

Smart Vision Aid-Currency and Object Recognition System for the Visually Impaired

Abhijith Krishnan¹, Jogimol Joseph²

¹Department of Computer Applications, Musaliar College of Engineering & Technology, Pathanamthitta, Kerala, India Email: [abhijithkrishnan5657\[at\]gmail.com](mailto:abhijithkrishnan5657[at]gmail.com)

²Associate Professor, Department of Computer Applications, Musaliar College of Engineering & Technology, Pathanamthitta, Kerala, India

Abstract: *Visual impairment is a significant challenge that affects the ability of individuals to perform daily tasks independently. Identifying objects and recognizing currency are common difficulties faced by visually impaired people, often requiring assistance from others. Traditional methods such as touch-based identification are not always reliable and can lead to confusion, especially in complex environments. Visual impairment is a significant challenge that affects the ability of individuals to perform daily tasks independently. Identifying objects and recognizing currency are common difficulties faced by visually impaired people, often requiring assistance from others. Traditional methods such as touch-based identification are not always reliable and can lead to confusion, especially in complex environments. Once the detection is completed, the system converts the recognized information into audio output using text-to-speech technology, enabling users to understand the results without visual interaction. The system operates in real time, providing instant feedback and improving user independence in daily activities. Once the detection is completed, the system converts the recognized information into audio output using text-to-speech technology, enabling users to understand the results without visual interaction. The system operates in real time, providing instant feedback and improving user independence in daily activities.*

Keywords: Smart Vision Aid, Object Detection, Currency Recognition, CNN, YOLO, Computer Vision, Assistive Technology, Text-to-Speech, Real-Time Processing, Image Recognition

1. Introduction

Visual impairment is a major challenge that affects the daily life of many individuals around the world. Visually impaired people often face difficulty in identifying objects and recognizing currency in their surroundings. These limitations reduce their independence and require them to depend on others for basic activities. Traditional methods such as touch-based identification or human assistance are not always reliable and can lead to confusion or mistakes. Visual impairment is a major challenge that affects the daily life of many individuals around the world. Visually impaired people often face difficulty in identifying objects and recognizing currency in their surroundings. These limitations reduce their independence and require them to depend on others for basic activities. Traditional methods such as touch-based identification or human assistance are not always reliable and can lead to confusion or mistakes.

The Smart Vision Aid system is designed to provide an intelligent solution for object detection and currency recognition. The system uses a camera to capture images, processes them using deep learning models, and provides output through audio. This helps visually impaired users to perform daily activities more independently and confidently. Traditional assistive methods such as touch-based identification and manual support have several limitations. Objects with similar shapes or textures can be difficult to distinguish, and worn-out currency notes can lead to incorrect identification. In addition, depending on others for simple tasks is not always practical or safe. These challenges highlight the need for a smart and reliable solution that can assist users in real-time.

With the rapid advancement of artificial intelligence,

machine learning, and computer vision, it has become possible to develop intelligent assistive systems. Technologies such as Convolutional Neural Networks (CNN) and YOLO enable accurate detection and classification of objects from images. These systems can process visual data quickly and provide meaningful outputs, making them suitable for real-time applications.

The Smart Vision Aid system is designed to address these challenges by combining object detection, currency recognition, and audio feedback into a single platform. The system captures images using a camera, processes them using deep learning models, and provides output through text-to-speech technology. This approach allows visually impaired individuals to understand their surroundings more effectively and perform daily activities with greater independence and confidence.

The Smart Vision Aid system is designed to address these challenges by combining object detection, currency recognition, and audio feedback into a single platform. The system captures images using a camera, processes them using deep learning models, and provides output through text-to-speech technology. This approach allows visually impaired individuals to understand their surroundings more effectively and perform daily activities with greater independence and confidence.

2. Objectives

The primary objective of this project is to design and develop an intelligent assistive system that helps visually impaired individuals identify objects and recognize currency notes using computer vision and audio feedback.

