Review of Computer Engineering Research
2025 Vol. 12, No. 3, pp. 182-194
ISSN(e): 2410-9142
ISSN(p): 2412-4281
DOI: 10.18488/76.v12i3.4438

EAPR-Net: A lightweight framework for voice-assisted pill recognition in visually impaired users

Rashmi R. Shrirao[1+]
Shubhangi M. Handore[2]

[1,2]Department of Electronics and Telecommunication, Trinity College of Engineering and Research, Pune, India.
[1]Email: rashmideulkar1186@gmail.com
[2]Email: hodentc.tcoer@kjei.edu.in

(+ Corresponding author)

Article History
Received: 29 May 2025
Revised: 23 August 2025
Accepted: 15 September 2025
Published: 2 October 2025

Keywords
Android application
CNN
Contrast enhancement
Drug pill recognition
EAPR-net
REST API
Voice assistance.

ABSTRACT

Elderly individuals and visually impaired people are often at risk in managing their medication since they forget to take their medication, experience memory loss, deteriorating eyesight, and difficulty distinguishing between different pills. Misidentification of drugs can result in adverse health effects, and thus, accessible solutions are essential. To solve these problems, we propose a novel framework, called Enhanced Accessibility Pill Recognition Network (EAPR-Net), embedded in an Android mobile application. In addition to a CNN-based model, models have previously been proposed to overcome challenges like variability in lighting (through contrast enhancement), limits in power consumption with a network that is lighter than traditional CNNs for working on mobile devices, and alert feedback features with a voice assistant made possible through REST API. Users need only take a picture of their pills with a smartphone and receive immediate, audible, hands-off identification. The model was developed and tested on a synthetic dataset of 900 images from 14 common categories of pills, achieving a remarkable 98% test accuracy. The design of EAPR-Net not only ensures computational efficiency but also guarantees high classification accuracy for diverse real applications. By combining advanced computer vision methodologies with user-friendly accessibility features, this framework advances medication safety to a new level. Finally, EAPR-Net enables the elderly and visually impaired to become self-sufficient in prescription management and less dependent on caregivers, resulting in an overall better quality of life.

Contribution/Originality: This study exclusively combines contrast-enhanced preprocessing, a lightweight CNN tailored for mobile efficiency, and real-time voice-assisted feedback in a single application, contributing an integrated, accessibility-focused solution for pill recognition that has not been previously implemented or evaluated in this form.

1. INTRODUCTION

As individuals age, both physical and cognitive faculties, such as memory and vision, tend to decline, posing substantial risks in everyday activities like medication management. The consequences of misidentifying or missing medications can lead to serious health complications, particularly for elderly individuals and visually impaired persons. Several studies have emphasized that errors in drug identification remain a significant contributor to patient harm and healthcare inefficiency [1, 2]. Despite advancements in medical support technologies, a clear gap persists in providing accessible, reliable solutions tailored specifically for these vulnerable populations.

Healthcare systems, especially in regions with aging demographics like Europe and Asia, are under increasing pressure to manage growing patient needs without proportional increases in personnel resources [3]. Expecting healthcare professionals to oversee routine tasks such as daily medication identification for each patient is impractical. Therefore, there is a critical need for assistive technologies that empower individuals to independently manage their prescriptions without external dependency.

Recent advances have highlighted the importance of mobile healthcare technologies for elderly care [4] and AI-enabled accessibility solutions for visually impaired individuals [5].

In response to this gap, this paper proposes a novel approach termed the Enhanced Accessibility Pill Recognition Network (EAPR-Net). Unlike traditional systems that primarily focus on general image classification, EAPR-Net integrates contrast-enhanced image preprocessing, a lightweight Convolutional Neural Network (CNN) optimized for mobile deployment, and a real-time voice-assisted feedback system. This unique combination is designed to ensure high accuracy in pill recognition under diverse real-world conditions while maintaining ease of use for visually impaired users.

By incorporating EAPR-Net into an Android application, the system allows users to capture a pill image using their smartphone camera and instantly receive audible feedback regarding the identified medication. This initiative aims to reduce medication errors, foster greater independence among users, and bridge the accessibility gap in healthcare technology.

The remainder of this paper is structured as follows. Section 2 provides a comprehensive review of existing literature on pill recognition technologies. Section 3 details the proposed methodology, including dataset preparation, preprocessing, and model design. Section 4 presents experimental results and performance analysis. Section 5 describes the integration of the system into an Android application. Finally, Section 6 concludes the paper and discusses future research directions.

2. LITERATURE REVIEW

For medication pill recognition, a number of publications employ computer vision and machine learning approaches. This section covers some of the most recent and innovative reviews of the literature in this field.

The medicine pill recognition system presented by Rádli, et al. [1] uses a multi-stream, two-phase metric embedding neural model. It improves recognition accuracy by employing a dynamic margin adjustment in the loss function to incorporate both textual information from medication pamphlets and visual aspects of tablets. Medical pill verification and other object identification tasks can benefit from this method, which overcomes issues such as different viewing circumstances and pill similarity, with 1.6% and 2.89% increases in top-1 accuracy on the Cure dataset, respectively [6].

