function.	filtering, and executing the human tracking algorithm.		measurement errors.				
Anis Koubaa 2023	The goal of the ROSGPT solution is to enhance human-robot interaction by leveraging large language models (LLMs) like ChatGPT and integrating them with the Robot Operating System (ROS). The specific problem that ROSGPT aims to solve is the challenge of seamless communication between humans and robots.	Large language models (LLMs) such as ChatGPT, GPT-3, and GPT-4, the Robot Operating System (ROS), the ROS Parser module, prompt engineering technique s, ontology-based approache s, the integration of ChatGPT with ROS.	The paper introduces ROSGPT, which combines large language models with robotic systems for human-robot interaction. ROSGPT integrates ChatGPT with ROS2 to convert natural language instructions into structured robotic commands using zero-shot and few-shot learning.	ROSGPT innovatively integrates Large Language Models (LLMs) with the Robot Operating System (ROS), fostering intuitive human-robot interaction. Guided by ontology, it translates unstructured commands into task-specific instructions.	The paper critically analyses ROSGPT, focusing on ChatGPT's performance in human-robot interaction without ontology-based prompting. It commends ChatGPT's strong elicitation and adaptability, noting accurate JSON command generation but highlights limitations, including deviation from ontology affecting responses.	The model demonstrates the ability to make appropriate calculations and generate accurate JSON structures based on given information and can adapt to new tasks with only a few training examples, making human-robot interaction more flexible and adaptable. By leveraging structured knowledge from the ontology, the model can more effectively process natural language commands and translate	Without the guidance of an ontology, the model may struggle to understand complex relationships between entities and concepts, leading to potential inaccuracies or misunderstandings in the generated commands. The effectiveness of the system heavily relies on the quality and completeness of the ontology. Inaccuracies or limitations in the ontology can impact the performance of the model.	ROSGPT enhances human-robot interaction, boosts efficiency through LLMs, enables real-time command execution, and bridges ROS and ChatGPT for seamless communication.

						them into accurate actions.		
Xiao-Feng Zhao, Zi-Heng Chen, He-Feng Yin, Xiao-Jun Wu 2023	To develop an intelligent visitor system that can utilize the advantages of both cloud computing and edge computing to efficiently register and manage visitors in an unmanned environment using AI technologies like face recognition and speech recognition.	Edge Computing Layer: RK3399 motherboard running Android 8.0, camera, microphone array, ID card reader, infrared temperature sensor, etc. Cloud Service Layer: AI cloud services (face recognition, speech recognition), visitor application cloud services (data storage, management).	Edge devices collect data (face images, voice, ID card) from sensors/hardware. Cloud AI services perform face recognition, speech recognition on the data from edge. Edge devices get results from cloud and provide visitor self-registration using face/voice. Cloud stores and manages all visitor data which can be accessed remotely.	Unmanned visitor registration using face recognition and voice interaction. Real-time visitor identity and temperature verification. Remote data storage, management and access via cloud application.	High recognition accuracy (over 80% for registered faces) Short recognition delay (average 382ms after initial 1.1s connection).	Improves efficiency and reduces labor cost compared to manual visitor management. Ensures data authenticity and prevents human errors. Integrates temperature screening for COVID prevention.	Requires decent computing power and internet connectivity at the edge for interaction with cloud. Cloud service used has limited throughput of 2 queries per second.	The system has been trialed at enterprise front desks. Shows high stability and is well-received for automating visitor reception during COVID times.

CHAPTER 3

PROPOSED SYSTEM

3.1 PROPOSED SYSTEM

The primary objective of the proposed system is developing a conversational robot with the advanced functionalities like speech, object detection, face recognition. This innovative system aims to engage users in friendly conversations while identifying objects and faces, enhancing user experiences across various domains. With advanced technology and user-centric design, it promises to redefine human-robot interactions.

3.2 OBJECTIVES OF PROPOSED SYSTEM

The objectives of the proposed system include the following:

- A face recognition module capable of recognizing specific users and extracting relevant information about them.
- An object recognition system to identify and categorize various objects based on visual input.
- A tailored visitor tracking system to keep track of people who are approaching the user in his absence.
- An "infobot" functionality to provide tailored information and assistance to recognized users based on their individual needs.
- A memory storage system to retain information about known users, them preferences, and relevant data.

3.3 ADVANTAGES OF PROPOSED SYSTEM

The proposed system has the following advantages:

- Enhanced security with personalized access control.
- Streamlined object categorization for efficient organization.
- Real-time visitor monitoring for improved security.
- Personalized assistance for recognized users' needs.

- Efficient data retention and retrieval for user preferences.
- Seamless integration with existing systems for scalability.
- Reduced manual intervention through automated processes.
- Improved user experience with tailored information delivery.
- Enhanced data analysis capabilities for insights.
- Flexibility to adapt to evolving user requirements.

3.4 SYSTEM REQUIREMENTS

The system requirements for the development and deployment of the project as an application are specified in this section. These requirements are not be confused with the end-user system requirements.

S.NO	Requirements	Requirement type	Explanation
1.	Python	Programming language	Used to write the code and run.
2.	Visual Studio Code	Development Environment	Environment to write and execute the code.
3.	YOLO (You Only Look Once) model	Pretrained object detection model	Used for object detection and recognition
4.	ChatGPT API	API for chatbot	Used for response generation.

Table 3.4.1 Requirements for developing and deploying the application

3.4.1 SOFTWARE REQUIREMENTS

Below are the software requirements for application development:

- Operating System : Windows, Raspberry pi 64bit OS.
- Programming Language : Python for implementing face recognition, object detection, speech analysis, tracking visitor information
- Development Environment : Visual Studio Code.
- Libraries : OpenCV, NumPy, Face Recognition, OpenAI, Speech Recognition, pyttsx3, smtplib, email

3.4.2 HARDWARE REQUIREMENTS

Below are the hardware requirements for the application development:

- Device : Raspberry pi 5
- Operating System : Raspberry pi 64bit OS
- Processor : 64-bit 2.4 GHz quad-core ARM Cortex-A76
- Ram : 4 GB (minimum)
- Hard Disk : 256GB (minimum)
- Output Devices : Speaker.
- Input Devices : Keyboard, mouse, Camera, Mic.

3.4.3 FUNCTIONAL REQUIREMENTS

1. Input: The robot should accept video frames and speech as input.
2. Face Recognition: The robot should detect and classify the users as known and unknown.
3. Object Recognition: The robot should detect and recognize objects.
4. Visitor tracking: The robot should detect visitors and gather information about the visitors.
5. Response Generation: The robot should give relevant responses to the user's queries.
6. Output: Speech output is given by the robot as the response.

3.4.4 NON-FUNCTIONAL REQUIREMENTS

1 Reliability:

a) The system should be reliable, producing consistent results in face and object classification across multiple executions.

b) It should handle errors gracefully and provide informative error messages to users.

2 Security:

a) The system should ensure the confidentiality and integrity of user data throughout processing.

b) It should implement appropriate access controls to prevent unauthorized access to sensitive information.

3 Accuracy:

a) The face recognition and object recognition should be accurate to ensure reliable image classification.

b) The YOLO model should achieve high accuracy in object classification, minimizing misclassification errors.

3.5 IMPLEMENTATION TECHNOLOGIES

3.5.1 FACE RECOGNITION MODULE IN PYTHON:

Face recognition is a crucial aspect of computer vision applications. It involves identifying and verifying individuals based on their facial features. A face recognition module in Python utilizes various algorithms and techniques to achieve accurate recognition.

Face Recognition Algorithms:

1. Preprocessing:

a. Normalize images for consistent features.
b. Detect facial landmarks for precise alignment.

2. **Feature Extraction:**
    a. Extract discriminative features using methods like Eigenfaces or Convolutional Neural Networks (CNNs).

