Infobot- A Robot that recognizes Objects/Faces and extracts information about them

Mr. Vijay Gopal Jagadam
Associate Professor
MLR Institute of Technology
Telangana, India
vijayjagadam@gmail.com

Mr. Sai Prasad Kashi
Professor
MLR Institute of Technology
Telangana, India
saiprasad.kashi@mlrinstitutions.ac.in

Chigiri Varsha
Computer Science Artificial Intelligence and Machine Learning
MLR Institute of Technology
Telangana, India
varshapoornachandar@gmail.com

V Arun Kumar
Computer Science Artificial Intelligence and Machine Learning
MLR Institute of Technology
Telangana, India
arunpwr2712@gmail.com

Matam Jahnavi
Computer Science Artificial Intelligence and Machine Learning
MLR Institute of Technology
Telangana, India
matams1108@gmail.com

S Sai Krishna Reddy
Computer Science Artificial Intelligence and Machine Learning
MLR Institute of Technology
Telangana, India
surukuntisaikrishnareddy@gmail.com

ABSTRACT

The purpose of this paper is to introduce the development of a robot called 'Infobot'. Infobot is an advanced chatbot with cutting-edge objects and human face recognition capabilities. The primary objective of Infobot is to smoothly recognize certain items and faces, as well as collect important data that will enable people to communicate on personal grounds. Therefore, it serves as an "infobot" for familiar individuals or faces, surpassing regular chatbots by delivering personal information and help to known users. Conversely, Infobot can also identify strangers' faces to aid in obtaining some useful information about their purposes and whom they are meeting. Our project taps into contemporary AI and computer vision technologies for creating user-friendly smart robots capable of improving interactions between humans and machines in various settings.

KEYWORDS

AI, Chatbot, Computer Vision, Face Recognition, Human-Robot Interaction, Infobot, Object Recognition.

1. INTRODUCTION

Infobot stands out as a robotic entity equipped with advanced object and facial recognition functionalities, aiming to redefine the development of personalized chatbots for enriched human-robot interactions. Users can seamlessly navigate through the perception of objects and faces, extracting relevant information with precision. What sets Infobot apart from traditional chatbots is its advanced capability to identify unknown faces among visitors and extract detailed information regarding their purpose of visit. This unique feature empowers the robot to compile comprehensive information about both familiar and unfamiliar individuals, providing tailored insights into each visitors intent.

Infobot represents cutting-edge advances in robotics, offering sophisticated object and facial recognition capabilities that elevate the landscape of personalized chatbots and human-robot interactions. Unlike conventional chatbots, Infobot provides users with an intuitive experience characterized by smooth navigation of objects and face detection areas, enabling accurate extraction of relevant information.

One of the main differentiating features of Infobot is its unparalleled ability to identify visitors from unfamiliar faces and extract complex details about the purpose of their visit. This innovative feature allows the bot to build in-depth profiles of both known and unknown individuals, providing personalized insight into the purpose of each visitor's

presence. Finally, Infobot overcomes the limitations of traditional chatbots by leveraging advanced technology beyond recognition—faces, but also the motives of each meeting. By providing personalized insights and facilitating enriched human-robot interactions, Infobot heralds a new era in robotics where the seamless integration of advanced features increases the quality and depth of the human-machine connection.

The base of this project lies in utilizing edge-cutting technologies, by using superior facial recognition algorithms and real-time processing, Infobot is a continuing and shrewd human-robotic interplay. we are using AI and computer vision, to develop a user-friendly and intelligent robot. To personalize assistance, enhance face and object recognition, and create friendly conversation will be modified to lead to altered scenarios by Infobot. We are going to explore the techniques in detail by inspecting their first-rate packages, technological components, and implications. In this, we're capable of contributing to the growing literature on artificial intelligence and human-robotic relationships with a watch on clever robots

2. BRIEF INTRODUCTION

A. FACE RECOGNITION

Facial identification is the technology that is used to identify or verify a person from an image or video. It may be that facial recognition systems operate by comparing chosen facial features on one image with faces in a database.

It can also be called a Biometric Artificial intelligence-based application that can distinguish somebody uniquely through analysis of patterns predicated on the face's texture and shape.

It is often described as a process that first involves four steps: face detection, face alignment, feature extraction, and finally face recognition.

Face Detection. Find one or more faces in the picture and indicate them with a bounding box.

Face Alignment. Normalize the face to match with the database for example geometry and photometrics.

Feature Extraction. Extract features of a face that can serve for recognition purposes.

Face Recognition. Match this face against any other known faces located in the already prepared database.

In some cases, such as traditionally where each step had a separate module or program in some systems while others may combine several or all of these steps into one single process.

