



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH


IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 5, May 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 ijirset@gmail.com

 www.ijirset.com

Smart Glasses with Blind Assistance System: Real-Time Object Detection with Voice Alerts

Jadhav Soham Ganesh, Dange Asmita Chandrakant, Bhujbal Bhavesh Gangadhar,
Walke Shivam Ramesh, Prof. Gunaware NJ

U.G. Student, Department of Computer Engineering, HSBPVT's GOI Faculty of Engineering, Kashti, Maharashtra, India

Professor, HSBPVT's GOI Faculty of Engineering, Kashti, Maharashtra, India

ABSTRACT: One prospective technological advancement to improve visually impaired people's mobility and freedom is the Blind Assistance System-equipped smart glasses. This research presents a unique technology that offers blind people comprehensive help by combining voice alarms with real-time object detection. The system is centered on a wearable gadget with a built-in camera, providing a discrete aid for individuals with vision impairments. Modern deep learning models are used by the real-time object identification component to recognize different items and barriers in the wearer's surroundings. Voice warnings are used to communicate this information; they are generated in real-time and sent to the user's earphones to minimize distractions. The goal of the Smart Glasses is to increase the user's safety and situational awareness so they can move more confidently in challenging environments. This technology not only increases the mobility of people with visual impairments but also helps them become more socially included by lowering their need on sighted assistance. The Blind Assistance System for Smart Glasses is a noteworthy advancement that shows how technology can be used to solve practical problems and enhance the quality of life for people who are blind or visually impaired. It also shows how the visually impaired community can be empowered, safer, and more independent.

KEYWORDS: Raspberry pi, Internet of Things (IoT), Object detection, YOL, Voice feedback, raspberry PI 5, wearable device, OpenCV, Tensor Flow Lite

I. INTRODUCTION

It is a well-known statistic that there are an estimated 285 million blind people in the world, or almost 20% of the Indian population. They have persistent and ongoing difficulties with navigation, particularly when they are by themselves. Most of the time, they rely on someone to even get the necessities for their daily existence. As a result, a technology solution is desperately needed for a difficult task. One attempt on our part was the development of an Integrated Machine Learning System that enables blind victims to recognize and categorize common everyday objects in real time, provide voice feedback, calculate distance, and generate warnings when they are approaching or far from the object. The same system can be used for Obstacle Detection Mechanism.[1] Wireless technology has been making tremendous progress over the past few years this report's focus, the Smart Glasses with Blind Assistance System, is an example of this kind of technological breakthrough. With the combination of voice alerts and real-time object recognition, this wearable gadget provides a complete and discrete solution to meet the specific requirements of those who are visually impaired. The system can identify a wide range of objects and obstacles in the user's surroundings by utilizing cutting-edge computer vision and deep learning techniques. Most importantly, the system communicates this information through voice alerts, guaranteeing that users are consistently informed about their surroundings in an understandable and clear manner.

There are numerous advantages to such a system. It encourages greater freedom in visually impaired people, allowing them to move more confidently in different environments. It promotes independence and social inclusion by lessening dependency on sighted help. The technology is a promising solution for improving safety and overall quality of life for individuals with visual impairments because to its real-time updates, immediate feedback, and diverse indoor and outdoor applicability.

II. RELATED WORK

In this segment, we examine research projects related to object detection and smart glasses. One of the best and most practical ways to support individuals with brain-related illnesses in moving around independently and carrying out every day personal duties is through wearable assistance devices. Smart glasses are being used in a wide range of disciplines, including education, industry, sports, computer science, social science, healthcare, and agriculture. We emphasize the positive aspects of blind and visually disable people in this literature review. [2] Recognizing the face and identify information of friends and family members is one of the most crucial and critical jobs for Blind and Visually disable individuals. They developed a facial recognition system that can interpret facial photos taken with a smart glass camera in response to user inputs, process the results on a server, and then provide an audible response. A pair of cellphones, smart glasses, and a back-end server are used in the client-server architecture of the system to do face recognition utilizing deep CNN models like FaceNet and Inception-ResNet. It will take longer for this face recognition system to operate since it must be retrained to identify new faces that are not stored on the server.[2] For Blind and Visually disable individuals, a completely automated smart glass system was created to help with object recognition at night. To the best of our knowledge, current smart glass systems are unable to manage issues during the night and do not assist walking in low light or noise environments. It uses real-time audio output to give users information about nearby items.