The specific objectives of the proposed system are as follows:

- To capture real-time images using a camera
- To detect and identify common objects using deep learning technique.
- To recognize different currency denominations accurately
- To convert detected results into speech output using text-to-speech technology
- To provide real-time assistance for visually impaired users
- To improve independence in daily activities.
- To design a simple and user-friendly interface
- To ensure fast and accurate detection results
- To develop a cost-effective assistive solution
- To create a scalable system for future improvements

3. Existing System

In the existing system, visually impaired individuals mainly depend on traditional methods such as touch and external assistance to identify objects and currency. These methods are not always accurate and can lead to confusion, especially when objects are similar in shape or size. Currency recognition using touch becomes difficult when notes are worn out or damaged.

Most traditional assistive systems are limited in functionality and do not provide real-time feedback. Some systems only focus on object detection or currency recognition separately, and they lack integration of both features in a single system. Additionally, many existing solutions are expensive and not easily accessible to all users.

In the existing system, visually impaired individuals mainly depend on traditional methods such as touch and human assistance to identify objects and recognize currency. These methods are often slow and not fully reliable, especially in unfamiliar environments. Identifying objects based only on shape or texture can be confusing, and mistakes are common when objects are similar.

Most existing assistive technologies are limited in functionality and usually focus on a single feature, such as only object detection or only currency recognition. They do not provide an integrated solution that combines multiple features into one system. Additionally, many systems do not offer real-time feedback, making them less effective for practical daily use.

Another major limitation of current systems is the lack of proper audio support. Without clear and instant audio feedback, visually impaired users may find it difficult to understand the results. Some systems also require internet connectivity, which reduces reliability in offline conditions and limits accessibility.

4. Proposed System

The proposed Smart Vision Aid system is designed to overcome the limitations of existing systems by providing an integrated and intelligent solution. The system uses a camera to capture real-time images from the user's surroundings. These images are then processed using deep learning algorithms such as CNN and YOLO for object detection and currency recognition.

Initially, the captured image undergoes preprocessing steps such as resizing and noise reduction to improve quality. After preprocessing, the image is passed to trained models that analyse and identify objects and currency denominations accurately. The system ensures fast processing to provide real-time results.

Once the detection process is completed, the identified information is converted into text and then into speech using a text-to-speech module. This audio output allows visually impaired users to understand the detected results easily without any visual interaction. The system continuously works in real time to provide instant assistance.

The proposed system is designed to be simple, efficient, and user-friendly. It is also cost-effective and can work in offline conditions, making it more accessible. In the future, the system can be enhanced with features such as mobile integration, multilingual support, and improved accuracy, making it a powerful assistive tool for visually impaired individuals.

The proposed Smart Vision Aid system is designed to overcome the limitations of existing systems by integrating object detection, currency recognition, and audio feedback into a single platform. The system uses a camera to capture images and processes them using deep learning models such as CNN or YOLO.

The captured image is analysed to detect objects and identify currency notes. Once the detection is completed, the system converts the result into speech output using a text-to-speech module. This allows visually impaired users to understand the detected information easily.

5. Methodology

The Smart Vision Aid system follows a structured methodology that includes continuous image capture, processing, detection, and audio output. The system works in a step-by-step process to ensure accurate and real-time assistance.

The camera captures images from the surrounding environment, which are then sent to the processing module. The image is preprocessed to improve quality and remove noise. After preprocessing, the image is passed to the object detection and currency recognition models.

The trained deep learning models analyse the image and identify objects and currency denominations. The detected results are then converted into text format. This text is further processed by a text-to-speech engine to generate audio output.

Under normal conditions, the system continuously captures and processes images without interruption. When an object or currency is detected, the system immediately provides audio feedback to the user. This ensures real-time interaction and assistance. The system is designed to be efficient, accurate, and responsive. It provides continuous support to visually impaired users and enhances their ability to perform daily tasks independently.

The Smart Vision Aid system follows a structured and efficient methodology to provide real-time assistance for visually impaired users. The process begins with capturing images from the surrounding environment using a camera. The system continuously collects visual data so that users can identify objects and currency whenever needed. This real-time image acquisition ensures that the system remains active and responsive at all times.