One of the most high-profile and common 'machine learning' techniques that have been used for face recognition was described in a paper published in 1991 called "Face Recognition Using Eigenfaces". The technique is known as "Eigenfaces" and marked a milestone in simple holistic methods as they proved that such approaches could produce impressive results

B. OBJECT DETECTION

Object Detection is a technology of computers that are under computer vision, image processing, and deep learning which is concerned with finding instances of objects in images or videos. Object detection on the other hand is a computer vision technique that aims to identify and locate objects within images or videos. It not only detects categories like cars, dogs, and people but it also recognizes their spatial positions inside the scene by using bounding boxes. Different applications require object detection such as autonomous vehicles, security systems, and facial recognition among others. Photos are identified using convolutional neural networks (CNNs) in object recognition based on deep learning. These techniques stand out best and can detect multiple objects as well as their classes simultaneously. One of these is the Single Shot MultiBox Detector (SSD), which offers a compromise between accuracy and speed among popular deep learning-based models used for object detection. How Learning Works in Object Detection: Feature Extraction: The CNN extracts features from the input image. This network learns how to capture hierarchical features starting from low-level patterns like edges up to high-level semantics such as object parts. Region Proposals: The network suggests possible regions containing objects in an image frame

C. CHATGPT

With an exciting possibility of human-robot interaction, artificial intelligence, and robotics integrate language models like ChatGPT. This paper explores distinctive ways of commanding robots to be directed utilizing ChatGPT through voice commands as a result having more instinctive communication amid machines and people. ChatGPT is a state-of-the-art language model developed by OpenAI using GPT-3.5 architecture. As it has been trained on diverse data sources it can comprehend and produce natural language-like texts applicable in many areas. With such an NLP system behind it, ChatGPT can understand user input almost like a real conversation between two humans.

A significant advantage of using ChatGPT is the ease with which it converts instructions in human languages into machine or robot-friendly commands. By connecting the control system of the robot with ChatGPT, users can provide orders through generalized language used regularly during chats. The integration of ChatGPT and the control system of a robot provides an opportunity for users to communicate with the machine in plain language. Giving commands to the robot, such as "go forward," "turn left," or "get item," is similar to talking to a human. A far easier approach to conversational programming than traditional coding and the use of quite complex interfaces is provided by ChatGPT. To direct the robot to do tasks, users specify desired actions, and ChatGPT uses these descriptions to generate executable robot instructions. Again, ChatGPT's ability to process new data allows for continuous improvements in robot control abilities. Chat-GPT improves its comprehension of instructions over time to enable more individualized and effective conversations by engaging with users and gathering feedback. In this way, through continued interactions, however, chat-GPT would understand user preferences and provide better responses based on them. The adaptability ensures that the robotic behavior matches the expectations of a user leading to a more intuitive and user-friendly interface.

3. EXISTING PROBLEMS

Current challenges for chatbots are aimed at achieving real human likeness that can be seen through a full understanding of the context given. Among these issues are errors that arise due to the lack of contextual knowledge and disruptions in conversation flow. In most cases, emotional intelligence is absent, so a chatbot may not properly understand what the user feels and what has caused his or her problem. The issue also arises from the use of ambiguous questions, which lead to problems when trying to find correct answers. The level of intelligence, capacity to adapt to changes, speed, and precision in response are constraints that influence any system. Before implementing a chatbot application, prejudice and privacy concerns must also be taken into account. Finally, data might become challenging if extensive information is required due to the complexity of compatibility with other systems for data retrieval. Building a rapport with the user and providing visibility into the bot's capabilities are both crucial. Among the areas of improvement for a more efficient and naturalistic chatbot conversation, multi-turn exchanges and addressing language and cultural gaps continue to emerge as frontiers. Robotic facial recognition faces significant challenges. It can be inaccurate because different lighting conditions, facial expressions, and angles can cause false positives. The sensitivity of facial data raises concerns about privacy, data security, and unauthorized access. The careful consideration of the widespread implementation of facial recognition technology should also take into account the ethical consequences it brings, such as the possibility of abuse and

invasion of privacy. Algorithms that exhibit racial and cultural biases have the potential to disproportionately and unjustly penalize specific groups. To develop and apply robotic facial recognition skills ethically, it is necessary to reconcile technological advancements with ethical considerations. Computer vision object identification has several difficulties. The difficulty of accurately identifying objects in a range of situations, including changes in illumination, occlusion, and viewing angles, is one major problem. Achieving real-time processing can be challenging, especially when processing high-resolution images or video streams. Anomalies such as false positives and false negatives are still common and affect the reliability of detection systems. Scalability issues arise when trying to handle a large number of objects or classes. Training robust models requires large labeled datasets, and a lack of diversity in training data can lead to performance bias and affect the generalization ability of the system. Numerous issues affect visitor monitoring systems' acceptability and effectiveness. Privacy considerations are important because persistent tracking might lead to ethical dilemmas and unease in the monitored person. Accuracy problems, especially in crowded environments, can result in misidentifications or incomplete tracking. Handling diverse entry and exit points presents a challenge for maintaining a comprehensive tracking system. Integration complexities with other security systems and databases can hinder seamless operation. Authentication mechanisms must be strong to prevent unauthorized access and identity theft. Striking the right balance between security measures and individual privacy remains a persistent challenge for the successful implementation and widespread acceptance of visitor tracking systems.

