iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: V Month of publication: May 2025

DOI: https://doi.org/10.22214/ijraset.2025.71388

Visionbridge: Enabling Independence Through Object, Face and Currency Recognition for Blind

Nandhini E[1], Nanjundeswari M[2], Sreehari P[3], Yokesh V S[4], Swathi H[5]
[1, 2, 3, 4]*UG Students, Department of Computer Science and Engineering, Dhirajlal Gandhi College of Technology, Salem, TamilNadu, India*
[5]*Assistant Professor, Department of Computer Science and Engineering, Dhirajlal Gandhi College of Technology, Salem, TamilNadu, India*

Abstract: Navigating daily life poses significant challenges for blind and visually impaired individuals, particularly in identifying obstacles, recognizing familiar faces, and handling currency transactions. These everyday tasks often require external assistance, leading to a dependency on others and a reduced sense of autonomy. Traditional tools like white canes and guide dogs offer limited functionalities and cannot address the dynamic and complex challenges faced by visually impaired individuals in real-time environments. This study introduces an innovative interface designed to empower blind and visually impaired individuals by enhancing their independence and safety. The proposed system integrates advanced AI powered functionalities such as face detection, obstacle detection, and currency recognition. By utilizing real-time image capturing and processing, the system provides users with immediate, context sensitive audio feedback to assist them in navigating their surroundings and performing essential tasks. The development of such a system is necessary to bridge the gap left by existing assistive technologies, which are often limited in functionality, integration, or affordability. By leveraging advancements in artificial intelligence, computer vision, and wearable technology, the proposed solution addresses critical challenges, fostering greater autonomy and confidence for visually impaired individuals in their daily lives.

Keyword: Assistive Technology, Visual Impairment, Object Detection, Face Recognition, Currency Recognition, Real-Time Image Processing

I. INTRODUCTION

For individuals who are blind or visually impaired, navigating the world presents numerous challenges. Everyday tasks such as recognizing familiar faces, avoiding obstacles, and managing currency transactions often require assistance from others. This dependency on external support reduces their autonomy and confidence in performing daily activities, leading to significant limitations in their social and personal lives. While traditional assistive tools such as white canes and guide dogs have provided valuable help, they come with limitations in functionality. These tools are primarily focused on physical navigation and safety but are unable to offer real-time information about the surrounding environment or identify specific objects, faces, and currency denominations. The inability to access comprehensive, real-time information limits the independence of visually impaired individuals. For example, identifying a familiar face in a crowded setting or detecting moving obstacles in a dynamic environment is beyond the capabilities of traditional tools. Currency handling, which is a critical part of daily life, can also be a significant challenge for blind individuals, leading to financial management difficulties. As a result, visually impaired individuals are often compelled to rely on others for assistance in these essential tasks. The lack of a fully integrated assistive system increases the risk of accidents and limits the opportunity for independent living. With the rapid advancements in artificial intelligence (AI), computer vision, and wearable technology, there is a promising opportunity to develop innovative assistive devices that can significantly improve the lives of visually impaired individuals. AI-powered solutions can combine multiple functionalities, such as face detection, obstacle detection, and currency recognition, into a single device. By processing real-time visual data and providing context-sensitive audio feedback, these systems empower users to navigate their environment and perform daily tasks more independently.

II. LITERATURE SURVEY

A. Moving Object Detection In Complex Scene Using Spatiotemporal Structured-Sparse RPCA

Author: JAVED, SAJID, ET AL.

This paper introduces a novel approach to moving object detection in complex scenes using Spatiotemporal Structured-Sparse Robust Principal Component Analysis (RPCA). The method aims to address challenges such as dynamic backgrounds, camera jitter, and varying illumination conditions that hinder accurate object detection in video sequences.

The proposed algorithm integrates spatiotemporal constraints and structured sparsity into the RPCA framework to better capture the underlying motion characteristics of foreground objects. This fusion enhances the separation of moving objects from the background, producing more accurate and stable detection results. The model effectively deals with occlusions and overlapping regions, which often mislead traditional detection methods.

B. Faster R-Cnn: Towards Real-Time Object Detection With Region Proposal Networks

Author: REN, SHAOQING, ET AL.

