

Vision Guide — A Smart AI-Based Assistive Application for the Visually Impaired: An Integrated Framework for Real-Time Face and Currency Recognition

Deeksha P¹, Divyadharshini S², Miruthula Sri S³, Rashmi Risha J⁴, Thivesha M S⁵ and Mrs. C Janani⁶

^{1,2,3,4,5}UG Student, ⁶Assistant Professor, Department of Computer Science and Engineering, School of Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore 641108, TamilNadu, India

Abstract Visually impaired individuals face persistent challenges in recognizing people and handling currency independently in daily life. This paper presents vision guide, a smart AI-based assistive mobile application that integrates real-time face recognition and Indian currency recognition into a single accessible platform. The face recognition module captures live video through the smartphone camera, detects human faces, extracts unique facial features using deep learning techniques, and compares them against a stored face dataset to identify known individuals. Unrecognized faces are flagged as unknown to improve situational awareness. The currency recognition module applies a Convolutional Neural Network (CNN) to classify multiple Indian currency denominations after image preprocessing steps such as noise removal and image enhancement. Both modules deliver their results through a text-to-speech engine, allowing users to receive audio feedback without depending on visual interaction. The application is designed with a simple and accessible interface tailored to the needs of visually impaired users. By combining computer vision, deep learning, and speech synthesis within a lightweight mobile solution, vision guide enhances the independence, safety, and confidence of visually impaired individuals in social interactions and financial transactions.

Key Words: Assistive Technology, Face Recognition, Currency Recognition, Convolutional Neural Network, Deep Learning, Text-to-Speech, Visually Impaired, Computer Vision.

1. INTRODUCTION

Visually impaired individuals encounter significant difficulty in performing everyday tasks that depend on visual perception, particularly recognizing people in their surroundings and identifying currency during financial transactions. Conventional aids such as walking sticks and human assistance offer only partial support and do not address the need for independent social interaction or

safe handling of money. Recent advances in Artificial Intelligence (AI), deep learning, and mobile computing have opened new possibilities for building lightweight, camera based assistive systems that operate entirely on a smartphone without specialized hardware.

This paper proposes vision guide, an AI based assistive mobile application that unifies two essential capabilities face recognition and Indian currency recognition into a single accessible platform. The face recognition module uses the smartphone camera to capture live video, detect faces, extract distinctive facial features through deep learning, and compare them against a stored dataset to identify known persons, while unknown faces are reported to increase user awareness. The currency recognition module applies a Convolutional Neural Network (CNN) to classify Indian currency notes of denominations ₹10, ₹20, ₹50, ₹100, ₹200, and ₹500 after preprocessing the captured image to improve clarity and remove noise. Both modules convert their output into speech using a text to speech engine, enabling the user to receive information audibly.

Unlike systems that require dedicated hardware or address only a single assistive function, vision guide adopts a unified, software only architecture built for standard smartphones, making it affordable and readily deployable. By integrating face recognition, currency recognition, and voice based feedback within one application, the proposed system aims to reduce dependency on others, improve safety, and enhance the overall independence and confidence of visually impaired users in daily life.

2. LITERATURE SURVEY

A number of studies have explored Artificial Intelligence (AI), deep learning, and machine learning techniques to develop assistive technologies for visually impaired individuals. Existing research has mainly focused on face recognition, currency recognition, and smartphone based

assistive systems. The following studies provide the foundation for the proposed VISION GUIDE system.

Schroff et al. [1] proposed FaceNet: A Unified Embedding for Face Recognition and Clustering, published in the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015. The methodology uses deep convolutional neural networks to recognize and verify human faces by learning distinctive facial features. The proposed approach achieves high recognition accuracy and robustness. However, it requires a large training dataset and significant computational resources, limiting its deployment on resource constrained devices.

Bansal and Sharma [2] presented Indian Currency Recognition Using Image Processing and Machine Learning, published in the International Journal of Computer Applications (IJCA). The methodology combines image processing and machine learning techniques to identify and classify Indian currency notes. The proposed system provides reliable recognition under normal conditions. However, its performance may decrease when currency notes are damaged, folded, or captured under poor lighting conditions.