Once the image is captured, it is sent to the preprocessing stage where the quality of the image is improved. This includes operations such as resizing, noise reduction, and image enhancement. Preprocessing helps in preparing the image in a suitable format for the deep learning models, thereby improving the accuracy of detection and recognition.

YOLO: YOLO (You Only Look Once) is a popular real-time object detection algorithm used in computer vision. Unlike traditional methods that detect objects in multiple steps, YOLO processes the entire image in a single pass, making it extremely fast and efficient. It divides an image into a grid and predicts bounding boxes and class probabilities directly, allowing it to identify multiple objects simultaneously. This makes YOLO highly suitable for applications that require real-time performance, such as surveillance systems, self-driving cars, and assistive technologies.

One of the key advantages of YOLO is its speed and accuracy balance. It can detect objects like people, vehicles, animals, and everyday items with high precision while maintaining real-time processing capabilities. YOLO uses deep neural networks, particularly convolutional neural networks (CNNs), to learn features from images. Over different versions (YOLOv1 to YOLOv8), the algorithm has improved significantly in terms of detection accuracy, handling small objects, and reducing errors. However, earlier versions sometimes struggled with detecting very small or closely packed objects.

6. System Architecture

The Smart Vision Aid system consists of a camera module, processing unit, object detection model, and audio output system. The camera captures real-time video of the user's surroundings.

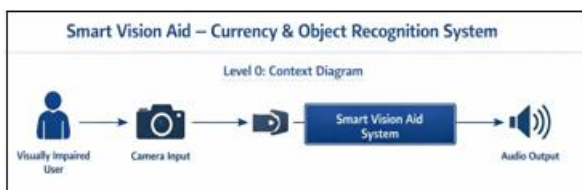


Figure 1: Block Diagram of the Proposed Smart Vision Aid – Currency and Object Recognition System for the Visually Impaired

The captured video frames are sent to the YOLO object detection model, which identifies multiple objects in each frame, such as people, vehicles, furniture, and obstacles. Detected objects are then processed to extract labels and positions, which are converted into audio feedback using a Text-to-Speech module.

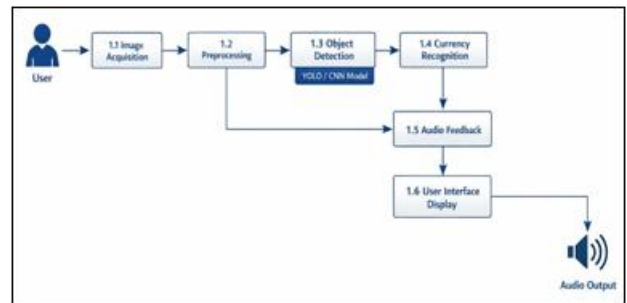


Figure 2: Level 1: Smart Vision Aid – Currency and Object Recognition System for the Visually Impaired

The system is designed to be lightweight and portable, making it suitable for deployment on embedded devices, mobile phones, or wearable platforms, allowing visually impaired users to navigate safely and independently.

7. Hardware Components

The major hardware components used in the proposed smart vision aid- currency and object recognition system for the visually impaired are described below.

- Provides audio feedback to the user. Connected via the processing unit for real-time alerts.
- Provides power to the camera, processing unit, and speaker. For portable versions, a rechargeable battery pack can be used.
- Ultrasonic or IR sensors for distance measurement. GPS module for outdoor navigation.

8. Implementation

The implementation starts with capturing video input using a camera connected to the processing unit or mobile device.

- The frames are processed by a pre-trained YOLO model, which detects multiple objects and predicts their locations in each frame.
- The detected object labels are converted to voice output using a Text-to-Speech engine, which continuously announces objects in real-time to the user.
- The entire system is integrated using Python and PyTorch for the detection model, OpenCV for video handling, and pyttsx3 or gTTS for audio output, creating a seamless pipeline for real-time assistance.

