



(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 5, May 2025

⊕ www.ijirset.com 🖂 ijirset@gmail.com 🖄 +91-9940572462 🕓 +91 63819 07438





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025 |DOI: 10.15680/IJIRSET.2025.1405395 A Blind Assistance System

N Dhivya, Salma Chandhini M

Assistant Professor, Vivekanandha Institute of Information and Management Studies, Tiruchengode, Tamil Nadu, India

Master of Computer Application, Vivekanandha Institute of Information and Management Studies, Tiruchengode,

Tamil Nadu, India

ABSTRACT: This initiative focuses on creating an innovative assistive system tailored for individuals with visual impairments, utilizing state-of-the-art technologies to improve their independence and safety. The system integrates YOLO (You Only Look Once) for object detection, Grassmann algorithms for facial recognition, and CNN (Convolutional Neural Networks) for identifying currency. Real-time data from a camera is analyzed to deliver auditory feedback, assisting users in navigation, recognizing familiar faces, and determining monetary values. Individuals with visual impairments face considerable obstacles in executing tasks such as safe navigation, identifying acquaintances, and handling financial transactions. Current solutions frequently do not meet their needs due to prohibitive costs, absence of real-time capabilities, and limited accessibility, resulting in a significant segment of this demographic being underserved. To tackle these issues, the project aspires to create an affordable and efficient realtime object detection system for visually impaired users, facilitate seamless face recognition and identification, ensure precise currency recognition, and provide intuitive, voice-based outputs for user-friendly interaction, while guaranteeing scalability and compatibility across various platforms and devices. Real-time data is captured using a camera, processed through the YOLO model for object detection to recognize obstacles and offer navigation assistance, Grassmann algorithms for facial recognition to distinguish between known and unknown individuals, and CNN for the accurate identification of currency denominations. The outcomes are conveyed through voice commands, promoting effective interaction. The system adeptly detect objects, recognize faces, and identifies currency with a high degree of accuracy, providing clear and actionable auditory feedback. It greatly enhances user autonomy and safety. This project delivers a comprehensive, real-time assistive system that connects the gap between visual impairments and practical independence for visually impaired individuals. By merging advanced algorithms and intuitive design, it aims to empower users in their daily lives.

KEYWORDS: Blindness, Visual impairment, Machine Learning, Real-time objects

I. INTRODUCTION

Visually impaired individuals face numerous challenges in their daily lives, including difficulties with navigation, recognizing people, and handling financial transactions. These challenges can lead to a lack of independence, reduced quality of life, and social isolation. The advancement of technology presents an opportunity to address these issues by providing practical, affordable, and efficient solutions. The motivation for this project stems from the need to enhance the autonomy and safety of visually impaired individuals by leveraging modern technologies. Existing assistive systems often lack real-time functionality, are costly, or are inaccessible to a large portion of the population. By integrating innovative technologies such as YOLO (You Only Look Once), Grassmann algorithms, and CNN (Convolutional Neural Networks), this project aims to offer an effective and user-friendly system that bridges the gap between technological possibilities and the needs of the visually impaired community. This initiative is driven by the potential to empower individuals to lead more independent and fulfilling lives, making daily tasks easier and more manageable.

Blind people usually have trouble identifying objects and accessing text-based information, which are fundamental aspects of daily life. Traditional tools, such as braille or white canes, offer very limited support and are insufficient in recognizing products or reading printed materials. Though some advanced systems exist, they are usually quite costly, complicated, and out of reach for many users. Furthermore, such systems usually require huge hardware setups, making them unsuitable for everyday applications.

The motivation for this project arises from the need to provide a cost-effective, efficient, and user-friendly solution. Advances in deep learning and neural networks now enable the development of systems that can efficiently detect





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

DOI: 10.15680/IJIRSET.2025.1405395

objects and recognize text in real time. The aim of this project is to use these technologies in developing systems that will assist the blind in performing tasks such as shopping, reading labels, or even navigating through an environment independently. This is all in the hope of improving their quality of life with technology.

II. OBJECTIVE

The main aim of this project is to create a real-time, intelligent assistive system that greatly improves the independence and quality of life for individuals with visual impairments. By utilizing advanced computer vision methods and deep learning models such as YOLO (You Only Look Once) for object detection, Grassmann algorithms for facial recognition, and CNN (Convolutional Neural Networks) for visual interpretation, the system will empower users to accurately identify objects, people, and text.