Liu et al. [3] proposed Smartphone Based Indoor Localization Using Sensors and Machine Learning, published in the IEEE Systems Journal. The methodology combines smartphone sensors with machine learning algorithms to improve indoor localization for visually impaired users. The proposed system enhances navigation accuracy in indoor environments. However, it mainly focuses on localization and does not integrate face recognition and currency recognition within a single mobile application.

The reviewed literature demonstrates that face recognition, currency recognition, and smartphone based assistive technologies have achieved promising results individually. However, most existing systems focus on only one specific functionality. The proposed vision guide application addresses this limitation by integrating face recognition, Indian currency recognition, and voice based feedback into a single lightweight mobile application, providing a comprehensive assistive solution for visually impaired individuals.

3. PROBLEM STATEMENT

Visually impaired individuals face considerable difficulty in recognizing people around them and verifying currency denominations without external assistance.

Existing support methods, such as walking sticks or help from other people, provide only limited, non-scalable assistance and cannot supply real-time identification of faces or currency. Where digital assistive tools do exist, they are typically single-purpose and may depend on specialized or costly hardware. They often lack reliable voice-based output, making them difficult for visually impaired users to operate independently.

In addition, recognition accuracy in many existing systems is highly sensitive to lighting conditions and camera quality. It is also affected by image clarity. The absence of an integrated, camera-only, software-based solution forces users to depend on multiple disconnected tools for different needs. There is a clear need for a unified, accessible, AI-powered mobile application that can reliably perform both face recognition and currency recognition and communicate results through speech, thereby improving the daily independence, safety, and confidence of visually impaired individuals.

4. OBJECTIVES

The objectives of this work are as follows:

- To develop a face recognition module that detects faces from a live camera feed, extracts facial features using deep learning, and identifies known individuals by comparison with a stored dataset.
- To develop a currency recognition module using a Convolutional Neural Network (CNN) that accurately classifies Indian currency denominations of ₹10, ₹20, ₹50, ₹100, ₹200, and ₹500.
- To integrate a text-to-speech engine that converts recognition results from both modules into clear audio output.
- To design a simple, accessible user interface suited to the operational needs of visually impaired users.
- To combine both modules into a single mobile application, reducing dependency on multiple tools and improving user independence and confidence.

5. EXISTING SYSTEM

Existing assistive solutions for visually impaired individuals mostly rely on traditional aids such as walking sticks and human assistance, which provide only limited situational information and require continued dependence on others. Where digital tools are available, they typically address a single function, either face identification or currency recognition, rather than combining both within one application, and some require additional hardware or complex operating procedures. Current currency recognition tools frequently use basic image processing techniques that are sensitive to poor

lighting, unclear images, or damaged notes, and many lack effective voice based output, making results difficult for visually impaired users to interpret without visual support.

Limitations of Existing Systems

- Limited user independence, since traditional aids do not provide identity or currency information.
- Single-function design, requiring separate applications or devices for different needs.
- Reduced accuracy under variable lighting, angle, or image-quality conditions.
- Inadequate or absent voice-based feedback for visually impaired users.
- Dependency on specialized hardware in some solutions, increasing cost and reducing portability.
- Slow processing that limits real-time usability.

6. PROPOSED SYSTEM

The proposed vision guide application overcomes these limitations by integrating face recognition and currency recognition into a single, software only mobile application that operates using only the smartphone camera. The face recognition module continuously analyzes the live camera feed, detects human faces, extracts distinctive facial features through deep learning, and compares them with a stored face dataset to identify known individuals, while flagging unrecognized faces to support user awareness. The currency recognition module applies image preprocessing followed by a CNN classifier to accurately identify Indian currency denominations. Both modules deliver their output through a text to speech engine, giving the user immediate audio feedback.

The system follows a modular architecture in which the User Interface, Camera Input, Face Recognition, Currency Recognition, and Text to Speech components interact through a coordinated processing pipeline. This design keeps the application lightweight, avoids dependency on external hardware, and allows each module to be maintained, retrained, or extended independently, for example, by expanding the stored face dataset or adding new currency denominations, without disrupting the overall system.

