Akash Sn, 2026, 14:1
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

AI-based companion system for medicine and currency recognition for visually impaired users

Akash Sn, Adithiya S, Muni Praharsha M
Anna university

Abstract - Visual impairment restricts independent recognition of everyday objects such as medicines and currency, leading to medication errors and financial dependency. This paper presents an AI-based companion system designed to assist visually impaired users in real-time medicine and currency recognition. The system utilizes computer vision and Convolutional Neural Networks (CNNs) to classify medicines based on packaging features and identify currency denominations using distinctive visual patterns. Optical Character Recognition (OCR) is integrated to extract drug names and dosage information from printed text. The system provides audio feedback through text-to-speech and supports voice commands for hands-free interaction. Designed for deployment on low-cost smartphone or embedded platforms, the proposed solution ensures affordability and portability. Experimental results demonstrate high recognition accuracy under varying environmental conditions. The system enhances user independence, reduces reliance on caregivers, and improves safety in medication management and financial transactions, highlighting the potential of AI-driven assistive technologies.

Keywords - Assistive Technology, Computer Vision, OCR, Deep Learning, Currency Recognition, Medicine Identification, Text-to-Speech.

I. INTRODUCTION

Visual impairment affects millions of people worldwide and limits their ability to perform daily activities independently. Tasks such as recognizing medicines and handling currency require accurate visual perception, which visually impaired individuals lack. Incorrect medicine identification may lead to serious health issues, while difficulty in handling money increases dependency and reduces confidence.

Recent advancements in artificial intelligence and computer vision enable the development of smart assistive systems. Convolutional Neural Networks (CNNs) provide accurate image classification, and Optical Character Recognition (OCR) allows extraction of printed text from images. Combined with voice interfaces, these technologies can assist visually impaired users in real time.

This project proposes an AI-based companion system for recognizing medicines and currency notes. The system captures images using a camera-enabled device, processes them using trained deep learning models, and delivers results through audio feedback. Voice command functionality ensures hands-free interaction. The solution is portable, affordable, and deployable on smartphones or embedded platforms, improving independence and safety for visually impaired users.

II. PROCEDURE FOR PAPER SUBMISSION

Authors must prepare the manuscript using the IEEE template and follow the prescribed format for headings, figures, tables, and references. The paper should be written in technical language with proper citation and plagiarism compliance. The manuscript is submitted in PDF format through the online system. After submission, the paper undergoes peer review. Based on reviewer feedback, revisions may be required. Upon acceptance, the camera-ready version and copyright forms are submitted for publication.

Review Stage
After submission, the paper is evaluated by reviewers for originality, technical quality, and relevance.

Authors may be asked to revise based on feedback before final acceptance.

Final Stage
Once the paper is accepted, authors prepare the camera-ready version by incorporating reviewer comments and ensuring IEEE formatting compliance. The final manuscript is submitted along with copyright forms and conference registration for publication.

III. PROPOSED SYSTEM

The proposed system consists of a camera module, deep learning models, OCR engine, voice interface, and text-to-speech unit. The camera captures images of medicines or currency notes placed in front of the user. The captured image is preprocessed to remove noise and enhance clarity.

For medicine recognition, CNN models analyze packaging features such as color, shape, and text. OCR extracts printed information like medicine names to improve accuracy. For currency recognition, the system classifies denominations based on unique visual patterns and symbols.

Once recognized, the system converts the output into speech using text-to-speech technology. Voice commands allow users to control the system without physical interaction, making it accessible and user-friendly.

IV. METHODOLOGY

The system begins with image acquisition using a camera-enabled device. The image is resized and normalized for processing. Feature extraction is performed using trained CNN models. OCR is applied to extract textual information from medicine labels.

Classification algorithms determine the medicine type or currency denomination. The result is then passed to the text-to-speech module, which generates audio output for the user. Voice recognition modules process user commands and control system operations.

The system is implemented on smartphones or Raspberry Pi to ensure portability and low cost.
How to create a PostScript File

Results and Discussion
The proposed system was tested with different medicine strips and currency notes under varying lighting conditions. The recognition accuracy was high for commonly used medicines and currency denominations. OCR successfully extracted medicine names from clear packaging.