III. PROPOSED SYSTEM

The recommended and proposed system is able to identify any trained objects from the available data sets (or create our own) as well as the model that is generated following the training process. Typically, data sets are available online for use. Once there, we can apply any function to the data according to our requirements. Yolo serves as a database for object identification.[3]

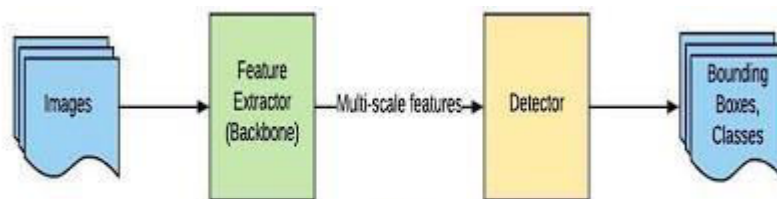


Figure 1: Proposed plan

Our mission is to give blind and visually disable people the ease and opportunity to move independently throughout the day and at night. The most efficient ways to accomplish this goal are wearing smart glasses and a multipurpose system that can take pictures with a tiny camera and provide users with speech feedback along with object recognition findings. According to the demands of blind individuals, it is also possible to perceive visual information by touching the contours of detected salient objects via a tactile display that may be refreshed. Deep CNNs must be used by the system to recognize objects with high accuracy, and a strong CPU must be used to complete the tasks quickly enough in real time. [2]

IV. PROBLEM DEFINITION

In recent years, there have been notable advancements in the creation of assistive devices for those with visual impairments. However, there are still significant issues and unfulfilled requirements that call for creative solutions. Developing a complete Smart Glasses with Blind Assistance System that combines audio alerts and real-time object identification in a seamless manner is one such problem. The main issues that this technology seeks to solve will be outlined in this section.

- 1) **Mobility and Independence:** The fundamental challenge faced by individuals with visual impairments is the limitation of mobility and dependence on others. Navigating through spaces, crossing streets, and identifying objects in the environment present daily hurdles.
- 2) **Safety Concerns:** Visually impaired individuals are at an elevated risk of accidents and collisions due to obstacles and objects they cannot perceive. The problem revolves around the need to enhance safety by providing real-time information about the proximity and identity of objects and obstacles in their path, reducing the risks associated with

navigation.

3) Customization and Adaptability: The needs of individuals with visual impairments can vary significantly. Therefore, the problem is to design a system that is adaptable to individual preferences and requirements, allowing users to customize settings, sensitivity, and the nature of voice alerts.

4) Real-Time Processing: Achieving real-time object detection and voice alerts poses a significant technical challenge. The system must process and analyse video footage, identify objects, estimate distances, and generate voice alerts with minimal latency. The problem is to ensure that the technology operates swiftly and accurately in real-world environments.

5) Ethical and Privacy Considerations: The integration of technology into assistive devices brings forth ethical considerations, such as data privacy and security. The problem is to address these concerns by implementing safeguards to protect user data and privacy.

A number of intricately linked problems served as the driving force behind the creation of the Smart Glasses with Blind Assistance System: Real-Time Object Detection with Voice Alerts. By offering visually impaired people a wearable solution that encourages independence, improves safety, encourages social inclusion, and provides an amazing user experience, it seeks to address these issues. When these issues are successfully resolved, people with visual impairments will have far better lives, and society will become more accessible and inclusive.

V. SCOPE OF THE PROJECT

The main factor influencing a project or model's accuracy is how well we train it using different data sets, which in turn affects the model's accuracy and useful results. Increased capabilities can also be used to image-detecting systems, self-driving cars, and all security cameras. Moreover, object detection is used in video sequencing things, such as automatically annotating faces in real-time video for subsequent analysis (label and identification), predicting the position of the next item after identifying it in earlier frames, etc. [3] this project can also be done against the backdrop of image processing, character recognition, etc. For the majority of computer and robotic detection systems, object detection is a critical skill. In this topic, recent research has made significant progress in a number of areas. In this article, we provide an overview of the findings of earlier studies, emphasize important research criteria for the present, and discuss potential problems and possible future directions. In addition to handling fundamental information ranging widely from security concerns with a single individual to its application in a real-time scenario, things recognition will be divided into a wide range of organizations. A number of image processing applications, such as photo retrieval, security, observation, computerized vehicle systems, and machine research, use thing recognition.