Advantages of the Proposed System

- Improves user independence by enabling recognition of people and currency without external help.

- Provides real-time face recognition with voice-based identification of known and unknown individuals.
- Accurately classifies Indian currency denominations using a trained CNN model.
- Delivers all results as speech, removing dependence on visual or text-based interfaces.
- Uses only the smartphone camera, avoiding the cost and inconvenience of dedicated hardware.
- Combines two assistive functions in one lightweight application, simplifying daily use.

7. SYSTEM ARCHITECTURE (BLOCK DIAGRAM)

The vision guide architecture begins at the User Interface Module, which provides visually impaired users with a simple, accessible entry point for selecting either the face recognition or currency recognition function. Once a function is selected, the Camera Input Module activates the smartphone camera to capture live video frames or a still currency image. The captured data is then passed to the corresponding recognition pipeline.

For face recognition, captured frames are preprocessed and passed through a deep learning based feature extractor. The resulting facial features are compared against the stored Face Dataset to determine whether the person is known or unknown. For currency recognition, the captured note image undergoes preprocessing, including resizing and noise reduction. It is then classified by the trained CNN model, which outputs the identified denomination. In both cases, the recognition result is passed to the Text to Speech Conversion Module. The module generates the corresponding audio message. The Voice Assistance Output Module then delivers this feedback to the user in real time.

This modular pipeline begins with the User Interface Module, followed by the Camera Input Module, the Face Recognition or Currency Recognition Module, the Text to Speech Conversion Module, and finally the Voice Assistance Output Module. This design keeps processing efficient and allows each stage to be optimized independently. It ensures the application remains responsive on standard smartphone hardware.