4. IMPLEMENTATION

The solution is achieved by integrating face recognition, object detection, a Visitor tracking system, and ChatGPT modules.
The process of integrating modules follows different steps.

4.1 Face recognition

In Python, face recognition typically involves using OpenCV for face detection using Haar cascade classifiers to identify facial features like eyes, nose, and mouth, resulting in bounding boxes around detected faces. Optionally, pre-trained models for facial landmark detection can pinpoint specific points on faces for tasks like alignment or advanced recognition. After detection and optional landmark identification, deep learning models extract facial encodings, numerical representations capturing unique face characteristics. Recognition involves comparing these encodings to known individuals' encodings using a distance metric, such as Euclidean distance, with a threshold for confident matches. Libraries like face_recognition simplify this process by abstracting internal workings, but accuracy relies on factors like lighting, pose, and image quality, following a workflow of model loading, face detection, landmark detection, encoding extraction, and comparison for recognition.
Acquisition and preprocessing: As the robot can perform the process in runtime before the process, the visual input is captured through a camera needed so the video frames can be preprocessed.
Face detection: The training of the system is done on visual data thus faces have to be detected in live video streaming shots so OpenCV face recognition's Python module is used by the system.
Feature extraction: After the recognition, the feature extraction is done of the surroundings of an unknown person by capturing the example of an eye, nose, mouth, and forehead for feature extraction.
Feature matching: The identification of known and obscured faces as the features are extracted from the captured face. From features of the captured face and comparing among with the facial features already present in the database guarantees that the system will recognize the known faces

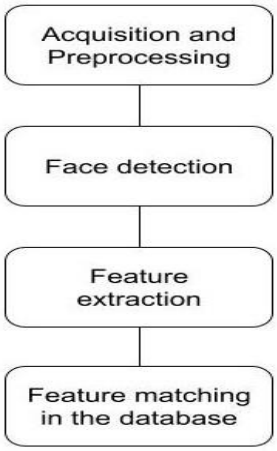

Fig 1: Steps followed in face recognition

A: Algorithm of face recognition

```
frames=videocapture()
faces=detect_the_faces(frames)
if faces
do
features=extract_face_encodings(frames)
user=face_recognition.compare(features,user_features)
return (user)? True: False
end
```

4.2 Object Detection

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

Feature extraction: the next step is about the processing and extraction of a set of attributes from each frame feature set, which allows the determining of the content of the frames. These elements might happen in the context of the objects such as their shape, color as well as texture, or their spatial relationships.
Bounding boxes: the extracted features are applied to locate and obtain the coordinates of diverse objects in the image and to inform the robot about the bounding boxes around them. A rectangular boundary box can be drawn that will show where the position and size of the object in the image. They are called boxes because they enclose objects identified with them and also with the surroundings.
Classification: the labeling step on the other hand consists of choosing the category that represents the bottle or choosing its class assuming there is a camera image that contains an object. Classification is attaching a label or a category to each box depending on the visual features of this bounding box.
Output: the culmination of it is going to be the necessity of demonstrating the results to the user.

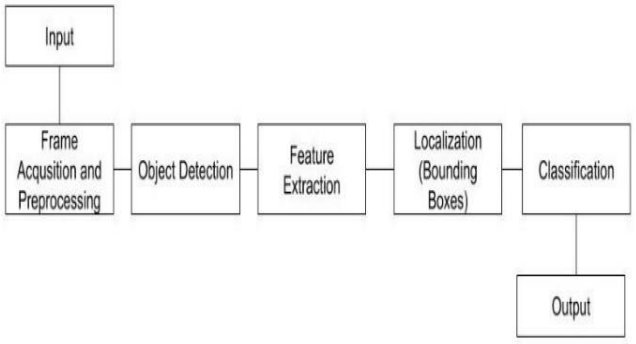

Fig 2: Flow diagram of Object Detection

B: Algorithm of object detection

```
frames=videocapture()
model=load(Yolov5)
object=model.predict_middle_object(frame)
return object?object_label:None
```