VI. HARDWARE COMPONENTS USED

1. Laptop or PC
2. Raspberry Pi 4B
3. ESP 32 cam
4. Wearable Glasses
5. DC power bank

Development environment with necessary software tools (IDEs, code editors).

1. Laptop or PC –Laptop with an i5 processor minimum 8 GB RAM keyboard and Mouse with the help of this we developed the initial program for object detection.
2. Raspberry Pi 5-A Raspberry Pi 5 is a key component in object detection as it plays a central role in detecting objects and giving the voice feedback process.

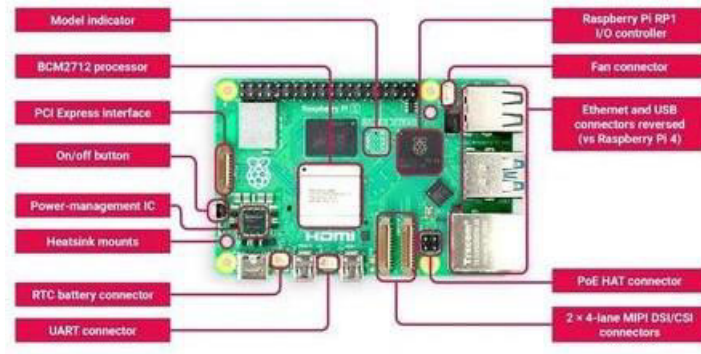


Figure 2.Raspberry Pi

1. 64-bit quad-core Arm Cortex-A76 processor at 2.4GHz for superior CPU performance.
2. Dual 4Kp60 HDMI outputs and advanced Video Core VII GPU for exceptional graphics.
3. Full-size design with Raspberry Pi's own RP1 chip enhancing I/O capabilities.
4. Supports high-bandwidth peripherals with PCIe 2.0, dual-band Wi-Fi, and Bluetooth 5.0.
5. Ideal for educational, hobbyist, and industrial applications.
6. Available in 4GB and 8GB variants for diverse project needs.
7. 1 Year warranty

3. ESP32 cam- ESP 32 cam helps in to give image/video input wirelessly to the raspberry pi 5 for detection using Wi-Fi, which is available inbuilt in ESP32 cam.



Figure 3.ESP32 Cam Module

1. Module Model:ESP32-CAM
2. Custom IO Port:9
3. SPI Flash: Default 32Mbit
4. RAM:520KB SRAM +4M PSRAM5. WIFI:802.11 b/g/n/e/i
5. Bluetooth Version:4.2
6. Antenna Type: IPEX and PCB
7. Input Supply Range (VDC):3.6 ~ 5
8. Current Dissipation (mA):180-310

4. Wearable Glasses - Wearable Glasses is used to place the ESP 32 cam



Figure 4. Wearable Glasses With ESP32 Cam



Figure 5. DC power bank

5. DC power bank – this is used to give power supply to ESP32 cam

VII. SOFTWARE REQUIREMENTS

Python:

Python is commonly used for running scripts, data preparation, and working with deep learning frameworks.

1. OpenCV:
OpenCV (Open Source Computer Vision Library) is a valuable tool for image and video processing, which is often used in conjunction with YOLO for real-time object detection.
2. TensorFlowLite-
The Tensor Flow platform helps you implement best practices for data automation, model tracking, performance monitoring, and model retraining. Using production-level tools to automate and track model training over the lifetime of a product, service, or business process is critical to success.
3. Training Hardware:
Depending on the scale of project, we need a hardware with a capable GPU for training (Raspberry pi 5)
4. Raspbian OS- Raspbian OS for installing in raspberry pi 5
5. Thonny IDE- IDE required for writing and developing python program.