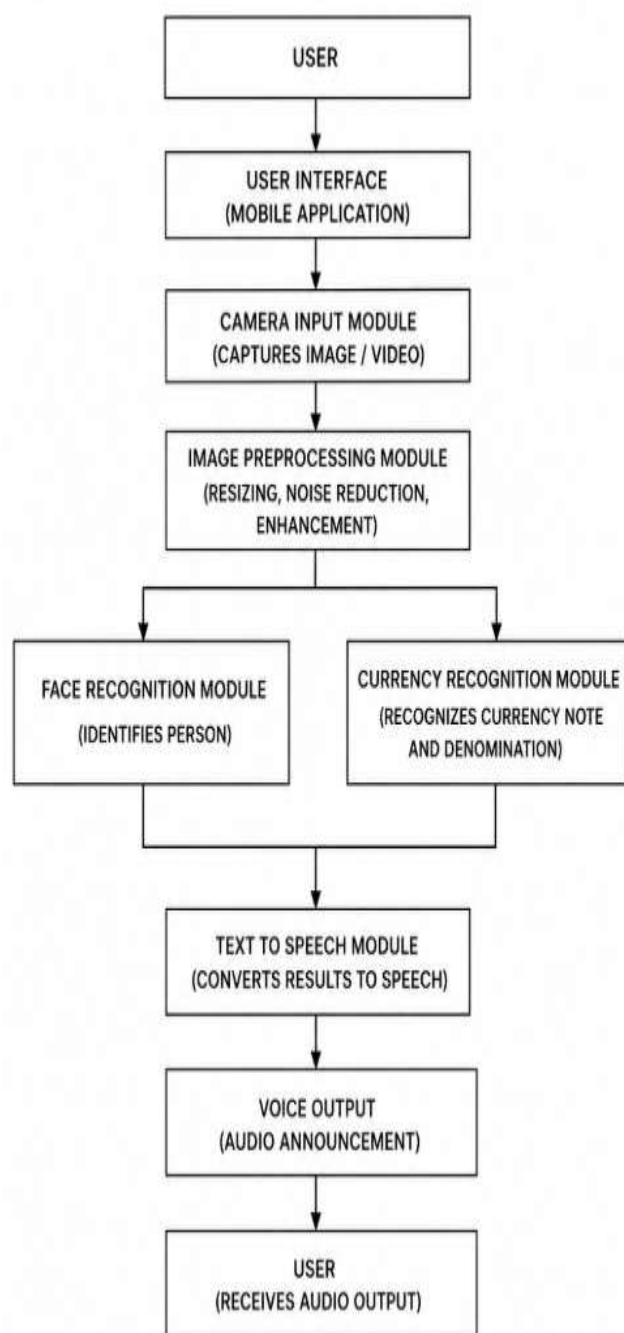


Fig. 1 — Block Diagram of VISION GUIDE System Architecture

8. DATA FLOW DIAGRAM (DFD)

The data flow diagram below illustrates the step-by-step movement of data through the vision guide application, from user interaction at the mobile interface through camera capture, recognition processing, and speech-based delivery of results back to the user.

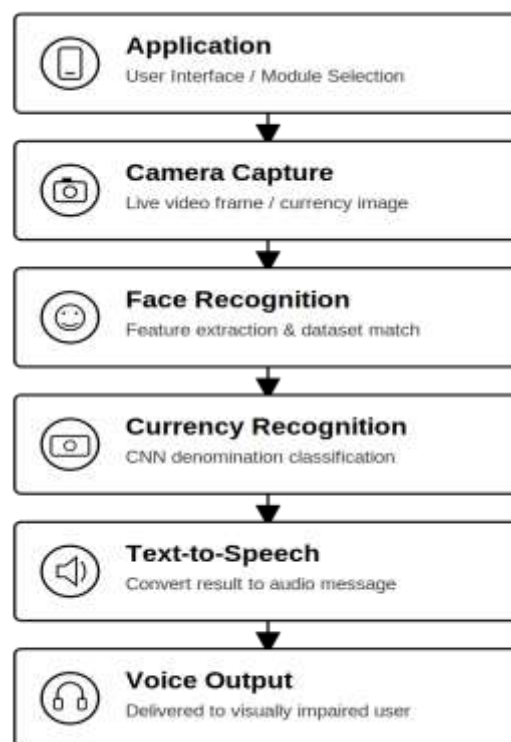


Fig. 2 — Data Flow Diagram of VISION GUIDE

9. METHODOLOGY

The vision guide methodology begins when the user launches the application and selects either the face recognition or currency recognition function from a simple, accessible interface. Selecting a function activates the smartphone camera, which supplies real-time visual input to the corresponding AI processing pipeline.

9.1 Face Recognition Workflow

The camera continuously captures live video frames from the surrounding environment. Detected faces are preprocessed to improve image quality, and deep learning techniques extract unique facial features from each detected face. These features are compared against the stored face dataset; a match identifies the known individual by name, while the absence of a match classifies the face as unknown. The result is converted into speech, allowing the user to identify people nearby without visual confirmation.

9.2 Currency Recognition Workflow

The user captures an image of an Indian currency note using the smartphone camera. The image undergoes preprocessing, including resizing, noise reduction, and quality enhancement, before being analyzed by the

trained CNN model. The CNN extracts distinguishing visual features, including patterns, colour distribution, and texture, to classify the note into one of the supported denominations (₹10, ₹20, ₹50, ₹100, ₹200, ₹500). The recognized value is converted into an audio announcement through the text-to-speech engine.

This two-pipeline methodology, unified under a single voice-driven interface, allows vision guide to deliver consistent, real-time assistance for both social recognition and financial independence within one lightweight mobile application.

10. MODULE DESCRIPTION

10.1 User Interface Module

Provides a simple, accessible entry point for selecting the required function, connects user input to the backend AI processing pipeline, and presents results to the user through voice-based feedback.

10.2 Camera Input Module

Captures real time visual information by acquiring live video frames for face recognition and still images for currency recognition, then forwards the captured input to the respective recognition module for preprocessing and analysis.

10.3 Face Recognition Module

Detects human faces from live camera frames, extracts unique facial features using deep learning, and compares them with the stored face dataset to identify known individuals. Unrecognized faces are reported as unknown, and the result is passed to the speech module.

10.4 Currency Recognition Module

Processes the captured currency image, extracts visual features such as colour, pattern, and texture, and classifies the note into the correct denomination using a trained CNN model.