3. **Classification:**
    a. Utilize machine learning models such as Support Vector Machines (SVMs) or deep learning models for classification.

4. **Verification and Identification:**
    a. Verify identities by comparing extracted features with stored templates.
    b. Identify individuals by matching against a database of known faces.

Python Libraries for Face Recognition:

1. **OpenCV**: Provides image processing tools and pre-trained models for face detection and recognition.
2. **Dlib**: Offers facial landmark detection and face recognition functionalities.
3. **Face Recognition Library**: A Python library specifically designed for face recognition tasks, built on top of dlib.

Algorithm for Face Recognition:

1.Data Collection:
    a. Gather a dataset of face images with labeled identities.

2.Preprocessing:
    a. Resize images and detect facial landmarks for alignment.

3.Feature Extraction:
    a. Extract facial features using a chosen algorithm (e.g., CNN-based feature extraction).

4.Model Training:
    a. Train a classification model on the extracted features using machine learning techniques.

5.Face Recognition Process:
    a. Detect faces in new images using the trained model.
    b. Extract features and compare against the stored templates.
    c. Determine the identity based on similarity scores or classification results.

3.5.2 YOLO PRETRAINED MODEL:

YOLO is a popular object detection algorithm known for its real-time performance and accuracy. The pretrained YOLO model in Python leverages deep learning techniques to detect objects in images or videos efficiently.

YOLO was proposed by Joseph Redmond et al. in 2015. It was proposed to deal with the problems faced by the object recognition models at that time, Fast R-CNN is one of the state-of-the-art models at that time but it has its own challenges such as this network cannot be used in real-time, because it takes 2-3 seconds to predicts an image and therefore cannot be used in real-time. Whereas, in YOLO we have to look only once in the network i.e. only one forward pass is required through the network to make the final predictions.

YOLO Algorithm Overview:

1. **Single Shot Detection:**
   - YOLO follows a single shot detection approach, predicting bounding boxes and class probabilities in a single pass through the neural network.

2. **Anchor Boxes:**
   - Utilizes anchor boxes to predict multiple bounding boxes for each grid cell, improving accuracy in detecting objects of different sizes.

3. **Feature Extraction:**
   - Employs a deep convolutional neural network (CNN), such as Darknet or YOLOv3, to extract features from input images.

4. **Bounding Box Prediction:**
   - Predicts bounding box coordinates (x, y, width, height) and objectness score for each grid cell.

5. **Class Prediction:**
   - Assigns class probabilities to each bounding box, indicating the presence of specific objects (e.g., person, car, dog).

Python Libraries for YOLO:

1. **Darknet Framework:** Original framework for YOLO implementation, written in C and CUDA, with Python bindings available.

2. **YOLOv5:** A popular Python-based implementation of YOLO with pretrained models and easy-to-use APIs.

3. **OpenCV:** Used for image processing tasks and integration with YOLO models for object detection.

YOLO Pretrained Model Usage:

Model Loading:

1. Load a pretrained YOLO model using available libraries or custom implementations.

2. **Input Preparation**: Prepare input images or video frames for detection, resizing them to the model's input size.

3. **Inference:** Perform inference on the input data using the loaded YOLO model to detect objects and their bounding boxes.

4. **Post-processing**: Apply non-maximum suppression (NMS) to filter out overlapping bounding boxes and improve detection accuracy.

5. **Output Visualization:** Visualize the detected objects and bounding boxes on the input images or video frames for analysis and decision-making.

3.5.3 CONVOLUTIONAL NEURAL NETWORKS (CNN):

A Convolutional Neural Network (CNN) is a type of Deep Learning neural network architecture commonly used in Computer Vision. Computer vision is a field of Artificial Intelligence that enables a computer to understand and interpret the image or visual data. When it comes to Machine Learning, Artificial Neural Networks perform really well. Neural Networks are used in various datasets like images, audio, and text. Different types of Neural Networks are used for different purposes, for example for predicting the sequence of words we use Recurrent Neural Networks more precisely an LSTM, similarly for image classification we use Convolution Neuralnetworks. In this blog, we are going to build a basic building block for CNN. Convolutional NeuralNetwork (CNN) is the extended version of artificial neural networks (ANN) which ispredominantly used to extract the feature from the grid-like matrix dataset. For example: visual datasets like images or videos where data patterns play an extensive role.

CNN architecture

Convolutional Neural Network consists of multiple layers like the input layer, Convolutional layer, Pooling layer, and fully connected layers.

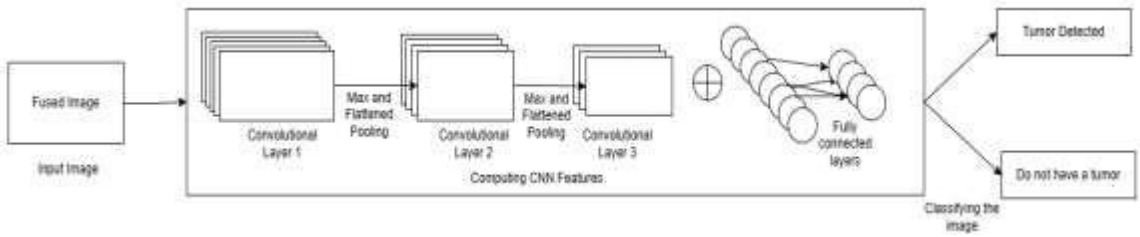


Figure 1: CNN Architecture

The Convolutional layer applies filters to the input image to extract features, the Pooling layer down samples the image to reduce computation, and the fully connected layer makes the final prediction. The network learns the optimal filters through backpropagation and gradient descent.

How Convolutional Layers works

Convolution Neural Networks or covnets are neural networks that share their parameters. Imagine you have an image. It can be represented as a cuboid having its length, width (dimension of the image), and height (i.e. the channel as images generally has red, green, and blue channels).

Now imagine taking a small patch of this image and running a small neural network, called a filter or kernel on it, with say, K outputs and representing them vertically. Now slide that neural network across the whole image, as a result, we will get another image with different widths, heights, and depths. Instead of just R, G, and B channels now we have more channels but lesser width and height. This operation is called **Convolution**. If the patch size is the same as that of the image it will be a regular neural network. Because of this small patch, we have fewer weights.

Now let's talk about a bit of mathematics that is involved in the whole convolution process.

- Convolution layers consist of a set of learnable filters (or kernels) having small widths and heights and the same depth as that of input volume (3 if the input layer is image input).

- For example, if we have to run convolution on an image with dimensions 34x34x3. The possible size of filters can be axax3, where 'a' can be anything like 3, 5, or 7 but smaller as compared to the image dimension.

- During the forward pass, we slide each filter across the whole input volume step by step where each step is called stride (which can have a value of 2, 3, or even 4 for high-dimensional images) and compute the dot product between the kernel weights and patch from input volume.

- As we slide our filters, we'll get a 2-D output for each filter and we'll stack them together as a result, we'll get output volume having a depth equal to the number of filters. The network will learn all the filters.

Layers used to build ConvNets

A complete Convolution Neural Networks architecture is also known as covnets. A covnets is a sequence of layers, and every layer transforms one volume to another through a differentiable function.

Types of layers:

Let's take an example by running a covnets on of image of dimension 32 x 32 x 3.

- **Input Layers:** It's the layer in which we give input to our model. In CNN, Generally, the input will be an image or a sequence of images. This layer holds the raw input of the image with width 32, height 32, and depth 3.

- **Convolutional Layers:** This is the layer, which is used to extract the feature from the input dataset. It applies a set of learnable filters known as the kernels to the input images. The filters/kernels are smaller matrices usually 2×2, 3×3, or 5×5 shape. it slides over the input image data and computes the dot product between kernel weight and the corresponding input image patch. The output of this layer is referred as feature maps. Suppose we use a total of 12 filters for this layer we'll get an output volume of dimension 32 x 32 x 12.