Applications
Medicine identification for visually impaired users Currency recognition for financial independence Assistive smart devices Healthcare support systems Portable vision assistance tools

Advantages Easy to use Portable and affordable Real-time recognition Audio feedback support Hands-free interaction Improves safety and independence

Literature Survey
Several research works have focused on assistive technologies for visually impaired users. Early systems mainly used barcode scanning for medicine identification, which required precise alignment and was limited in usability. Later, image processing techniques were introduced for recognizing currency notes using handcrafted features such as edges, texture, and color histograms.

Recent advancements utilize deep learning models such as Convolutional Neural Networks (CNNs) for improved accuracy in object classification. OCR-based systems are also used for reading printed text on medicine labels. However, many existing solutions lack real-time performance, portability, and user-friendly voice interfaces. The proposed system overcomes these limitations by integrating CNN-based recognition with OCR and audio feedback in a portable and affordable platform

System Architecture

The system architecture consists of image acquisition, preprocessing, feature extraction, classification, OCR processing, and audio output modules. A camera captures the image of the object placed in front of the user. The image is preprocessed to remove noise and improve contrast. CNN models extract visual features for classification. OCR is applied to extract textual information from medicine packaging. The recognition result is then passed to the text-to-speech module to generate audio feedback. Voice command modules allow users to interact with the system hands-free.

Hardware and Software Requirements

Hardware Requirements:

- Camera-enabled device
- Raspberry Pi / Smartphone
- Microphone
- Speaker / Headphones
- Software Requirements:
- Python
- OpenCV
- TensorFlow / Keras
- Tesseract OCR
- Text-to-Speech Engine
- Speech Recognition API

Algorithm

- Capture image using camera.
- Preprocess the image (resize, noise removal).
- Extract features using CNN model.
- Apply OCR for text extraction.
- Classify medicine or currency denomination.
- Convert output to speech.
- Accept voice commands for next operation.

Performance Analysis

The system performance is evaluated based on recognition accuracy, response time, and usability. Tests were conducted under different lighting and background conditions. The CNN model achieved high accuracy for common medicines and currency notes. OCR performance was effective on clear packaging. The system provided fast audio responses, enabling real-time assistance. Overall, the solution demonstrated reliable performance and improved user experience.

Limitations

The system performance may reduce under poor lighting conditions or when medicine labels are damaged. OCR accuracy depends on image clarity. Similar-looking packaging may cause misclassification. Internet connectivity may be required for certain voice processing modules. These limitations can be addressed in future improvements.

Future Enhancements

Future work includes multilingual support, integration with wearable devices, cloud-based model updates, larger medicine databases, improved recognition under complex environments, and mobile app deployment. Advanced AI models can further enhance accuracy and usability.

Social Impact

The proposed system supports inclusive technology by empowering visually impaired individuals to perform daily tasks independently. It reduces dependency on caregivers, increases confidence, and improves safety in health and financial activities. The project contributes toward building smart and socially responsible assistive technologies.

Problem Statement

Visually impaired individuals face difficulties in independently identifying medicines and currency notes, which can lead to medication errors, health risks, financial dependency, and reduced confidence in daily activities. Existing assistive technologies are often expensive, limited in functionality, or not designed for real-time usability. Therefore, there is a need for an affordable, intelligent, and portable system that can accurately recognize medicines and currency notes and provide accessible audio feedback for visually impaired users.

Objectives of the Project

The primary objectives of the proposed system are:

- To develop an AI-based real-time medicine recognition system.
- To design an accurate currency denomination identification model.

- To integrate OCR for extracting printed information from medicine labels.
- To implement text-to-speech for audio feedback.
- To enable hands-free interaction using voice commands.
- To ensure affordability and portability using smartphones or embedded platforms.

Dataset Description

The dataset used in this project consists of images of commonly used medicine strips, bottles, and Indian currency notes of different denominations. Images were captured under various lighting conditions and backgrounds to improve model robustness. The medicine dataset includes packaging variations, brand differences, and label orientations. The currency dataset includes front and back images of notes to capture unique design patterns.

The dataset was divided into training, validation, and testing sets to evaluate model performance effectively.

Image Preprocessing Techniques

Before classification, images undergo preprocessing to improve recognition accuracy. The preprocessing steps include:

- Image resizing to a standard resolution
- Noise removal using filtering techniques
- Contrast enhancement
- Image normalization
- Background segmentation (if required)

These techniques improve feature extraction and ensure consistent input to the CNN model.

Deep Learning Model Description

The system uses Convolutional Neural Networks (CNNs) for feature extraction and classification. CNN layers include convolution layers, pooling layers, and fully connected layers. The convolution layers extract spatial features such as edges, textures, and shapes. Pooling layers reduce dimensionality and improve computational efficiency. Fully connected layers perform final classification.