4.3 GPT Chatbot

While calling ChatGPT via an API, you first ship an API request to OpenAI with a set of non-obligatory parameters like temperature, max tokens, and pinnacle p sampling to influence the reaction's creativity and period. OpenAI preprocesses your request, together with responsibilities like textual content cleansing and tokenization, earlier than selecting the best ChatGPT version variant primarily based on elements like set-off complexity and request parameters. the selected model, possibly based on transformers, then strategies the set-off, is aware of its context and generates a reaction following the prompt's cause. put up-processing steps inclusive of filtering and blunders correction may be 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.
Question: A question or query from the individual is the enter. this is inside the form of voice entry. The question serves as the preliminary entry that the chatbot will use and respond to.
Similarity: After receiving the purchaser's question, the chatbot may additionally carry out a similarity search to find comparable questions or queries in its understanding base or database. This step lets the chatbot discover applicable statistics or responses that have been furnished in the past for comparable inquiries.
Related records: once similar questions or queries are identified through the similarity are attempting to find, the chatbot retrieves related facts or responses associated with those queries. This data might also moreover include relevant articles, documents, FAQs, or formerly gen excellent responses that cope with similar subjects or questions.
ChatGPT: The chatbot uses ChatGPT, an advanced language version, to create responses. ChatGPT draws on its first-rate information of language to generate extremely good coherent and applicable replies based on the customer's question and the records collected earlier.
Output: The generated reaction from ChatGPT, alongside any more relevant statistics retrieved in advance steps, is furnished to the character due to the fact the output of the chatbot. This output is in the form of speech synthesized via a voice synthesizer appropriate for verbal exchange with the client.

C: Algorithm for the chatbot

```
function chatbot(question)
bot=chatgpt()
response=bot.get_result(question)
return response
```

4.4 Visitor Tracking System

Input: The non-prevent video that is captured through the digital digicam of the robotic is considered because of the entry. The video frames are preprocessed for the identity of the faces.
Enhancement: The technique of improving the video or an image is called enhancement. This is completed to beautify the remarkable of the video to understand the faces. Noise reduction, polishing, and Normalization are the enhancement techniques.
Face Detection: With the usage of the Face reputation module the faces are detected and the capabilities are extracted from the detected faces.
Face popularity: The detected faces are diagnosed with the use of the Python face reputation module. that is finished via evaluating the detected face to a database of recognized faces.
Now via the detected featured face, the robotic classifies them as recounted or unknown customers.
Notify individual: at the same time as there is an unknown person, it collects the records and notifies the recognized man or woman via electronic mail or maybe via the use of informing verbally the stated person when they go to again.

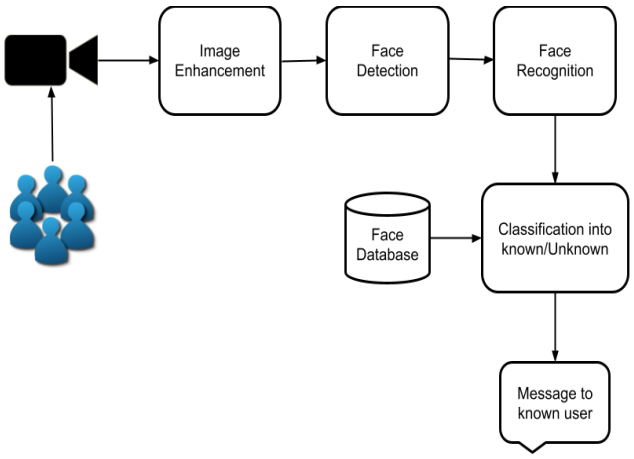

Fig 3: Flow chart of Visitor tracking system

```
C: Algorithm Visitor Tracking System
function Infobot():
frames=videocapture()
user=face_recognition(frames)
if (user)
do
query=listen()
 if "detect The object" in query
do
object=object_detection_and_recognition()
end
else if "tell me about the object" in query
do
response=chatbot("tell me about the "+object)
end
else
 do
 response=chatbot(query)
 end
 speak(response)
 end
```

4.5 Design of Robot

Fig 4.5.1: Hand of Humanoid robot

The humanoid robot's hands include several degrees of flexibility in their articulation which allows for deft manipulation and gripping replicating human movement and object contact the hand is powered by motors or servos and equipped with feedback sensors its open-source construction modular design and customization enable accessibility and further advancement in the robotics community.

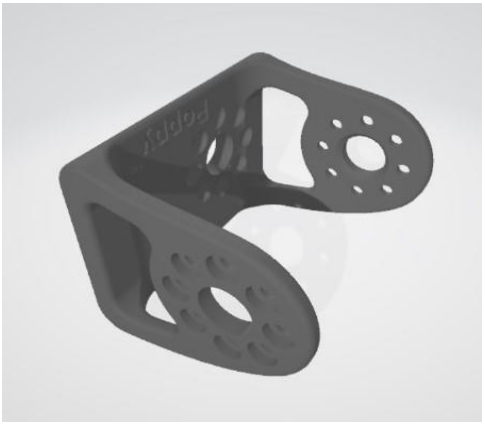

Fig 4.5.2: Shoulder of Humanoid robot

The shoulder of the Humanoid robot incorporates numerous degrees of freedom joints driven by actuators allowing for a wide range of upper body movements it is intended to give stability and flexibility as well as feedback for precise motions through sensor integration this modular component is easily adaptable and integrates effortlessly with the robot s control system allowing for natural and coordinated interactions.