- **Activation Layer:** By adding an activation function to the output of the preceding layer, activation layers add nonlinearity to the network. it will apply an element-wise activation function to the output of the convolution layer. Some common activation functionsare **RELU**: max (0, x), **Tanh**, **Leaky RELU**, etc. The volume remains unchanged hence output volume will have dimensions 32 x 32 x 12.

- **Pooling Layer:** This layer is periodically inserted in the covnets and its main function is to reduce the size of volume which makes the computation fast reduces memory and also prevents overfitting. Two common types of pooling layers are **max pooling** and **average pooling**. If we use a max pool with 2 x 2 filters and stride 2, the resultant volume will be of dimension 16x16x12.

- **Flattening:** The resulting feature maps are flattened into a one-dimensional vector after the convolution and pooling layers so they can be passed into a completely linked layer for categorization or regression.x

- **Fully Connected Layers:** It takes the input from the previous layer and computes the final classification or regression task.
- **Output Layer:** The output from the fully connected layers is then fed into a logistic function for classification tasks like sigmoid or softmax which converts the output of each class into the probability score of each class.

3.5.4 CHATGPT AND ITS USAGE WITH API:

ChatGPT is an advanced language model developed by OpenAI based on the GPT (Generative Pre-trained Transformer) architecture. It is trained on a diverse range of text data and can generate human-like text responses to various prompts.

ChatGPT Algorithm Overview:

Transformer Architecture:

ChatGPT is built upon the Transformer architecture, a deep learning model specifically designed for natural language processing tasks. This architecture enables ChatGPT to understand complex linguistic patterns, dependencies, and contexts, leading to coherent and contextually relevant responses.

Pretraining:

During pretraining, ChatGPT learns from a vast amount of text data across various domains. This process allows ChatGPT to capture diverse language nuances, common phrases, and contextual information, enhancing its ability to generate human-like responses across different topics.

Fine-tuning:

When prompted with input text or questions, ChatGPT leverages its pretrained knowledge and language modeling capabilities to generate text-based responses. It uses contextual understanding and pattern recognition to produce coherent and meaningful replies that match the input context.

Response Generation:

Given a prompt or input text, ChatGPT generates coherent and contextually relevant responses using its learned knowledge and language modeling capabilities.

ChatGPT API Usage:

API Integration:

Integrating the ChatGPT API into applications or platforms involves connecting to OpenAI's servers, enabling real-time communication with the ChatGPT model. This integration allows seamless access to ChatGPT's conversational abilities within your software environment.

Text Generation:

Users can send prompts or queries to the ChatGPT API, which processes the input and generates text-based responses. This functionality is beneficial for applications requiring interactive conversations, virtual assistants, chatbots, or content generation.

Customization:

The ChatGPT API supports model customization through fine-tuning. By providing domain-specific data or fine-tuning parameters, developers can enhance ChatGPT's performance for specific use cases, ensuring that generated responses are accurate and relevant to the application's context.

Multilingual Support:

ChatGPT and its API offer multilingual support, allowing developers to create applications that cater to users speaking different languages. This feature expands the reach of applications to global audiences and facilitates multilingual conversations seamlessly.

Contextual Understanding:

ChatGPT maintains context across interactions, understanding the flow of conversations and retaining information from previous messages. This contextual understanding enables ChatGPT to provide coherent and engaging conversations, mimicking human-like conversational dynamics.

Error Handling:

Integrating error handling mechanisms within the ChatGPT API ensures robustness and reliability. By handling errors gracefully, such as detecting ambiguous queries or out-of-context requests, developers can improve the overall user experience and maintain the conversational flow effectively.

CHAPTER 4

SYSTEM DESIGN

4.1 PROPOSED SYSTEM ARCHITECTURE

The proposed system encompasses the development of a robot that can recognize faces and objects and also extract information about them.

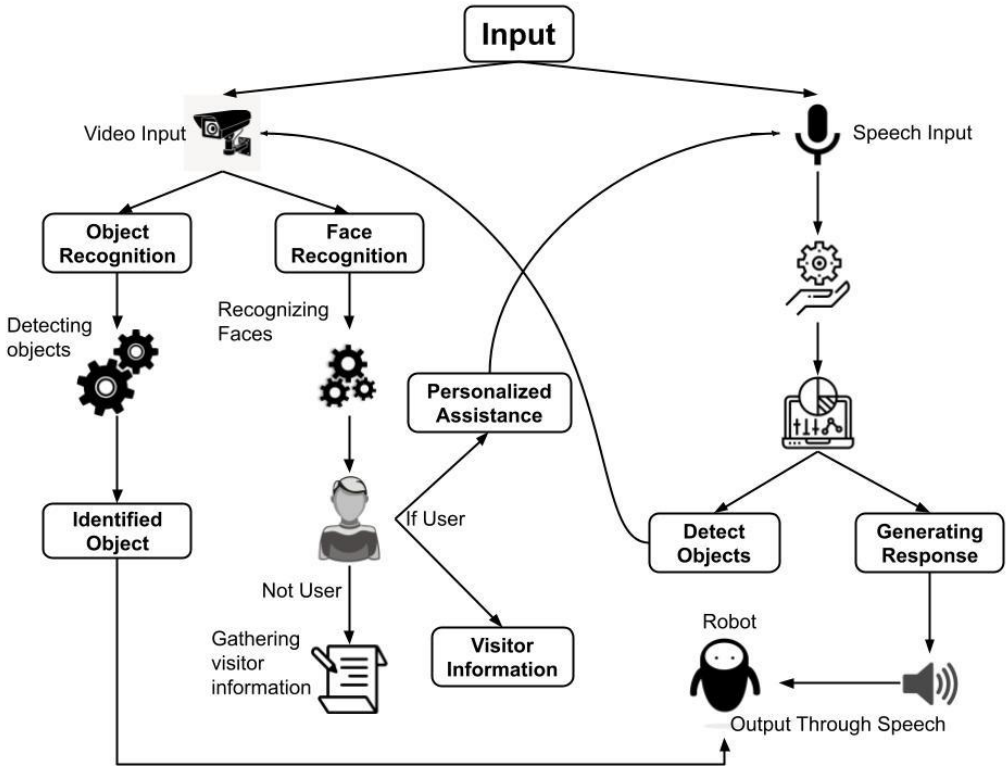


Figure 2: Proposed Architecture

4.2 APPLICATION MODULES

On The application on an overall involves Three main modules, which cater to the four main functions of this application, i.e. to recognize face, detect objects, interact with user through ChatGPT, visitor tracking system.

4.2.1 Face Recognition Module

The robot is equipped with an advanced facial recognition system leveraging deep learning models to differentiate between different users accurately. Using a Raspberry Pi-compatible camera, it captures facial features and processes them through neural networks for precise identification. Upon identifying a user, the robot accesses its internal database to retrieve relevant information, enabling

it to offer a tailored and personalized experience. This sophisticated system not only enhances security by verifying user identities but also allows the robot to adapt its interactions based on individual preferences and past interactions, creating a more engaging and customized interaction for users.