9. Results and Comparison

Evaluation of the Smart Vision Aid System shows significant improvements in real-time object and currency recognition, responsiveness, and user accessibility compared to traditional methods. Conventional solutions, such as tactile markings, asking for help from others, or relying on memory, are limited in independence and do not provide intelligent recognition. In contrast, the proposed system utilizes machine learning to automatically identify objects and currency in real time. The implementation of on-device TensorFlow/Keras inference ensures low latency and complete offline functionality, enhancing reliability in daily use. Audio feedback allows users to understand results immediately, improving confidence and independence. Performance testing demonstrated accurate

recognition under various lighting conditions, angles, and object orientations. The modular architecture also enables easy scalability for adding new objects, currencies, or features in future updates. Overall, the Smart Vision Aid System proved to be more efficient, portable, intelligent, and privacy-preserving than traditional methods, significantly enhancing independence, usability, and safety for visually impaired users.

ROC CURVE: ROC (Receiver Operating Characteristic) curves were plotted for each class, It shows the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) at different threshold values. The True Positive Rate (also called sensitivity or recall) indicates how well the model correctly identifies positive cases, while the False Positive Rate shows how often the model incorrectly identifies negative cases as positive.

10. Conclusion

The Smart Vision Aid provides a real-time AI-based assistive solution for visually impaired individuals, helping them navigate safely and independently. By combining YOLO object detection with Text-to-Speech feedback, the system identifies objects in the surroundings and informs the user through audio alerts, enhancing situational awareness. The project demonstrates the effectiveness of modern computer vision and deep learning techniques in practical applications, showing

daily life. The project demonstrates the effectiveness of modern computer vision and deep learning techniques in practical assistive technologies. By combining YOLO object detection and Text-to-Speech feedback, the system enhances situational awareness and helps users navigate safely.

References

- [1] PyTorch Documentation. (2023). PyTorch: Tensors and Deep Learning.
- [2] OpenCV Documentation. (2023). Open Source Computer Vision Library. <https://opencv.org/>.
- [3] Sayfullina, A., and Aliyev, R. (2020). Text-to-Speech Conversion in Python using pyttsx3. Journal of Computer Science Applications.
- [4] Bochkovskiy, A., Wang, C.Y., and Liao, H.Y.M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934.
- [5] Redmon, J., and Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.
- [6] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. CVPR 2016.

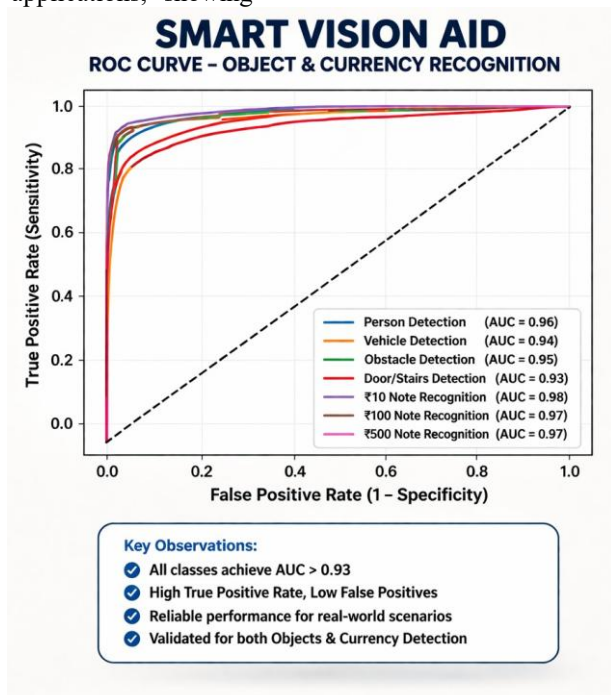


Figure 3: ROC CURVE

that technology can significantly improve the daily life of visually impaired people.

With further improvements, such as distance measurement, GPS integration, and portable deployment, the system has the potential to become a reliable, everyday companion for visually impaired users. With further improvements and deployment on portable devices, the system can become a reliable companion for visually impaired individuals in their

- [7] Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C.L. (2014). Microsoft COCO: Common Objects in Context. ECCV 2014.
- [8] Goodfellow, I., Bengio, Y., and Courville, A. (2013). Deep Learning. MIT Press.