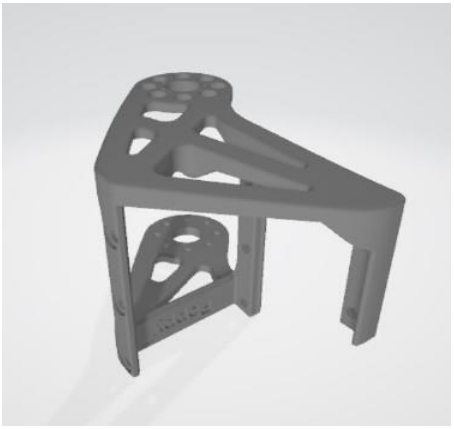

Fig 4.5.3: Hip of Humanoid robot

The hip of the humanoid robot is a pivotal joint connecting the torso to the legs allowing for dynamic movement and stability it features multiple degrees of freedom powered by actuators and integrates sensors for feedback modular and customizable it seamlessly integrates into the robot's control system for coordinated motion during locomotion and tasks

Fig 4.5.4: Foot of Humanoid robot

The foot of the humanoid robot is a complex structure designed to mimic the functionality of a human foot it typically includes multiple joints for flexibility sensors balance actuators for movement and compliance features for stability the design aims to facilitate natural and efficient locomotion allowing the robot to walk and navigate various terrains with ease.

Fig 4.5.5: Head of the Humanoid robot

The head of the humanoid robot is equipped with cameras microphones sensors and processing units for perception interaction and control it enables the robot to see hear and process information from its environment while also facilitating communication with users through speech recognition and synthesis additionally it may feature actuators for expressive movements and communication interfaces for connectivity with external devices and networks.

Fig 4.5.6: Chest of Infobot

The chest of the humanoid robot is a vital detail housing critical hardware and systems It typically consists of manage structures for interpreting instructions and coordinating moves, strength sources for electricity, sensors for perceiving the surroundings, actuators for motion manipulation, communication components for interplay, and protection capabilities for dependable operation. This compact area serves as the nexus of the robotic's functionality, allowing it to interact intelligently and correctly with its surroundings**.**

Fig 4.5.7: MX-28 servo motor

The mx 28 servo motor is a high-performance actuator commonly used in robotics and automation applications manufactured by the robot it offers precision control and high torque output making it suitable for a wide range of tasks from robotic manipulators to humanoid robots with durable construction and built-in sensors for feedback the mx 28 provides accurate position control and smooth motion contributing to the efficiency and reliability of robotic systems.

5. SOLUTION FOR THE PROBLEM

Our assignment is a ground-breaking aggregate of human-targeted format and cutting-edge-day technology, resulting in a robot with notable interactive and conversational talents. pushed via GPT, a complex language version, our tool is prepared to convert human-robotic interactions through way of fusing computational energy with intelligence, flexibility, and understanding. Our goal has continuously been to carry collectively a robotic that might speak obviously and meaningfully, however additionally has the highbrow capability to select new abilities over time and alter to precise situations, making every interplay precise.

The robot's conversational technique is powered through manner of GPT, a complex language version designed to have interaction in natural and human-like conversations. Leveraging GPT's computational techniques enhances conversational skills, making interactions greater fluid and huge. From its inception, our motive has been to create a robot that no longer best engages in herbal and considerable conversations but moreover possesses the cognitive capabilities to test, adapt, and personalize interactions over the years. The chatbot goes with the waft is examined in Fig 1.

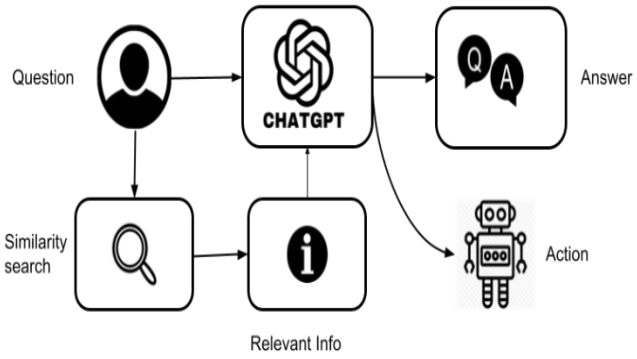

Fig 1. GPT chatbot workflow

One unique feature of the tool is its reminiscence retention and attention, facilitated by way of GPT. It shall we the robot apprehend and take into account people, even in some unspecified time in the future of initial encounters. information about past interactions is stored, contributing to a greater personalized and context-conscious engagement. GPT allows the protection of pertinent information approximately client encounters, ensuring the entire data of the person's statistics. This data aids in tailoring responses and interactions based on the consumer's preferences and beyond engagements, all facilitated seamlessly via the GPT interface on the front end.

important functionalities including item detection and face reputation are included in the machine. item detection capabilities permit the robotic to emerge as privy to and interact with physical gadgets in its environment. right here first the video frames are taken from a digicam and then every frame is sent to the object detection model which uses YOLOv8 in the history. The YOLOv8 version runs on the idea of CNN (convolutional neural network). After processing the photograph frames the use of the YOLOv8 version, we get a listing of sophistication names in which all the detected objects are protected. The architecture of YOLOv8 can be visible.