Faster R-CNN represents a major milestone in the evolution of object detection by introducing the Region Proposal Network (RPN) to generate region proposals directly from convolutional feature maps. This method significantly boosts the speed and accuracy of object detection tasks. The RPN shares convolutional layers with the detection network, allowing nearly cost-free region proposals and enabling an end-to-end training architecture. Prior to Faster R-CNN, object detection frameworks like R-CNN and Fast R-CNN relied on separate and often slower region proposal techniques like selective search.

C. Deep Residual Learning For Image Recognition

Author: HE, KAIMING, ET AL.

Deep Residual Networks (ResNet) introduced by this paper mark a significant advancement in deep learning, particularly in training very deep neural networks for image recognition tasks. The core innovation lies in the use of residual learning, which allows networks to learn identity mappings that help in the effective training of deep models without degradation. By incorporating shortcut connections that skip one or more layers, the model alleviates the vanishing gradient problem that often hinders the performance of deep architectures. ResNet achieved groundbreaking results in the ImageNet classification challenge, setting new benchmarks for accuracy.

III. EXISITING SYSTEM

In the current technological landscape, visually impaired individuals often rely on traditional tools such as white canes, guide dogs, or tactile walking surfaces for navigation and interaction with their environment. While these tools have been instrumental in assisting the blind, they have limitations in terms of real-time awareness, facial recognition, and identifying small or dynamic objects. These traditional methods lack the ability to provide information about the identity of people around or recognize currency, which are crucial for daily interactions and independent living. As a result, blind individuals still face challenges in tasks that require dynamic understanding of their surroundings. Some existing systems incorporate basic image processing or audio-based alerts but are often limited to either obstacle detection or simple speech feedback. Devices such as ultrasonic sensors and basic wearable cameras are employed to detect obstacles, yet they are unable to differentiate between object types or provide context-specific guidance. Furthermore, these systems are not efficient in identifying faces or detecting currency, and they may fail in crowded or complex environments. The feedback is often limited and not intelligent enough to convey complete awareness, which reduces the system's practical effectiveness and user-friendliness.

A. Disadvantages

- Lack of integration between different functionalities, leading to the need for multiple devices.
- Limited real-time adaptability in detecting dynamic obstacles and changes in the environment.
- Difficulty in distinguishing between various object types, which can lead to inaccurate feedback.
- High cost of specialized assistive devices, limiting accessibility for many users.

IV. PROPOSED SYSTEM

The proposed system introduces an advanced AI-powered assistant tailored for visually impaired individuals, integrating object detection; face recognition, and currency identification into a single, user-friendly solution. Utilizing deep learning algorithms such as YOLO (You Only Look Once), Grassmann for face recognition, and Convolutional Neural Networks (CNN) for currency detection, the system captures real-time input through an external or wearable camera. This live video feed is analyzed to detect obstacles, recognize known individuals, and identify currency notes providing critical feedback via voice output. The combination of these technologies offers a more holistic and intelligent assistive experience that goes beyond traditional tools. The object detection module, driven by YOLO, allows the system to analyze the surrounding environment and detect obstacles like vehicles, people, or other objects with high accuracy and speed. Based on the object's position and movement, the system guides the user to move left or right, enhancing navigation and safety in real-world scenarios. Simultaneously, the face recognition module leverages the Grassmann algorithm to differentiate between known and unknown individuals, which are especially useful in social or professional settings. When a known face is detected, the system announces the name of the person through audio output, fostering better interaction and awareness. In addition, the currency detection module is built using CNN, trained on datasets of various denominations. It enables blind users to identify currency notes simply by holding them in front of the camera, which the system then recognizes and announces audibly.

A. Advantages

1) Provides real-time audio feedback to assist visually impaired individuals.
2) Helps users recognize faces, avoid obstacles, and identify currency independently.
3) Combines multiple features into one wearable device.
4) Uses advanced AI to accurately detect and recognize objects and people.
5) Enhances safety by alerting users to nearby obstacles.
6) Learns and improves over time through machine learning algorithm.

V. METHODOLOGY

A. Requirements Analysis

Objective: Identify essential functionalities required to assist blind users in recognizing faces, detecting obstacles, and identifying currency in real-time with minimal latency.

Approach:

- Data Collection: Acquire datasets for Indian currency notes, human faces (with variations in lighting/pose), and various obstacle objects.
- User Need Assessment: Gather input from visually impaired individuals, caretakers, and accessibility specialists to understand key requirements.
- Functional Specification: Define system modules including camera input, object/face detection, currency recognition, and audio feedback.
- Constraints & Environment: Consider challenges like low-light conditions, varying distances, and limited processing power for real-time on-device performance.