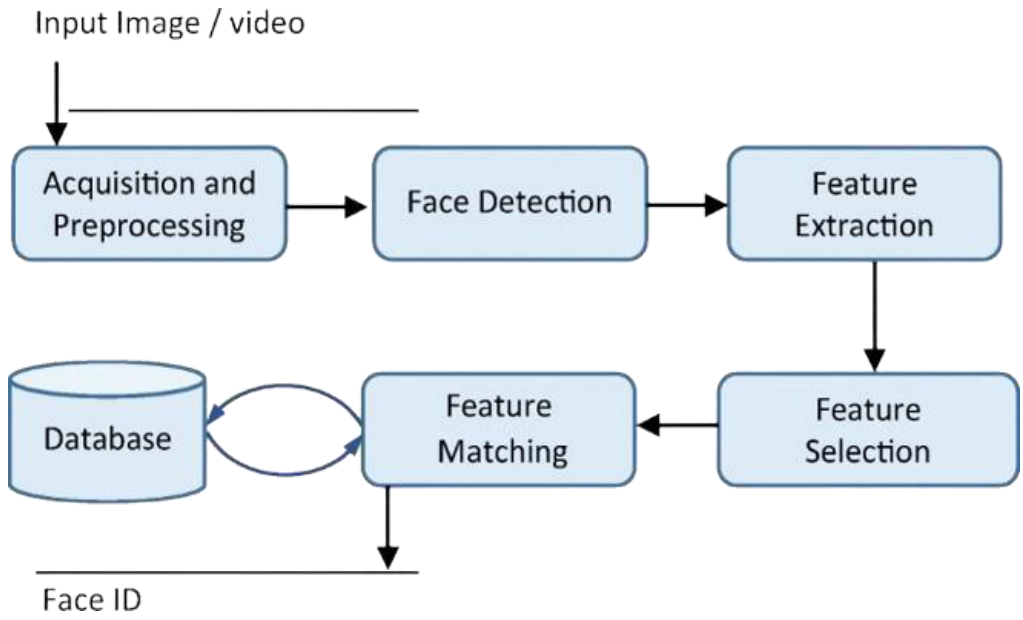


Figure 3: Workflow of Face Recognition process

4.2.2 Object Detection Module

Frames Acquisition and Preprocessing: The robot receives frames by means of continuously shooting the video feed this is furnished with the aid of the use of its camera. preprocessing including resizing, normalization, or color area conversion can be administered to the captured body to in addition prepare the enter to the item detection model and enhance its high quality. Object Detection: The robot uses an object detection version, known as YOLOV8, to grow to be privy to and localize gadgets at the frames capturing the scene. The item detection version is professional on a large-scale dataset of classified images and might apprehend items primarily based on their visible look. It additionally strategies each body and extracts capabilities of every item visible.

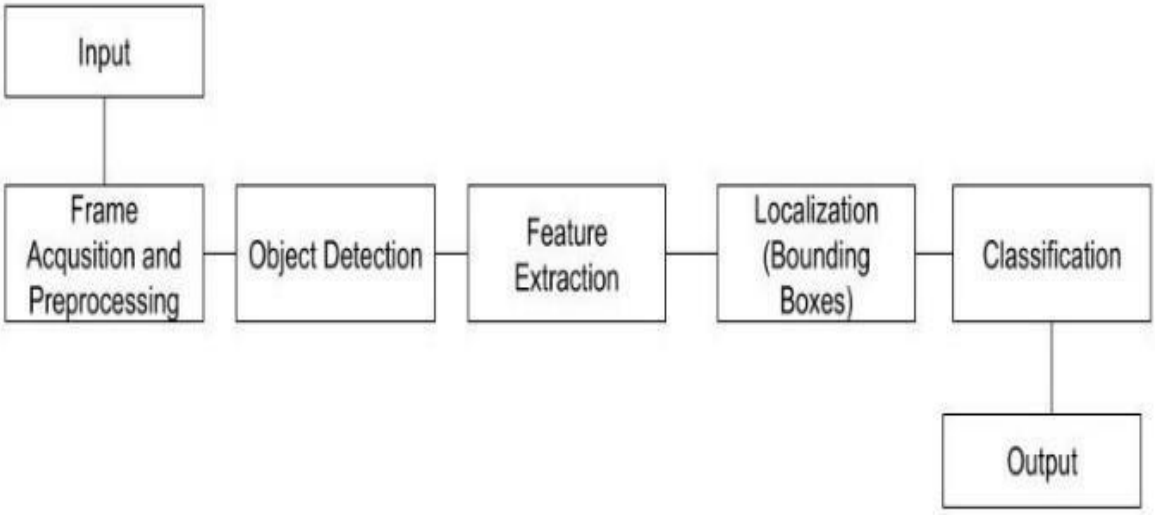


Figure 4: Workflow of Object Detection and Recognition

4.2.3 Chatbot Module

While calling ChatGPT via an API, you first ship an API requestto OpenAI with a set of non-obligatory parameters like temperature, max tokens, and pinnacle p sampling to influence thereaction's creativity and period. OpenAI preprocesses your request,together with responsibilities like textual content cleansing andtokenization, earlier than selecting the best ChatGPT version variant primarily based on elements like set-off complexity andrequest parameters. the selected model, possibly based on transformers, then strategies the set-off, is aware of its context andgenerates a reaction following the prompt's cause. put up- processing steps inclusive of filtering and blunders correction maybe carried out to the uncooked output, resulting in a very last reaction textual content this is sent again to you because of the API response, probably including confidence ratings. it is essential to be aware that the precise inner workings of ChatGPT are proprietary,and OpenAI may additionally replace and improve the gadget through the years.

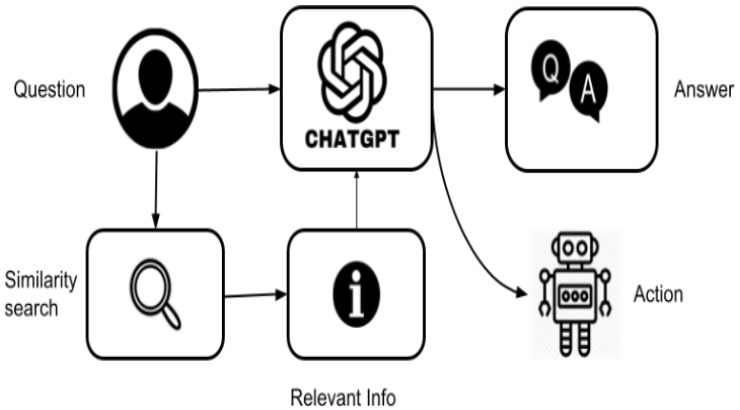


Figure 5: Workflow of Chatbot

4.2.4 Visitor Tracking System

The innovation tackles the challenge of interacting with unidentified guests by incorporating facial recognition capabilities. Upon detecting a new face, the robot initiates a conversation to gather essential details about the visitor and the purpose of their visit. This information is securely stored within the system, contributing to a comprehensive visitor tracking system. By proactively engaging with unfamiliar faces and capturing relevant data, the robot ensures efficient communication and a seamless experience for both guests and users. Our machine capabilities are a visitor monitoring machine that uses the face recognition version to pick out recognized and unknown individuals.

Video frames are processed to compare facial features with stored data. If recognized, the individual is identified otherwise, they're categorized as unknown. This complements personalization and safety, permitting the robotic to adapt interactions for that reason.