Face reputation complements the robotic's ability to apprehend and consider people, fostering a more personalized experience. right here additionally the video frames are taken as input and each body is processed with the use of the face popularity model. It makes use of the face recognition library of Python and identifies the faces of people with the aid of comparing the functions of the current user's face with the faces in its database. After comparing it gives an output with a boolean cost. An included chatbot element is included to facilitate seamless verbal exchange. you can see how the face is
detected in Fig 2.

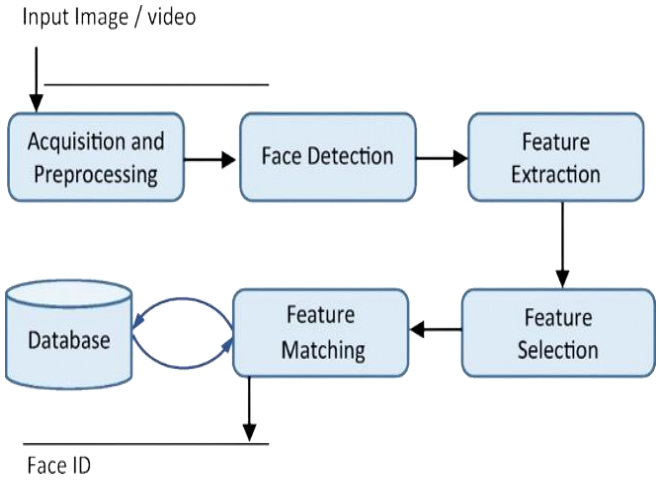

Fig 2. Face Recognition architecture

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. The architecture visitor tracking system can be seen in Fig 3.

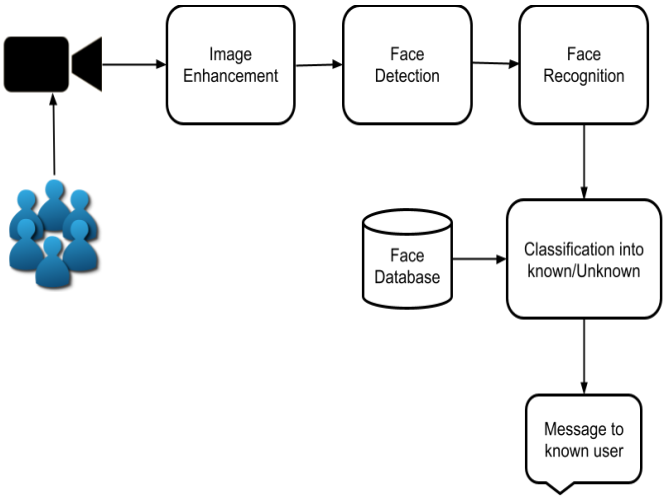

Fig 3. Visitor tracking system architecture

This chatbot enhances the robotic's conversational talents, permitting it to address an extensive variety of queries and interactions in significant dialogues with users. item detection and face reputation lay the foundation for more state-of-the-art interactions. The robot can reply to visual stimuli, pick out people, and interact with items, increasing its range of abilities. With its ability to apprehend individuals, interact with physical objects, and tune visitors, our machine heralds a new generation of personalized and context-aware robotics, poised to redefine the limits of human-robotic coexistence in numerous domain names. The infobot's architecture can be seen in Fig 3.