VIII. ALGORITHM USED

OpenCV: An open-source computer and software library is called Open-Source Computer Vision Library (OpenCV). OpenCV aims to expedite the adoption of machine vision in commercial goods by offering a complete infrastructure for computer vision applications. OpenCV is a BSD licensed tool that facilitates code modification and usage for enterprises. More than 2500 sophisticated algorithms are available in the library, including the full collection of machine learning algorithms and both ancient and new computer views. These algorithms may be used to recognize things, detect faces, monitor camera movements, track objects, extract 3D objects, create 3D point clouds on stereo

cameras, and combine images to achieve high resolution. They can also be used to detect human behaviors in films. photo of the entire scene, locate identical pictures on the picture website, eliminate red eyes from flash photos, track eye movement, identify the position and make tags to hide it with an unreal object, etc. OpenCV contains more than 47,000 members of the public and an estimated 18 million downloads. Governmental organizations, businesses, and research teams all make extensive use of the library. In addition to major corporations like Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, and Toyota, several startups also utilize OpenCV, including Applied Minds, VideoSurf, and Zeitera. [3]

Tensor flow-

Deep learning model development, training, and deployment are made easier with Google's open-source Tensor Flow machine learning framework. It provides an extensive ecosystem of materials, frameworks, and tools to facilitate several facets of machine learning development. Tensors, which are multi-dimensional arrays that represent data, are crucial to Tensor Flow's capabilities. In order to create intricate neural network topologies, these tensors flow through computational graphs where operations are carried out to alter and control the input. Tensor Flow offers low-level APIs for precise control over model implementation in addition to high-level APIs for simplified model creation, such as TensorFlow.keras and Tensor Flow Estimator. It is a favored option for corporations, developers, and researchers due to its multiplatform interoperability, rich documentation, and community support.[4]

1. Object detection-

Object detection algorithms are pivotal in computer vision, primarily relying on deep learning models. These algorithms are broadly classified into two categories. Firstly, the two-stage detection algorithm splits the process into distinct steps: generating candidate areas and then classifying and localizing them. Represented by the R-CNN series, including R-CNN, Fast R-CNN, and Faster R-CNN, this approach employs traditional computer vision methods initially, convolutional neural networks (CNNs) in subsequent stages, resulting in heightened complexity, and slower detection speeds. Conversely, the one-stage detection algorithm directly provides object categories and locations without the intermediate step of candidate region generation, prioritizing speed over precision. The YOLO series and SSD series exemplify this category, with YOLO pioneering real-time detection, broadening the application scope of deep learning in object detection. SSD, amalgamating YOLO with anchor box techniques, significantly enhances accuracy over YOLO, achieving a mean Average Precision (mAP) of 72% and a detection speed of 58fps. Thus, while two-stage algorithms offer higher accuracy by integrating traditional and deep learning methods, one-stage algorithms prioritize speed by directly detecting objects from images. Notably, SSD represents a substantial improvement over YOLO, effectively balancing speed and accuracy [4]

2. Object Classification Algorithm-

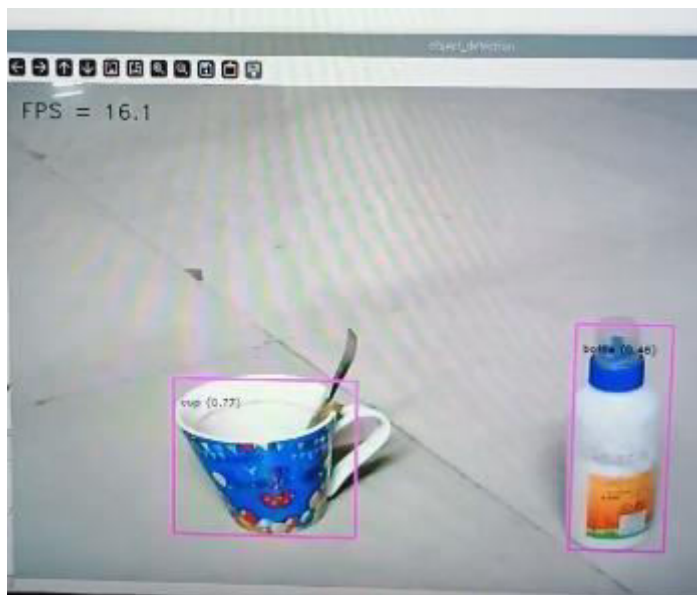
While TensorFlow and OpenCV use different approaches, their object categorization algorithms are based on similar ideas. In OpenCV, preparing input photos to guarantee consistency in size and format is usually the first step in the process. This is followed by feature extraction using conventional computer vision methods such as Haar cascades or Histogram of Oriented Gradients (HOG). The model learns to distinguish between several classes based on labeled data by being trained using these features in machine learning methods like Support Vector Machines (SVM) or k-Nearest Neighbors (k-NN). After being trained, the classifier uses the acquired model to categorize fresh photos. By using TensorFlow.keras or TensorFlow Estimator API, on the other hand, and convolutional neural networks (CNNs), TensorFlow's method entails preparing input data in tensor format and creating deep learning models. Putting the model through training in TensorFlow entails defining the model architecture, compiling it with appropriate loss functions and optimizers, and fitting it to the training data. After training, the model's performance is evaluated on a validation set, and it can then be used for inference on new data to classify images. While OpenCV leans towards traditional computer vision methods and simpler machine learning algorithms, TensorFlow is tailored for deep learning, offering greater flexibility and scalability for building complex models, albeit requiring more expertise and computational resources.[4]