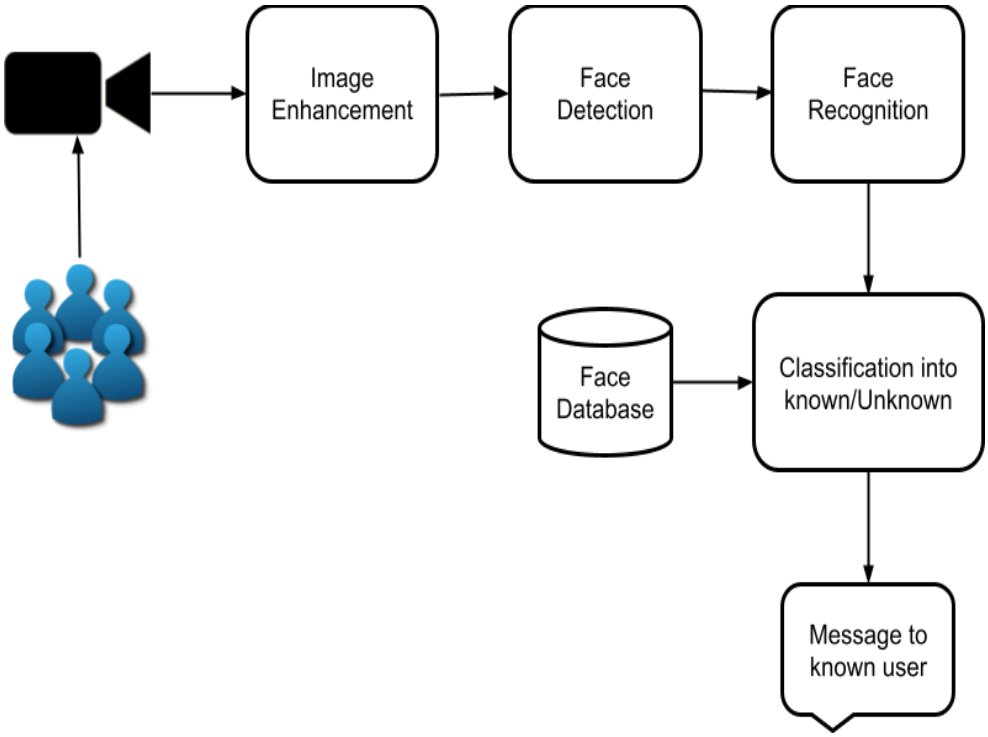


Figure 6: Workflow of Visitor Tracking System

4.3 UML Diagrams

UML stands for Unified Modelling Language. UML is a standardized fashionable-cause modelling language in the subject of object-oriented software engineering. In its modern shape, UML comprises of two essential components: a Meta-model and a notation. The Unified Modelling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software program machine, in addition to for commercial enterprise modelling and other non-software systems. The UML uses more often than not graphical notations to express the design of software program projects.

4.3.1 Use Case Diagram

In the Unified Modeling Language (UML), a use case diagram is a behavioral diagram that stems from use-case analysis. Its number one objective is to provide a visual summary of a gadget's capability, showcasing actors, their objectives (portrayed as use cases), and any relationships amongst those use cases. The fundamental aim of a use case diagram is to demonstrate which device capabilities are accomplished for each actor worried, while additionally illustrating the jobs played via these actors within the gadget.