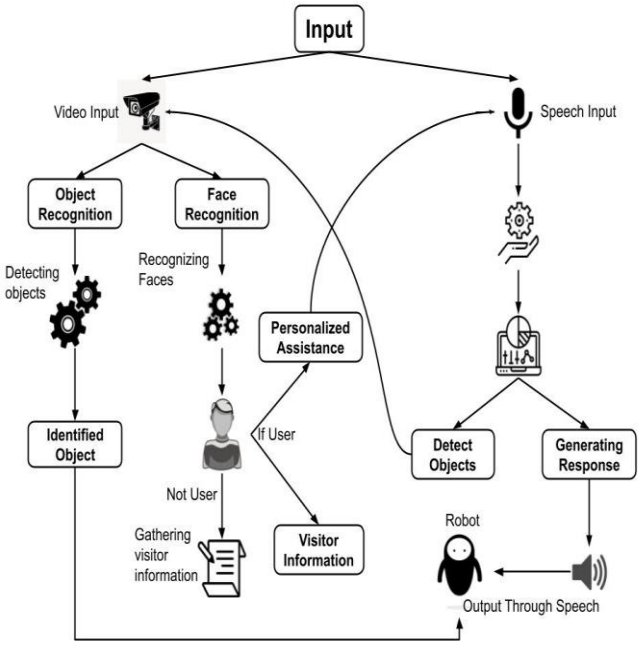

Fig 4. Infobot architecture

In summary, our mission embodies the convergence of innovation and practicality, showcasing the transformative potential of synthetic intelligence in improving human-robotic interactions. via the combination of GPT for conversation, YOLOv8 for item detection, and face reputation era, our robot stands as a testament to the electricity of interdisciplinary collaboration and technological advancement. With its ability to recognize individuals, engage with physical gadgets, and tune site visitors, our gadget heralds a brand-new technology of customized and context-aware robotics, poised to redefine the boundaries of human-robot coexistence in numerous domains.

6. RESULT ANALYSIS

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.

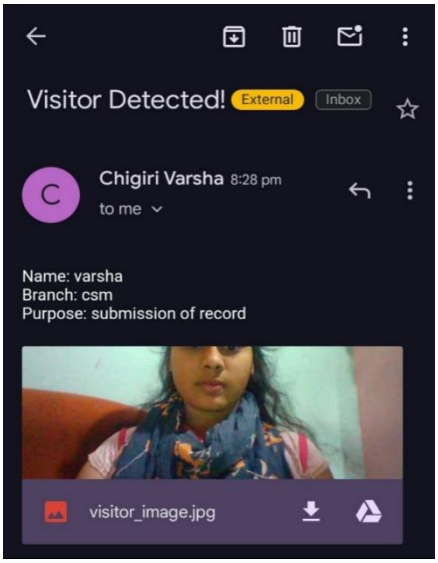

Fig 6.1 collected details from the unknown and notified to known user

	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				

Fig 6.2 visitor information stored in excel sheet

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.

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

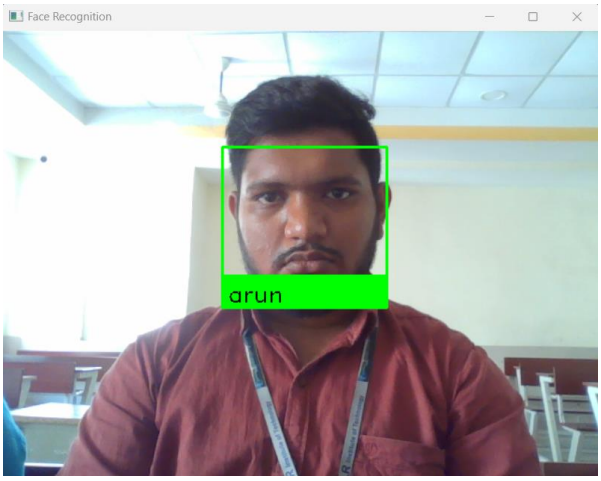

Fig 6.2 Face-recognized

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.

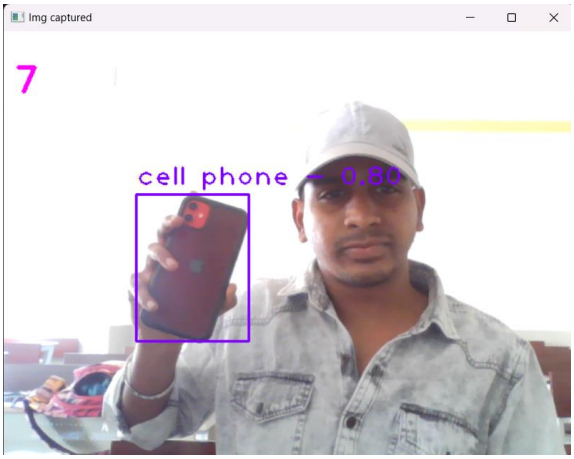

Fig 6.3 Object detection

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.

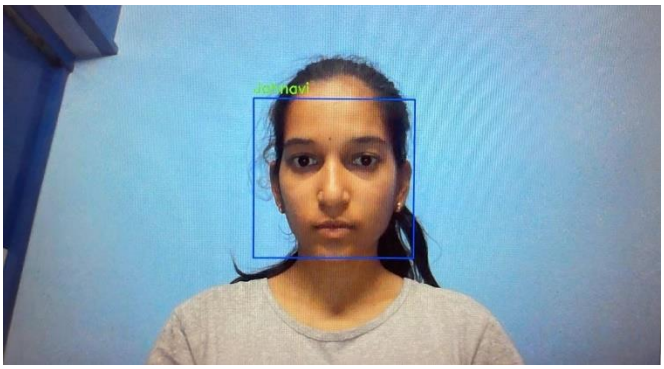

Fig 6.4 Recognized Face

To ensure object detection, users should prompt the system by asking, "What is this?" and then show the object. Without the user's prompt and presentation of the object, the system won't provide a detection output.
The robot acts as an infobot for registered users. It can be placed in the absence of the respective person to track visitors