3. Google text to speech (GTTs) –

Google Text-to-Speech (GTTS) serves as a versatile tool for converting text into natural-sounding speech. Its applications span various domains, from accessibility features tailored for visually impaired individuals to language learning aids facilitating pronunciation practice and comprehension. Developers seamlessly integrate GTTS into applications, enabling voice prompts in automated systems and enhancing user experiences in interactive interfaces. Moreover, GTTS finds utility in content creation, generating audio content for podcasts, audiobooks, and educational resources. Its adaptability extends to assistive devices, enriching smart home assistants and wearables with spoken

feedback capabilities. Additionally, content creators leverage GTTS to complement visual content with voiceovers in videos and presentations. In essence, GTTS empowers users across diverse contexts by providing accessible, engaging, and natural-sounding synthesized speech.[5]

IX. RESULT



Voice Feedback - One Cup One Bottle Detected

X. CONCLUSION

To help people with vision impairments navigate in real time, a Python application with sophisticated object tracking and audio feedback has been built. By sounding an alert when an object is detected, this model can function as an alarm alert system in addition to being a navigational assistance. In order to improve accessibility, it might be turned into a hardware device that is used in the biomedical industry to help blind people navigate different kinds of surroundings. To increase accuracy, future research and development initiatives should give priority to model retraining. While the accuracy of the current model is 80%, it may be improved to above 90% by adjusting it using data from the real world. Moreover, implementation on hardware devices such as smart glasses can greatly improve the navigational experience and user experience for people with vision impairments.

XI. FUTURE SCOPE

It can be made more accessible by turning it into a hardware device that can be used in the biomedical industry to help blind people navigate different kinds of situations. To increase accuracy, model retraining should be given top priority in future research and development projects. The accuracy of the current model is 80%, however it may be improved to 90% by using real-world data. Furthermore, the user experience and navigational capabilities for those with visual impairments can be greatly improved by implementation on hardware devices such as smart glasses. In the future, the ability to calculate the distance to identified objects will be taken into consideration. Ultimately, the project will be expanded to include a mobile app, increasing accessibility and userbase.

REFERENCES

1. Blind Assist System-Real Time object detection with distance alert and voice feedback –by- Aryan Gupta <https://medium.com/beingryaan/real-time-object-detection-along-with-distance-and-voice-alerts-for-blinds-a-blind-assistance-1708b97c3ecc>
2. Smart Glass System Using Deep Learning for the Blind and Visually Impaired - Mukhriddin Mukhiddinov,
3. Jinsoo Cho <https://www.mdpi.com/2079-9292/10/22/2756>



4. Tinier-YOLO: A Real-Time Object Detection Method for Constrained Environments -Wei Fang, Lin Wang,
<https://ieeexplore.ieee.org/abstract/document/8941141>
5. Object Detection and Tracking using YOLO- N. Murali Krishna, Ramidi Yashwanth Reddy, Mallu Sai Chandra Reddy, Kasibhatla Phani Madhav, Gaikwad Sudham <https://ieeexplore.ieee.org/abstract/document/9544598>
6. REAL TIME OBJECT DETECTION WITH SPEECH RECOGNITION USING TENSORFLOW LITE
7. -Ganesh Khekare ,Kalpeshkumar Solanki-
https://www.researchgate.net/publication/359393141_REAL_TIME_OBJECT_DETECTION_WITH_SPEECH_RECOGNITION_USING_TENSORFLOW_LITE



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details