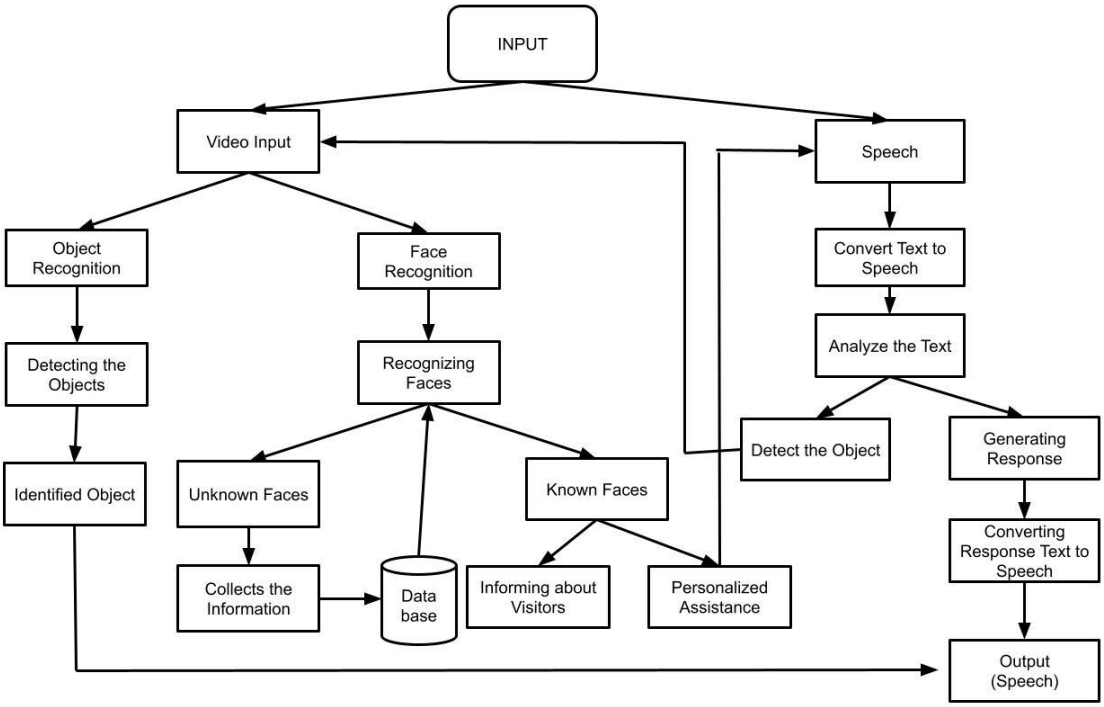


Figure 7: Use case Diagram

4.3.2 Sequence Diagram

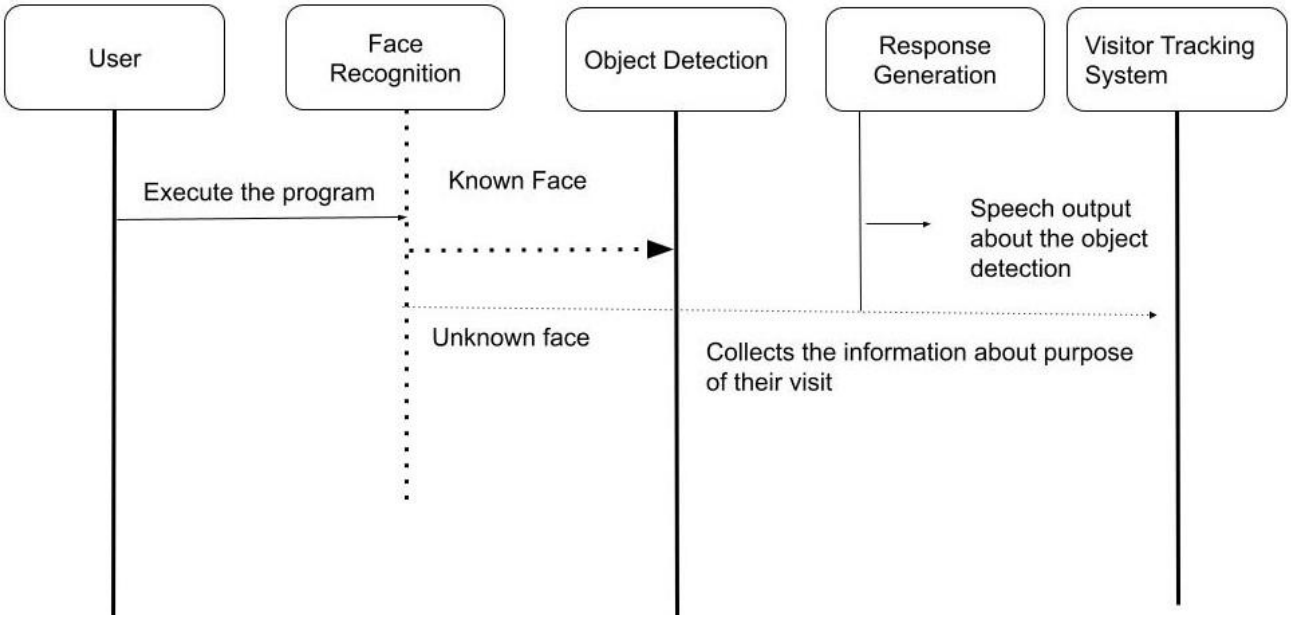


Figure 9: Sequence Diagram

4.3.3 Activity Diagram

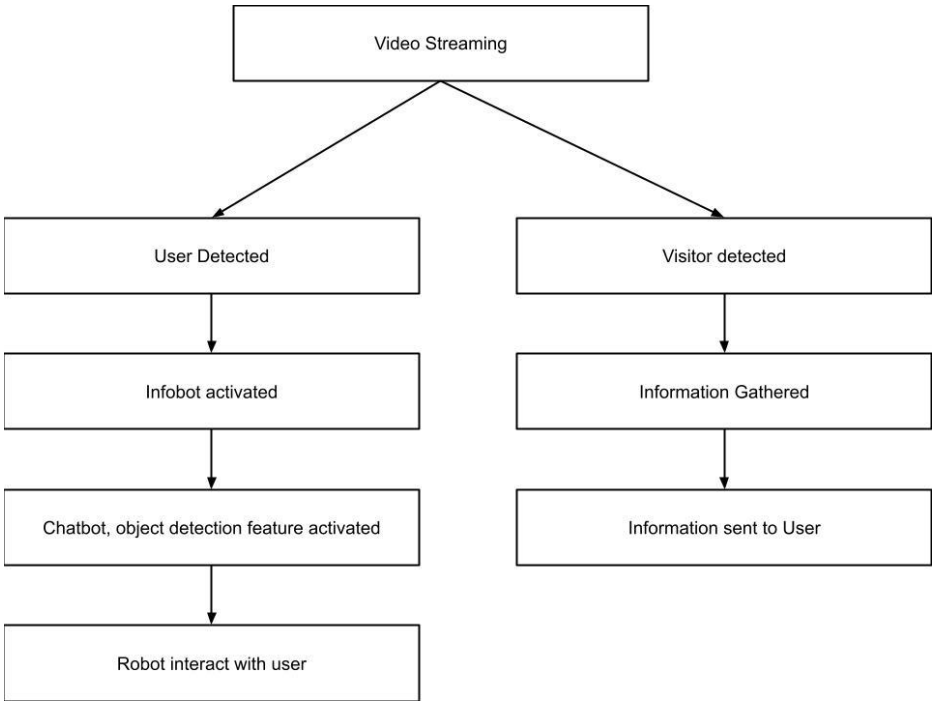


Figure 10: Activity Diagram

CHAPTER 5

IMPLEMENTATION

5.1 SOURCE CODE

5.1.1 Main Code

```
import cv2
import pyttsx3
import speech_recognition as sr
import face_recognition

import phrases as ph
import chatbot as cb
import face_recognize as l
import object as ob
import visitors as v

engine = pyttsx3.init("sapi5")
voices = engine.getProperty('voices')
engine.setProperty('voice', voices[1].id)

def speak(audio):
    engine.say(" "+audio)
    engine.runAndWait()

def listen():
    print("listening...")
    r = sr.Recognizer()
    with sr.Microphone() as source:
        r.adjust_for_ambient_noise(source)
        audio = r.listen(source)
    try:
        query = r.recognize_google(audio, language='en-in')
        query = query.lower()
        return query
    except Exception as e:
        print(e)
        return listen()

object=""
def main():
    classNames,encodeListKnown=l.img_list()
    cap = cv2.VideoCapture(0)
    while True:
        success, img = cap.read()
        face_locations = face_recognition.face_locations(img)
        print(face_locations)
        if len(face_locations)>0:
            img,isuser=l.face_main(success, img ,classNames,encodeListKnown)
```

```
            while isuser:
                question=listen()
                print(question)
                if question.lower() in ph.phrases:
                    i=5
                    while i>0:
                        success, img = cap.read()
                        object=ob.detect_and_find_middle_object(img)
                        i-=1
                    response=" Object is "+object[:-6]
                elif question in ph.phrases_about_object:
                    response=cb.ask_chatbot(question+object[:-6])
                else:
                    response=cb.ask_chatbot(question)
                print(response)
                speak(response)
                success, img = cap.read()
                face_locations = face_recognition.face_locations(img)
                print(face_locations)
                img,isuser=l.face_main(success, img ,classNames,encodeListKnown)
                cv2.imshow('Video', img)
            else:
                # v.handle_visitor(img,face_locations)
                v.handle_visitor(img, face_locations)
main()
```

5.1.2 Face Recognition

```
import cv2
import numpy as np
import face_recognition
import os
import shutil
from datetime import datetime
import datetime


def get_time_of_day():
    current_hour = datetime.datetime.now().hour

    if current_hour < 12:
        time_of_day = "Morning"
    elif current_hour < 17:
        time_of_day = "Afternoon"
    else:
        time_of_day = "Evening"

    return time_of_day

def capture_new_face(img,success,counter,name):
    #take image of new unknown faces
    capture_and_save_image(f'{counter}.jpg',img,success)
```

```
    # move the image to the correct place i.e images folder
    move_file_to_folder(f'{counter}.jpg', 'images')

    rename_file('images/'+f'{counter}.jpg',f'{name}.jpg')
    # encode the new face
    classNames,encodeListKnown=img_list()
    counter+=1
    return classNames,encodeListKnown

def findEncodings(images):
    try:
        encodeList = []
        for img in images:
            img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
            encode = face_recognition.face_encodings(img)[0]
            encodeList.append(encode)
        return encodeList
    except Exception:
        pass


def img_list():
    path = 'images'
    images = []
    classNames = []
    myList = os.listdir(path)
    print(myList)
    for cl in myList:
        curImg = cv2.imread(f'{path}/{cl}')
        images.append(curImg)
        classNames.append(os.path.splitext(cl)[0])
    print(classNames)

    encodeListKnown = findEncodings(images)
    print('Encoding Complete')
    return classNames,encodeListKnown

def f_recognize(img,imgS,encodeList,classNames,recognized_faces=[]):
    recognized_faces=set(recognized_faces)
    facesCurFrame = face_recognition.face_locations(imgS)
    encodesCurFrame = face_recognition.face_encodings(imgS,facesCurFrame)
    for encodeFace,faceLoc in zip(encodesCurFrame,facesCurFrame):
        matches = face_recognition.compare_faces(encodeList,encodeFace)
        faceDis = face_recognition.face_distance(encodeList,encodeFace)
        #print(faceDis)
        if len(faceDis)>0:
            matchIndex = np.argmin(faceDis)
            if matches[matchIndex]:
                name = classNames[matchIndex].upper()
                #print(name)
                recognized_faces.add(name) #add the recognized faces to the set
```

```
                y1,x2,y2,x1 = faceLoc
                y1, x2, y2, x1 = y1*4,x2*4,y2*4,x1*4
                cv2.rectangle(img,(x1,y1),(x2,y2),(0,255,0),2)
                cv2.rectangle(img,(x1,y2-35),(x2,y2),(0,255,0),cv2.FILLED)
                cv2.putText(img,name,(x1+6,y2-
6),cv2.FONT_HERSHEY_COMPLEX,1,(255,255,255),2)
                # markAttendance(name)

    return list(recognized_faces)

def face_main(success, img, classNames, encodeListKnown):
    imgS = cv2.resize(img,(0,0),None,0.25,0.25)
    imgS = cv2.cvtColor(imgS, cv2.COLOR_BGR2RGB)
    face_locations = face_recognition.face_locations(imgS)
    result=[]
    if f_recognize(img,imgS,encodeListKnown,classNames)!=None:
        result=f_recognize(img,imgS,encodeListKnown,classNames)
    return img,len(result)>0
```