B. System Design and Architecture

Objective: Develop a modular and scalable architecture that supports seamless communication between computer vision models and audio feedback systems.

Approach:

- Modular Design: Separate modules for image capture, object detection, face recognition, currency identification, and audio output.
- Embedded System Integration: Implement on microcontrollers like Raspberry Pi or ESP32 with camera module support.
- Model Management: Deploy pre-trained lightweight CNN models for real-time inference.
- Audio Feedback Engine: Integrate text-to-speech synthesis to convert identified objects/faces/currency into spoken audio output.
- Wearable Setup: Optimize for wearable formats like smart glasses or a belt-mounted camera with audio output via bone conduction or earbuds.

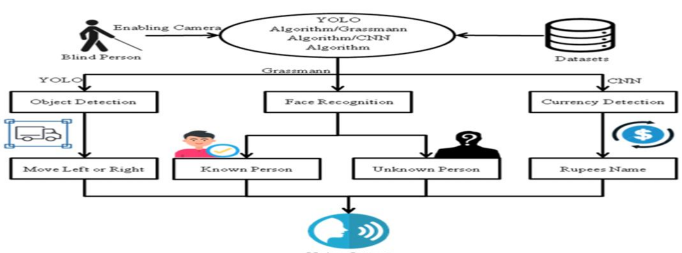

Fig 1:System Architecture

C. Development

Objective: Implement core functionalities for detecting faces, recognizing objects and currency, and delivering real-time feedback.
Approach:

- Camera Input & Image Processing
  - Use camera module for real-time frame capture.
  - Apply preprocessing (resizing, filtering) for better model accuracy.
- Object and Obstacle Detection
  - Employ YOLOv5 for lightweight and accurate detection.
- Face Recognition
  - Use OpenCV for face detection and recognition.
  - Store known faces in a local database for comparison.
- Currency Recognition
  - Train a CNN using Indian currency note images with different denominations and orientations.
- Audio Feedback System
  - Use gTTS (Google Text-to-Speech) or pyttsx3 to deliver spoken results to users.

D. Testing

Objective: Verify that each component performs accurately and reliably in varied real-world conditions.
Approach:

- Unit Testing: Evaluate detection accuracy of individual modules (face, object, currency).
- Integration Testing: Test real-time performance from image capture to audio feedback.
- User Testing: Conduct trials with visually impaired users to assess usability and impact.

VI. CONCLUSION

This project introduces a cutting-edge solution designed to empower visually impaired individuals by integrating face detection, obstacle detection, and currency recognition into a single wearable device. By combining these essential functionalities, the system provides real-time support, helping users overcome various challenges in their daily lives and enhancing their independence. The system utilizes advanced artificial intelligence (AI) technologies, including the Grassmann model for face recognition, YOLO (You Only Look Once) for object detection, and Convolutional Neural Networks (CNNs) for currency recognition. Each of these models has been meticulously implemented to ensure accuracy and real-time performance. The audio feedback module plays a critical role in delivering immediate and context-specific auditory information, guiding users through complex environments. This feature ensures that users can respond to obstacles, recognize familiar faces, and identify currency notes without needing visual assistance. One of the major strengths of this project is its focus on usability and practicality. The wearable device is designed with user convenience in mind, offering a lightweight and portable solution that can be easily adopted in everyday life.

REFERENCES

[1] Javed, Sajid, et al. "Moving object detection in complex scene using spatiotemporal structured-sparse RPCA." IEEE Transactions on Image Processing 28.2 (2018): 1007-1022.
[2] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems 28 (2015).
[3] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[4] Zhang, Han, et al. "Spda-cnn: Unifying semantic part detection and abstraction for fine-grained recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[5] Ouyang, Wanli, et al. "Deepid-net: Deformable deep convolutional neural networks for object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
[6] Elgendy, Mostafa, Cecilia Sik-Lanyi, and Arpad Kelemen. "Making shopping easy for people with visual impairment using mobile assistive technologies." Applied Sciences 9.6 (2019): 1061.
[7] Awad, Milios, et al. "Intelligent eye: A mobile application for assisting blind people." 2018 IEEE Middle East and North Africa Communications Conference (MENACOMM). IEEE, 2018.