The model is trained using labeled image datasets. Activation functions such as ReLU and Softmax are used for classification. The model is optimized using backpropagation and gradient descent algorithms.

OCR Integration

Optical Character Recognition (OCR) is integrated to extract text from medicine packaging. The OCR engine processes the detected text regions and converts image-based text into machine-readable format. This enhances reliability by verifying medicine names alongside visual classification. OCR improves accuracy especially when packaging designs are similar.

Speech Interface Module

The speech interface consists of:

- Speech Recognition Module (for voice commands)
- Text-to-Speech Module (for audio output)
- Speech recognition allows users to activate commands such as “Scan Medicine” or “Scan Money.” The text-to-speech module converts recognition results into natural spoken audio output, ensuring full accessibility without visual interaction.

Security and Privacy Considerations

The system processes images locally on the device to maintain privacy and prevent data leakage. No sensitive user data is stored without consent. Secure APIs are used for voice processing where required. Ensuring user privacy is critical in assistive technology applications.

Comparative Analysis

Compared to traditional barcode scanners and manual assistance systems, the proposed AI-based system provides:

- Real-time object recognition
- No need for barcode alignment
- Improved portability
- Audio-based accessibility
- Higher accuracy using deep learning

This makes the proposed system more flexible and user-friendly.

Evaluation Metrics

The system performance is evaluated using the following metrics:

- Accuracy
- Precision
- Recall
- F1-Score
- Response Time

High accuracy and low response time indicate effective real-time performance suitable for assistive applications.

Use Case Scenario

A visually impaired user places a medicine strip in front of the camera and gives a voice command to scan. The system captures the image, processes it through the CNN and OCR modules, and announces the medicine name through audio output. Similarly, when a currency note is scanned, the system identifies the denomination and announces the value instantly.

This workflow enables safe and independent daily functioning.

Experimental Setup

The system was implemented using Python programming language with OpenCV for image processing, TensorFlow/Keras for deep learning, and Tesseract for OCR. The hardware platform included a Raspberry Pi with a camera module and microphone-speaker interface. Testing was conducted under multiple lighting conditions to evaluate robustness.

Innovation and Contribution

The major contributions of this project include:

- Integration of CNN and OCR for dual verification
- Real-time voice-controlled assistive system
- Affordable implementation using low-cost hardware
- Focus on accessibility and inclusive design

Mathematical Model

The medicine and currency recognition system can be mathematically represented as a classification problem.

Let the input image be represented as:

$$I \in \mathbb{R}^{H\times W\times C}$$

where H is height, W is width, and C is the number of color channels.

The CNN model performs feature extraction:

$$F = f(I;\theta)$$

where f represents the convolutional operations and θ represents trainable parameters.

The classification output is given by:

$$Y = \text{Softmax}(W_f F + b)$$

where W_f and b are weight and bias parameters.

The predicted class is:

$$\hat{y} = \arg\max(Y)$$

This mathematical formulation represents the deep learning classification process used in the system.

Confusion Matrix Analysis

The performance of the classification model is evaluated using a confusion matrix. The matrix consists of:

- True Positives (TP)
- True Negatives (TN)
- False Positives (FP)
- False Negatives (FN)

Accuracy is calculated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision, Recall, and F1-Score are also computed to evaluate model effectiveness.

Implementation Challenges

During development, several challenges were encountered:

- Variations in lighting conditions
- Similar packaging designs among medicines
- Worn or folded currency notes
- OCR errors due to blurred text
- Background noise affecting voice recognition

These challenges were addressed using preprocessing techniques, data augmentation, and optimized deep learning models.

Data Augmentation Techniques

To improve model robustness, the following augmentation techniques were applied:

- Image rotation
- Brightness adjustment
- Zooming
- Horizontal flipping
- Noise addition

These techniques increase dataset diversity and reduce overfitting.

Real-Time Processing Considerations

The system is optimized for real-time performance by:

- Reducing model size
- Using lightweight CNN architectures
- Optimizing image resolution
- Running inference locally on device
- This ensures fast response time and smooth user experience.

Ethical Considerations

The system is designed with ethical responsibility by:

- Ensuring user privacy
- Avoiding misuse of captured images
- Promoting accessibility and inclusivity
- Maintaining transparency in AI decision-making

Assistive technologies must prioritize user dignity and safety.