5.1.3 Object Detection

```
import cv2
import numpy as np
import torch
from pathlib import Path

# Set the path to the yolov5 repository
yolov5_path = Path("yolov5")

# Load the YOLOv5 model
model = torch.hub.load(str(yolov5_path), 'custom', path='yolov5s.pt',
source='local')

class_names = model.names
label=None
# Function to perform object detection and find the object in the middle
def detect_and_find_middle_object(frame):
    # Perform object detection with YOLO
    results = model(frame)

    # Get detection results
    pred = results.pred[0]

    if len(pred) > 0:
        # Find the object with the highest confidence score
        max_confidence_idx = np.argmax(pred[:, 4])
        middle_object = pred[max_confidence_idx]

        # Extract object information
        x_center, y_center, width, height, confidence, class_id = middle_object

        # Calculate object coordinates
```

```
        x1, y1, x2, y2 = int(x_center - width / 2), int(y_center - height / 2),
int(x_center + width / 2), int(y_center + height / 2)

        # Draw a bounding box around the object
        cv2.rectangle(frame, (x1, y1), (x2, y2), (0, 255, 0), 2)

        # Display the class name and confidence score
        label = f'{class_names[int(class_id)]}: {confidence:.2f}'
        cv2.putText(frame, label, (x1, y1 - 10), cv2.FONT_HERSHEY_SIMPLEX, 0.5,
(0, 255, 0), 2)
    else:
        label=""

    return label
```

5.1.4 Visitor Tracking System

```
import cv2
import face_recognition
import smtplib
from email.mime.text import MIMEText
from email.mime.multipart import MIMEMultipart
from email.mime.image import MIMEImage
import pyttsx3
import speech_recognition as sr
import openpyxl
from datetime import datetime, timedelta

import pandas as pd
import os
import shutil
import datetime

import face_recognize as l

engine = pyttsx3.init()
voices = engine.getProperty('voices')
engine.setProperty('voice', voices[1].id)

def speak(audio):
    engine.say(audio)
    engine.runAndWait()

def listen():
    print("listening...")
    r = sr.Recognizer()
    with sr.Microphone() as source:
        r.adjust_for_ambient_noise(source)
        audio = r.listen(source)
    try:
        query = r.recognize_google(audio, language='en-in')
        query = query.lower()
```

```
        return query
    except Exception as e:
        print(e)
        return listen()

def rename_file(file_path, new_name):
    # Get the directory and current name of the file
    directory = os.path.dirname(file_path)
    current_name = os.path.basename(file_path)
    # Construct the new file path with the new name
    new_file_path = os.path.join(directory, new_name)
    try:
        # Rename the file
        os.rename(file_path, new_file_path)
        print("File renamed successfully!")
    except OSError as e:
        print(f"Error occurred while renaming the file: {e}")

def move_file_to_folder(file_path, destination_folder):
    # Specify the destination path for the file
    destination_path = os.path.join(destination_folder,
os.path.basename(file_path))
    # Move the file to the destination folder
    shutil.move(file_path, destination_path)
    print("file moved to destination")

def capture_and_save_image(image_name,img,success):
    if success:
        # Save the image with the custom name
        cv2.imwrite(image_name, img)
        print("new face saved successfully!")
    else:
        print("Failed to capture image from webcam")

def capture_new_face(img,success,counter,name):
    #take image of new unknown faces
    capture_and_save_image(f'{counter}.jpg',img,success)
    # move the image to the correct place i.e images folder
    move_file_to_folder(f'{counter}.jpg', 'images')
    rename_file('un_images/'+f'{counter}.jpg',f'{name}.jpg')
    # encode the new face
    classNames,encodeListunKnown=l.img_list()
    counter+=1
    return classNames,encodeListunKnown

def img_list():
    path = 'un_images'
    images = []
    classNames = []
    myList = os.listdir(path)
```

```
    print(myList)
    for cl in myList:
        curImg = cv2.imread(f'{path}/{cl}')
        images.append(curImg)
        classNames.append(os.path.splitext(cl)[0])
    print(classNames)
    encodeListunKnown = l.findEncodings(images)
    print('Encoding Complete')
    return classNames,encodeListunKnown

def save_to_excel(visitor_information):
    # Load existing workbook or create a new one
    try:
        workbook = openpyxl.load_workbook('visitor_information.xlsx')
        sheet = workbook.active
    except FileNotFoundError:
        workbook = openpyxl.Workbook()
        sheet = workbook.active
        sheet.append(["Name", "Branch", "Purpose", "Timestamp"])
    # Add visitor information to the sheet
    for visit in visitor_information:
        sheet.append([visit['name'], visit['branch'], visit['purpose'],
visit['timestamp']])
    workbook.save('visitor_information.xlsx')

# Function to delete visitor information from the Excel sheet after 24 hours
def delete_old_records():
    try:
        workbook = openpyxl.load_workbook('visitor_information.xlsx')
        sheet = workbook.active
        # Iterate through rows and delete old records
        current_time = datetime.datetime.now()
        for row in sheet.iter_rows(min_row=2):  # Start from the second row
(header is the first row)
            timestamp = row[3].value  # Assuming timestamp is in the fourth column
            if isinstance(timestamp, datetime) and current_time - timestamp >
timedelta(days=1):
                sheet.delete_rows(row[0].row, amount=1)  # Delete the entire row
        # Save the workbook after deleting old records
        workbook.save('visitor_information.xlsx')
    except FileNotFoundError:
        pass  # No need to delete records if the file doesn't exist

# Email credentials
sender_email = '20R21A6656@mlrinstitutions.ac.in'
sender_password = '@Runpwr2712'
user_email = '20R21A6610@mlrinstitutions.ac.in'


def send_email(subject, body):
    msg = MIMEText(body)
```

```
    msg['Subject'] = subject
    msg['From'] = sender_email
    msg['To'] = user_email

    # Connect to the SMTP server
    with smtplib.SMTP('smtp.gmail.com', 587) as server:
        server.starttls()
        server.login(sender_email, sender_password)
        server.sendmail(sender_email, [user_email], msg.as_string())


# Function to send an email with an image attachment
def send_email_with_attachment(subject, body, image):
    msg = MIMEMultipart()
    msg['Subject'] = subject
    msg['From'] = sender_email
    msg['To'] = user_email

    # Attach text body
    msg.attach(MIMEText(body, 'plain'))

    # Attach image
    image_attachment = MIMEImage(image)
    image_attachment.add_header('Content-Disposition', 'attachment',
filename='visitor_image.jpg')
    msg.attach(image_attachment)

    # Connect to the SMTP server
    with smtplib.SMTP('smtp.gmail.com', 587) as server:
        server.starttls()
        server.login(sender_email, sender_password)
        server.sendmail(sender_email, [user_email], msg.as_string())

# Load user's image and encode the face
user_image = face_recognition.load_image_file("images/arun.jpg")  # Replace with
the user's image
user_encoding = face_recognition.face_encodings(user_image)[0]

# Initialize variables
visitor_information = []
counter = 0
# Function to handle visitor interaction
def handle_visitor(frame, face_locations):
    for (top, right, bottom, left) in face_locations:
        face_encoding = face_recognition.face_encodings(frame)[0]

        # Compare with the user's face
        results = face_recognition.compare_faces([user_encoding], face_encoding)

        if not any(results):
            # Visitor detected
```

```
            visitor_image = frame
            # visitor_image = frame[top:bottom, left:right]
            cap = cv2.VideoCapture(0)

            success, img = cap.read()


            imgg = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
            face_encodings = face_recognition.face_encodings(imgg)[0]

            # Collect visitor information (photo, name, branch, purpose of visit)
            speak(" Visitor detected")
            speak(" Tell Your Name")
            visitor_name = listen()
            # visitor_name = input("Enter visitor's name: ")
            speak(" Tell Your Branch")
            visitor_branch = listen()
            # visitor_branch = input("Enter visitor's branch: ")
            speak(" Tell your purpose of visit")
            visitor_purpose = listen()
            # visitor_purpose = input("Enter purpose of visit: ")

            # Save visitor information
            visitor_information.append({
                'name': visitor_name,
                'branch': visitor_branch,
                'purpose': visitor_purpose,
                # 'face_encodings':face_encodings,
                'timestamp': datetime.datetime.now().strftime('%Y-%m-%d %H:%M:%S')
            })

            # Send an email to the user with visitor information and image
attachment
            subject = "Visitor Detected!"
            body = f"Name: {visitor_name}\nBranch: {visitor_branch}\nPurpose:
{visitor_purpose}"
            send_email_with_attachment(subject, body, cv2.imencode('.jpg',
visitor_image)[1].tostring())
            save_to_excel(visitor_information)
            delete_old_records()
            return
```

CHAPTER 6

RESULTS

Infobot is an innovative chatbot robot equipped with state-of the-art objects and facial recognition features. It also acts as a personalized assistant with the known users. The accuracy of the face recognition system in identifying known and unknown users is up to 99.38%. The insight challenge faced during the recognition is the lighting conditions. So, the face should be visible to the camera of the robot. The unknown user details are collected through the robot by taking the speech or audio from the user such as name, branch, and purpose of their visit. Those details are mailed to the respective known users with their captured faces. When we consider it for use in the college or school then the robot collects information such as the name of the person, the Branch of a student, and the purpose of their visit.