The objective is to deliver audio feedback that conveys information about the surrounding environment in real-time, enabling users to engage with the world more confidently and independently. Another essential goal is to develop a system that is both cost-effective and user-friendly. In contrast to conventional assistive devices that are often prohibitively expensive or necessitate complex hardware configurations, this solution will be lightweight, affordable, and easy to use via a wearable or mobile device interface. The system will employ a straightforward camera and microphone setup integrated with software algorithms that operate efficiently on smartphones or embedded hardware like Raspberry Pi. This approach aims to make the technology accessible to a wider audience, including those in low-resource environments who may lack access to high-end technology or specialized support systems. Additionally, the system will seek to include features such as text-to-speech (TTS) for reading printed materials, barcode or QR code recognition for product identification, and potentially GPS-assisted navigation for mobility support. These combined functionalities will work in unison to empower visually impaired individuals in performing daily activities such as shopping, identifying currency, reading medication labels, or navigating public spaces. The overarching aim is to construct a robust, real-time assistant that adapts to the user's surroundings, thereby providing a seamless and dependable technological companion that fosters inclusion and safety.

II. LITERATURE SURVEY

The literature review assesses the effectiveness of the proposed blind assistance system by examining findings from various studies that explore similar approaches using AI and image processing in both simulated and real-world settings. These assessments offer insights into the advantages, obstacles, and potential constraints of deploying an assistive system designed for individuals with visual impairments.

The author in [1] provides the auditory output and scalable training to assist blind persons read the text and recognizing objects using a Raspberry Pi microcontroller, OCR software, and Tensor-Flow Object Detection API. The research [2] proposes a deep attention network employing a ZED stereo camera to guide blind persons around outdoor areas, with an 81% accuracy rate compared to the naïve YOLO v3. The research detailed in [3] presents a robot module designed to aid visually impaired individuals during supermarket shopping. This module utilizes components such as the Raspberry Pi, RFID reader, headphones, and motors to detect and scan RFID tags, provide audio feedback, and record objects in the IoT. Furthermore, the device employs an ultrasonic sensor to identify obstacles, eliminating the requirement for queueing systems. The author of [4] proposes a cost-effective visual assistance system designed to meet the needs of individuals who are completely blind. The Raspberry Pi 3 Model B+ is employed in this system for its compact dimensions and seamless integration capabilities. Equipped with a camera, and sensors, the system enhances accessibility and comfort compared to traditional white canes. The paper [5] introduces a wearable technology solution engineered to aid visually impaired individuals through advanced visual processing capabilities. It utilizes Raspberry Pi to implement artificial vision on the OpenCV platform, allowing for object identification and navigation. The author of the paper [6] presents a system aimed at enabling visually impaired individuals to navigate smart cities independently. It utilizes white canes to interpret colour codes and transmit guiding tactile signals. In summary, AI and image processing hold the potential to enhance assistance systems for the blind by upgrading navigation and legibility for visually impaired people. Nonetheless, addressing issues like accessibility and costeffectiveness requires further research.





[www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405395

III. EXISTING YSTEM

To validate the effectiveness of the proposed model within the Blind Project, it is crucial to compare its performance with existing assistive technologies for the visually impaired. Previous methods, such as traditional OCR (Optical Character Recognition) for currency recognition or standalone face recognition systems, may lack real-time capabilities or require expensive hardware. For instance, conventional face recognition systems might achieve acceptable precision and recall but often struggle with real-time processing against dynamic backgrounds or varying lighting conditions. By contrast, the real-time performance of the YOLO model in detecting objects provides a significant advantage, allowing users to navigate environments more effectively. Moreover, existing systems may rely on user input for interactions or may not provide comprehensive auditory feedback for object classification. The Blind Project's integrated approach, combining real-time object detection, currency recognition, and face identification, offers a unique advantage in providing a seamless experience, thereby significantly improving user independence. Quantitative assessments, such as side-by-side comparisons of accuracy and F1-scores of the model against existing methods, can be conducted to highlight differences in performance. Qualitative insights, such as user satisfaction and usability studies, could also substantiate the claim that the new system is more effective, addressing key gaps left by pre-existing solutions.