Acknowledgment

The authors would like to thank the faculty members and project mentors for their guidance and support throughout the development of this project. Special gratitude is extended to the institution for providing the necessary resources and infrastructure.

Author Biography

Akash SN is currently pursuing a Bachelor of Technology in Computer Science and Business Systems. His research interests include Artificial Intelligence, Computer Vision, Assistive Technologies, and Embedded Systems. He aims to develop innovative AI-based solutions that contribute to social welfare and inclusive technology development.

V. CONCLUSION

An AI-based companion system for medicine and currency recognition offers a practical and empowering solution for visually impaired users. By leveraging technologies such as computer vision, machine learning, optical character recognition (OCR), and speech synthesis, the system can accurately identify medicines, read labels, detect expiry dates, and recognize different currency denominations in real time.

This type of assistive technology enhances independence, safety, and confidence in daily activities—reducing the risk of medication errors and financial confusion. It also promotes social inclusion by enabling users to manage personal health and transactions without constant assistance from others.

With continuous advancements in artificial intelligence and mobile computing, such systems can become more accurate, affordable, and widely accessible. Ultimately, an AI-based companion system represents a meaningful step toward improving quality of life and ensuring greater autonomy for visually impaired individuals.

REFERENCES

1. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, May 2015.
2. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in Proc. Advances in Neural Information Processing Systems (NIPS), 2012, pp. 1097–1105.
3. R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th ed. New Delhi, India: Pearson Education, 2018.
4. S. Haykin, Neural Networks and Learning Machines, 3rd ed. New York, NY, USA: Pearson, 2009.
5. R. Smith, “An overview of the Tesseract OCR engine,” in Proc. Int. Conf. Document Analysis and Recognition (ICDAR), 2007, pp. 629–633.
6. O. Russakovsky et al., “ImageNet large scale visual recognition challenge,” Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, Dec. 2015.
7. N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2005, pp. 886–893.
8. A. Howard et al., “MobileNets: Efficient convolutional neural networks for mobile vision applications,” arXiv:1704.04861, 2017.

9. World Health Organization (WHO), "World report on vision," Geneva, Switzerland, 2019.
10. S. K. Nayar and T. Mitsunaga, "High dynamic range imaging: Spatially varying pixel exposures," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2000, pp. 472–479.
11. D. Cireșan, U. Meier, J. Masci, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2012, pp. 3642–3649.
12. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.
13. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017.
14. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788.
15. A. Graves, A. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP), 2013, pp. 6645–6649.
16. M. Abadi et al., "TensorFlow: A system for large-scale machine learning," in Proc. 12th USENIX Symp. Operating Systems Design and Implementation (OSDI), 2016, pp. 265–283.
17. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2016, pp. 785–794.
18. L. Deng and D. Yu, "Deep learning: Methods and applications," Foundations and Trends in Signal Processing, vol. 7, no. 3–4, pp. 197–387, 2014.
19. A. Rosebrock, Practical Python and OpenCV. PyImageSearch, 2018.
20. M. Everingham et al., "The Pascal visual object classes (VOC) challenge," Int. J. Comput. Vis., vol. 88, no. 2, pp. 303–338, Jun. 2010.
21. S. Theodoridis and K. Koutroumbas, Pattern Recognition, 4th ed. London, U.K.: Academic Press, 2009.
22. C. Szegedy et al., "Going deeper with convolutions," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1–9.
23. F. Chollet, Deep Learning with Python. New York, NY, USA: Manning Publications, 2018.
24. A. Krizhevsky, "Learning multiple layers of features from tiny images," M.S. thesis, Univ. of Toronto, Toronto, ON, Canada, 2009.
25. J. Liang, D. Doermann, and H. Li, "Camera-based analysis of text and documents: A survey," Int. J. Doc. Anal. Recognit., vol. 7, no. 2–3, pp. 84–104, 2005.
26. G. Bradski, "The OpenCV library," Dr. Dobb's Journal of Software Tools, 2000.
27. S. Amershi et al., "Guidelines for human-AI interaction," in Proc. ACM Conf. Human Factors in Computing Systems (CHI), 2019, pp. 1–13.
28. J. K. Aggarwal and M. S. Ryoo, "Human activity analysis: A review," ACM Comput. Surv., vol. 43, no. 3, pp. 1–43, 2011.
29. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2001, pp. 511–518.
30. T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern Recognit., vol. 29, no. 1, pp. 51–59, 1996.