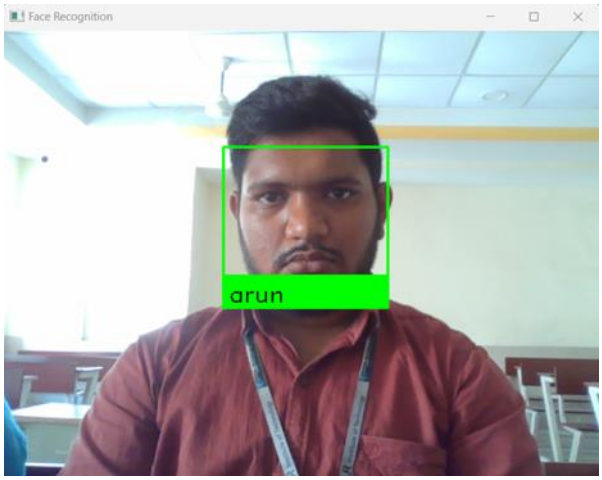


Figure 11: User Recognized as known

A visitor is detected and the information is sent to the known user. Here the unknown user's name is varsha, from the branch csm. The purpose of her visit is the submission of the record. In the same way, it is repeated with the others, and the infobot sends the details to the registered email ID i.e., the known user.

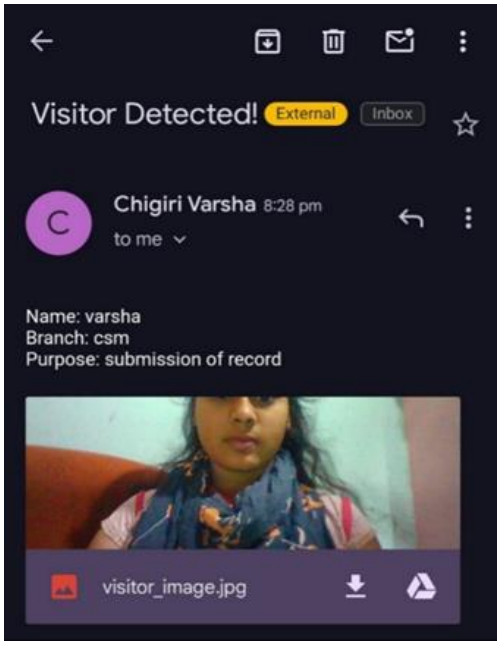


Figure 12: Visitor Detection main to User

In the frame, we observe a person, as his face is already registered, the infobot identifies his face through face recognition

	A	B	C	D
1	Name	Branch	Purpose	Timestamp
2	varsha	csm	record submission	2024-03-11 11:39:34
3	sai	csm	leave permission	2024-03-13 12.15.45
4	krishna	csm	hackathon permission	2024-03-15 10.30.12
5				
6				
7				
8				
9				
10				

Figure 13: Information Storing in Excel Sheet

For known users, additionally, it helps in detecting the objects. The objects are identified if the user specifically asks related to the objects by showing the object to the camera. The output of these is audio/speech.

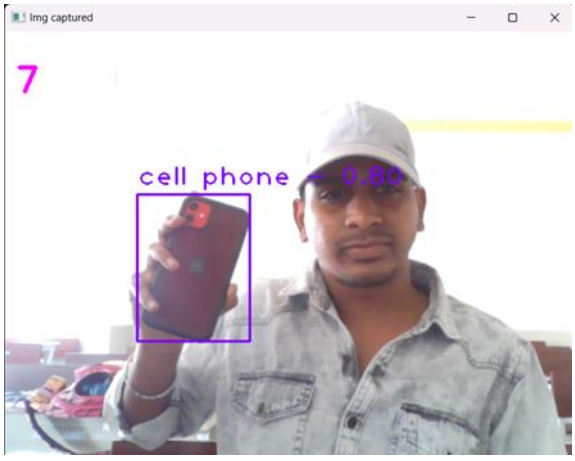


Figure 14: Object Detection and recognition

As Infobot already identifies the person as the known user it helps in object detection. Whenever the person asks for the robot a" detect the object in my hand" the object detection module helps in identifying the object and also provides the information regarding the object.

Figure 15: Infobot feature activated

CHAPTER 7

CONCLUSION

In conclusion, our project represents a significant advancement in robotics and artificial intelligence, focusing on enhancing user experience, security, and communication capabilities. The integration of a sophisticated facial recognition module allows the robot to identify specific users accurately, providing a personalized and tailored experience by retrieving relevant data from its database. This not only streamlines interactions but also strengthens security measures by verifying user identities. Moreover, the project addresses the challenge of engaging with unknown guests by leveraging facial recognition technology to initiate conversations and gather crucial information about visitors and their objectives. This data contributes to a robust visitor tracking system, enabling efficient monitoring and management of visitor interactions. The implementation of deep learning models, Raspberry Pi-compatible cameras, and neural networks underscores our commitment to leveraging cutting-edge technology for practical applications. The seamless integration of these components ensures smooth operation and accurate identification, enhancing the overall functionality and performance of the robot. Overall, our project represents a comprehensive solution for intelligent interaction, security enhancement, and data-driven decision-making. By combining innovative technologies with user-centric design principles, we aim to redefine the capabilities of robotic systems in providing personalized, secure, and efficient services in various domains, from home automation to hospitality and beyond.

FUTURE ENHANCEMENTS AND DISCUSSIONS

Moving forward, there are several avenues for enhancing our project to further elevate its capabilities and impact. One key area of focus is the refinement of the facial recognition module to improve accuracy and efficiency in identifying users, especially in diverse lighting conditions and with varying facial expressions. This could involve integrating advanced algorithms or exploring novel techniques in computer vision to enhance recognition performance. Additionally, enhancing the visitor tracking system by incorporating machine learning algorithms for predictive analytics could provide valuable insights into visitor behavior and preferences. This data-driven approach can inform decision-making processes and enable the robot to anticipate and cater to visitor needs more effectively. Furthermore, expanding the functionality of the robot to include natural language processing (NLP) capabilities would enable more sophisticated and context-aware conversations with users and guests. Integrating NLP technologies would allow the robot to understand and respond to complex queries, enhancing its utility in a wide range of scenarios. Moreover, exploring the integration of IoT (Internet of Things) devices and sensors could enhance

the robot's environmental awareness and enable it to perform tasks such as controlling smart home devices or monitoring environmental conditions. In terms of discussions, ongoing dialogue and collaboration with experts in robotics, AI, and human-computer interaction are crucial for refining the project's design, functionality, and user experience. Ethical considerations, such as data privacy and security, should also be prioritized to ensure responsible and transparent implementation of the technology.