IV. PROPOSED SYSTEM

The proposed model for the Blind Project integrates multiple advanced algorithms, each serving a specific purpose to create a comprehensive assistive system for visually impaired individuals. At the core of the system is the YOLO (You Only Look Once) algorithm, which is designed for real-time object detection. The architecture of YOLO consists of a single convolutional neural network that processes the entire image at once, allowing it to predict bounding boxes and class probabilities simultaneously. This approach significantly reduces the computation time compared to traditional methods that require multiple passes over the image. The mathematical formulation involves a loss function that combines localization and classification errors, enabling the model to learn to minimize both simultaneously.

For face recognition, the Grassmann algorithm is employed, which operates on the principle of subspace learning. This algorithm represents faces as points in a high-dimensional space and utilizes linear algebra concepts to compare and identify faces based on their geometric properties. The mathematical formulation involves eigenvalue decomposition to extract essential features from the face images, allowing for efficient recognition against a stored dataset of known faces.

To facilitate currency recognition, a Convolutional Neural Network (CNN) is implemented. The architecture typically consists of several convolutional layers followed by pooling layers, which extract hierarchical features from the input images. The final layers are fully connected, leading to a softmax output that classifies the currency denomination.

V. METHODOLOGY

- 5.1 MODULE
- Image Capture
- Object Detection
- Face Recognition
- Currency Detection
- Voice Output

5.2 MODULES DESCRIPTION

5.2.1 IMAGE CAPTURE

The Image Capture module serves as the primary interface for gathering visual information about the user's surroundings, specifically designed for urban and rural landscapes common in India. This module can utilize affordable cameras integrated into smartphones or smart glasses, making it accessible to a wide audience. The emphasis is on capturing images in real time to support unfamiliar environments, bustling markets, crowded public transportation, or serene rural areas. The camera leverages advanced image processing techniques using libraries like OpenCV to enhance image quality, especially in varying lighting conditions that are prevalent in many Indian settings. For instance, capturing images in bright sunlight or dimly lit interiors of shops can significantly impact visibility. The module





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405395

operates autonomously, consistently collecting visual data without requiring manual initiation, although users can also command the device to capture specific moments with phrases like "Capture now" or "Focus on this." This capability empowers users to maintain a fluid connection with their environment and respond effectively to dynamic and often unpredictable situations around them.

5.2.2 OBJECT DETECTION

The Object Detection module is crucial for interpreting and identifying objects in the user's environment, especially in densely populated urban areas or rural surroundings. Employing deep learning algorithms like YOLO (You Only Look Once), the system can recognize various entities in real time, including pedestrians, vehicles, street signs, and obstacles, which are common in Indian cities. For example, it should distinguish between a cycle rickshaw and a motorbike on a busy street or identify "Caution: cow ahead" in rural areas where animals might roam freely. The system generates auditory alerts to inform users about detected items, providing context-sensitive information tailored to common scenarios in India, such as "There's a tea stall to your left" or "You are approaching a busy intersection." This functionality greatly enhances the safety and mobility of visually impaired users, allowing them to navigate crowded markets, busy roads, or village paths with increased confidence. Continuous training using diverse datasets from various regions in India can improve the accuracy and responsiveness of object recognition, making it a robust tool for everyday navigation.

5.2.3 FACE RECOGNITION

The Face Recognition module plays a vital role in facilitating social interactions for visually impaired users within the Indian context, where community and family ties are significant. Utilizing advanced algorithms like the Grassmann algorithm, this module can recognize familiar faces and announce their names audibly to the user, improving their ability to navigate social settings such as family gatherings, community events, or workplace interactions. It can be particularly valuable for recognizing relatives, friends, or local acquaintances, as face recognition can enhance feelings of connection and belonging, which is essential in Indian society. Users will have the ability to manage this recognition database simply by adding new faces through voice commands or editing existing entries, allowing the system to adapt to their social circles. Privacy concerns are also addressed by providing options to disable alerts for unfamiliar faces, ensuring that users feel secure in their interactions. This module fosters a greater sense of independence and social engagement, empowering users to participate fully in their communities and reduce feelings of isolation.