10.5 Text-to-Speech Conversion Module

Converts the recognition results from both modules, including identified persons and recognized currency values, into clear and understandable speech output.

10.6 Voice Assistance Output Module

Delivers the generated audio response to the user in real time, completing the recognition-to-feedback pipeline and enabling the user to act on the information immediately.

11. TECHNOLOGIES USED

- Python — core programming language for AI model development and application logic.
- Convolutional Neural Network (CNN) — used for currency image classification.
- Deep learning-based feature extraction — used for facial detection and recognition.
- OpenCV — used for image preprocessing operations such as resizing and noise reduction.
- TensorFlow / Keras — used for building, training, and deploying the deep learning models.
- Text-to-Speech (TTS) engine — converts recognition output into audio feedback.
- SQLite — lightweight local database for storing face and model-related data.

12. RESULTS AND DISCUSSION

The vision guide application was successfully implemented as an integrated mobile assistive system combining face recognition and Indian currency recognition with voice-based output. The Face Recognition Module reliably detected human faces from live camera input and correctly distinguished known individuals from unknown ones, with recognized names delivered promptly through the text-to-speech engine, supporting improved social interaction for the user.

The Currency Recognition Module accurately classified Indian currency notes across the supported denominations (₹10, ₹20, ₹50, ₹100, ₹200, and ₹500) after preprocessing, with the CNN model successfully learning distinguishing visual patterns across the note images. The Text-to-Speech and Voice Assistance modules consistently converted both types of recognition results into clear audio feedback, confirming the practicality of a fully voice-driven interaction model for visually impaired users.

Overall, the integration of face recognition and currency recognition into a single, camera-only mobile application demonstrates that visually impaired users can be supported through one lightweight, affordable tool rather than multiple disconnected applications or hardware devices, simplifying daily use and reducing dependency on others.

13. CONCLUSION

This paper presented vision guide, a smart AI-based assistive mobile application that integrates face recognition and Indian currency recognition to support visually impaired individuals in daily life. By combining deep learning-based face identification, CNN-based

currency classification, and text-to-speech output within a single, accessible application, the system enables users to recognize people and handle money independently, without relying on additional hardware or external assistance. The results confirm that a unified, voice-driven, camera-only approach can meaningfully improve the independence, safety, and confidence of visually impaired individuals in both social and financial contexts.

14. FUTURE SCOPE

Future enhancements to vision guide may include obstacle detection, text reading, and navigation assistance to extend the application beyond its current face and currency recognition capabilities. Recognition accuracy can be further improved through larger, more diverse training datasets and more advanced deep learning architectures, along with better robustness under challenging lighting conditions and damaged or folded currency notes. Additional planned directions include cloud-based storage for the face dataset, automatic learning of new faces with user consent, multilingual voice assistance, and integration with wearable devices such as smart glasses for hands-free operation.

REFERENCES

- [1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 779–788, 2016.
- [2] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.
- [3] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," Computer Vision and Pattern Recognition, 2020.
- [4] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Int. Conf. Learning Representations (ICLR), 2015.
- [5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.
- [6] R. Szeliski, Computer Vision: Algorithms and Applications, 2nd ed., Springer, 2022.
- [7] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th ed., Pearson Education, 2018.
- [8] F. Chollet, Deep Learning with Python, 2nd ed., Manning Publications, 2021.
- [9] M. Abadi et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," 2016.
- [10] G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.
- [11] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.
- [12] A. Sharma and R. Kumar, "Artificial Intelligence Based Assistive Technologies for Visually Impaired People: A Review," Int. J. Computer Vision and Applications, 2023.
- [13] D. Patel and M. Shah, "Deep Learning Approach for Currency Recognition System Using Convolutional Neural Networks," Int. J. Intelligent Systems and Applications, 2022.
- [14] Google Developers, "Text-To-Speech Technology and Android Accessibility Services Documentation," 2024.
- [15] TensorFlow Developers, "TensorFlow Lite: Machine Learning Framework for Mobile and Embedded Devices," 2024.