7. CONCLUSION AND FUTURE SCOPE

Our vital purpose is to accumulate a smart, kind, sympathetic robot companion. we are hoping to acquire this through the use of concentrating our efforts on several areas, together with item identity, facial reputation, memory retention, and the aggregate of chatbot skills. we are confident that through taking a holistic technique, we're able to decorate the robot's abilities and flip it properly right into a useful and amiable assistant in a spread of settings. Our research responsibilities will redesign human-robot collaboration, ensuring customers a greater seamless and thrilling enjoyment. Our final cause is to create a complex chatbot that might understand words and images. Our cause is to push the boundaries of human-robot cooperation with the beneficial resource of tackling modern-day troubles in human-robotic interaction. This initiative can convert technology.

8. REFERENCES

[1] 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. 1816-1821, doi: 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: 0.1109/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.

[4] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.
[10] Koubaa, A. (2023). ROSGPT: Next-Generation Human Robot Interaction with ChatGPT and ROS Preprints.https://doi.org/10.20944/preprints202304.0827.v3
[11] 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
[12] Zhao X-F, Chen Z-H, Yin H-F, Wu X-J. Design of intelligent visitor system based on cloud and edge collaborative computing. Journal of Algorithms & Computational Technology. 2023;17. doi:10.1177/17483026231169154
[13] 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
[14] Y. Ye, H. You and J. Du, "Improved Trust in Human-Robot Collaboration with ChatGPT," in IEEE Access, vol. 11, pp.5574855754, 2023, doi: 10.1109/ACCESS.2023.3282111
[15] 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, doi: 10.1109/ICOIACT50329.2020.9332074
[16] 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
[17] H. Polat Et Al., "A bibliometric analysis of research on ChatGPT in education," International Journal of Technology in Education (IJTE), vol.7, no.1, pp.59-85, 2024
[18] 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
[19] Shafeeg, A., Shazhaev, I., Mihaylov, D., Tularov, A., & Shazhaev, I. (2023). Voice assistant integrated with chat gpt. Indonesian Journal of Computer Science, 12(1), DOI: https://doi.org/10.33022/ijcs.v12i1.3146
[20] 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.
[21] Stephen Balaban "Deep learning and face recognition: the state of the art", Proc. SPIE 9457, Biometric and Surveillance Technology for Human and Activity Identification XII, 94570B (15 May 2015); https://doi.org/10.1117/12.2181526
[22] L. Jiao et al., "A Survey of Deep Learning-Based Object Detection," in IEEE Access, vol. 7, pp. 128837-128868, 2019, doi: 10.1109/ACCESS.2019.2939201.
[23] Pal, S.K., Pramanik, A., Maiti, J. et al. Deep learning in multi-object detection and tracking: state of the art. Appl Intell 51, 6400–6429 (2021). https://doi.org/10.1007/s10489-021-02293-7
[24] R F Rahman and I D Sumitra 2020 IOP Conf. Ser.: Mater. Sci. Eng. 879 012062 DOI: 10.1088/1757-899X/879/1/012062
[25] L. Aziz, M. S. B. Haji Salam, U. U. Sheikh and S. Ayub, "Exploring Deep Learning-Based Architecture, Strategies, Applications and Current Trends in Generic Object Detection: A Comprehensive Review," in IEEE Access, vol. 8, pp. 170461-170495, 2020, doi: 10.1109/ACCESS.2020.3021508
[26] GAJARE, P., KADLAG, J., SALUNKE, M., & AWARE, V. (2019). VISITOR GUIDANCE AND TRACKING SYSTEM USING WIRELESS COMMUNICATION.
[27] Kadu, R. N., Wale, T., Lohale, S., & Sawant, S. VISITOR GUIDANCE AND TRACKING SYSTEM USING WIRELESS COMMUNICATION
[28] N. Wake, A. Kanehira, K. Sasabuchi, J. Takamatsu and K. Ikeuchi, "ChatGPT Empowered Long-Step Robot Control in Various Environments: A Case Application," in IEEE Access, vol. 11, pp. 95060-95078, 2023, doi: 10.1109/ACCESS.2023.3310935
[29] T. Wu et al., "A Brief Overview of ChatGPT: The History, Status Quo and Potential Future Development," in IEEE/CAA Journal of Automatica Sinica, vol. 10, no. 5, pp. 1122-1136, May 2023, doi: 10.1109/JAS.2023.123618.
[30] M. Khan, S. Chakraborty, R. Astya and S. Khepra, "Face Detection and Recognition Using OpenCV," 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, 2019, pp. 116-119, doi: 10.1109/ICCCIS48478.2019.8974493.