5.2.4 CURRENCY DETECTION

The Currency Detection module is specifically designed to help users identify Indian currency denominations accurately, which is crucial for independent financial transactions. Leveraging a convolutional neural network (CNN), this module can recognize different Indian Rupee notes and coins, addressing issues like confusion between similar denominations, which is common in daily transactions. When a user presents an Indian Rupee note or coin to the camera, the system analyzes the image and provides immediate feedback, such as "Two hundred Rupees note detected" or "Ten Rupees coin detected." This feature is especially beneficial in diverse transactional environments, including market stalls, ATMs, and shops, where users need to handle cash confidently. The module can also account for the various designs and security features present in Indian currency, ensuring accuracy even as new notes are introduced. Additionally, users could have a feature to recognize various payment methods prevalent in India, such as mobile wallets or QR codes, enhancing their overall financial independence. Continuous updates will ensure the system evolves with currency changes, fostering user confidence during monetary exchanges.

5.2.5 VOICE OUTPUT

The Voice Output module serves as the key communication channel for users, translating visual information gathered from other modules into clear, contextual audio messages. By employing high-quality text-to-speech (TTS) technology, such as those available in regional languages in India, the system can communicate detected objects, recognized faces, or verified currency in a natural manner. Users will receive real-time feedback such as "There is a vendor selling fruits ahead" or "Gaurav is approaching," which helps in building situational awareness. The system supports multiple languages, including Hindi, Tamil, Bengali, and other local dialects, making it accessible to users across different regions. Users can interact with the system using simple voice commands to query the environment, such as "What is nearby?" or "Who is that?" promoting a conversational and engaging experience. Customization options for voice characteristics, speed, and volume ensure that the interactions are tailored to user preferences, enhancing comfort and





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405395

usability. By providing immediate feedback and enabling user engagement, this module is fundamental to promoting independence and confidence among visually impaired individuals.

SYSTEM ARCHTECTURE



VI. EXPERIMENTAL RESULTS

Object detection is the emphasis of the proposed system. Wearable and portable technology has been developed. The person's chest is hooked to the system. The Raspberry Pi camera captures video of the scene, which is subsequently translated into frames by the processor. The auditory output from the system directs the user to the object. Figure 6 shows the detection of a person and a thing (blue cell phone). In Figure 7, a chair is seen with a person. The system displays the object's name its probability as a percentage. As a result, the system will only detect items with a probability more significant than the set threshold. Because You Only Look Once (YOLO) is employed to implement the design on the Android platform, the accuracy of object recognition is reduced. It shows the object's name as well as its probability. The program also informs the user of the class identification and the distance between the object and the camera over the device's speakers







www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405395



VII. CONCLUSION

Several technologies have been created to aid visually impaired persons. One such attempt is that we would wish to make an Integrated Machine Learning System that allows the blind victims to identify and classify real-time objects generating voice feedback and distance. Which also produces warnings whether they are very close or far away from the thing. For visually blind folks, this technology gives voice direction. This technique has been introduced specifically to assist blind individuals. The precision, on the other hand, can be improved. Furthermore, the current system is based on the Android operating system; it can be altered to work with any device that is convenient.

VIII. FUTURE WORK

The future research directions for the assistive system for visually impaired individuals include enhancing data augmentation techniques to address class imbalance and improve model robustness, developing adaptive algorithms to optimize performance in varying environmental conditions, and investigating lightweight model architectures to reduce computational demands. Additionally, exploring explainable AI methods can improve output interpretability, while cross-device compatibility may enhance accessibility. Finally, expanding the application of similar technologies to other disabilities can promote inclusivity and empower various user groups.

REFERENCES

[1] Choi D., and Kim M. (2018). Trends on Object DetectionTechniques Based on Deep Learning, Electronics and Telecommunications Trends, 33(4): 23-32.

[2] Dai Jet al., (2016). R-FCN: Object Detection via Region-based Fully Convolutional Networks. Conf. NeuralInform. Process. Syst.,Barcelona,Spain,Dec. 4-6, p. 379-387.

[3] Dalal N. and Triggs B., Histograms of oriented gradients for Human Detection (2015). IEEE Comput. Soc.Conf. Comput. Vision Pattern Recogn., San Diego, CA, USA, June 20-25, p. 886-893.

[4] Russakovsky O et al., (2015). ImageNet Large Scale Visual Recognition Challenge, Int..J.Comput. Vision, 115(3): 211-252.

[5] Rajeshvaree Ravindra Karmarkar (2021). Object Detection System for the Blind with Voice Guidance, International Journal of Engineering Applied Sciences and Technology,6(2): 67-70.









INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY





www.ijirset.com