A Roge et al. [7] Drug Pill Recognition System (DPRS) automatically recognizes and categorizes drugs using artificial intelligence, particularly computer vision and neural networks. By improving their accuracy and reducing the possibility of prescription errors, this innovative solution addresses the challenges faced by visually impaired individuals in managing their medications. The DPRS enhances daily health routines for people with visual impairments by providing a reliable method for correctly identifying tablets through advanced image recognition technology [7].

Boonthep, et al. [2] presented a drug picture classification system for identifying drug tablets that uses the Fast Region-based Convolutional Neural Network (Fast R-CNN). It draws attention to the difficulties presented by visually identical pills and suggests a deep learning method for classifying them according to attributes including packing, colour, shape, and markings. By using convolution, pooling, and fully connected layers to learn spatial hierarchies, the system has demonstrated its efficacy in drug recognition by recognizing 20 distinct types of medicines with over 98% accuracy [2].

Ashraf et al. [8] used a collection of 26,880 photos of the top 30 solid oral dosage forms to create a pill recognition model using code-free deep learning (CFDL) on the Microsoft Azure Custom Vision platform. During internal testing, the model's performance metrics were strong, with 98.7% accuracy and 95.1% recall. It demonstrated that employing AI-based systems to recognize pills in healthcare settings is feasible, highlighting the importance of customizing models to regional formularies for optimal practical applicability [8].

A smart pill detection system that utilizes advanced deep learning techniques specifically, the MobileNet architecture for drug pill recognition was presented by Dhayanithi et al. [9]. By automating drug detection from photographs, this system overcomes the challenges associated with manual identification methods. The system aims to achieve high accuracy and efficiency by employing a diverse dataset of pill samples, ultimately enhancing patient safety and optimizing the time and resources of healthcare professionals in medication management [9].

In order to accurately identify tablets and capsules in medical photographs, Tenneti et al. [10] describe a deep learning framework that utilizes the YOLO object detection technique. By initiating an alarm mechanism upon detecting a pill, this device aims to prevent confusion in the pharmaceutical sector. It enhances automation and reliability in drug detection, manufacturing, quality control, and medication identification operations by employing rigorous training on annotated datasets to recognize distinctive visual characteristics of prescription tablets [10].

An automatic pill recognition system that makes use of deep learning and computer vision techniques was proposed by Ponte et al. [3]. In addition to image preparation using an image data generator, it employs a deep learning model built with Keras. To examine important pill characteristics, including shape, color, and imprint, the system utilizes OpenCV and Paddle OCR. The ultimate goal is to develop a real-time identification system that works with video cameras to improve drug dispensing efficiency and accuracy [3].

In their discussion of the application of Convolutional Neural Networks (CNN) for medication identification, Pilania, et al. [11] focus on identifying tablets based on their size, color, shape, and imprint. When trained on a Kaggle dataset of drug images, the CNN model outperformed traditional image recognition techniques. Automating the drug identification process enhances efficiency and accuracy in the pharmaceutical sector, which is crucial for distinguishing between genuine and counterfeit drugs [11].

An enhanced YOLOv8s-based model for drug detection and identification is presented by Zhu et al. [12], addressing issues such as class imbalance and small target sizes. With an accuracy of 95.1% and a mAP@50 of 87.4%, this model outperforms the original YOLOv8s by 2.2% in mAP and 3.0% in precision. The improvements demonstrate the potential of deep learning and computer vision in developing effective drug pill identification systems by significantly increasing detection accuracy and recognition rates [12].

Using Multi-Combination Pattern Labelling (MCPL), Svetlana Kim et al. introduce a unique pill categorization system that incorporates deep learning approaches to improve medication pill identification. This technique efficiently extracts feature points that are scale- and rotation-invariant, concentrating on the distinctive edges and curves of tablets. To increase recognition performance and improve pill identification accuracy and reliability both of which are critical for patient safety and medical efficiency the system addresses issues with 2D imagery and limited datasets [13].

Two major processes make up the image retrieval technique used in the work by Al-Hussaeni, et al. [14]: a preprocessing phase that involves feature extraction and a classification phase. The goal of this approach is to improve the efficiency and accuracy of pill identification using photos taken under less-than-ideal circumstances. The study proposes three neural network architectures: one ResNet-50 network and two hybrid models that integrate convolutional neural networks (CNN) with classification techniques (CNN+SVM and CNN+KNN). When these models are tested on a real-world dataset from the National Library of Medicine, they achieve 90.8% accuracy, and the CNN+KNN architecture outperforms current models by 10% [14].

Drugs are identified by Siripraiwan et al. [15] using machine learning and image recognition methods based on their outward characteristics, including colour, size, and form. The goal of this strategy is to reduce drug abuse caused by ambiguous prescriptions and fading labelling. The support vector machine, a classification model, is trained using characteristics taken from pictures of ten distinct kinds of medications. The model showed a 100% accuracy rate for detecting new pharmaceutical photos collected under normal settings and an average test fold accuracy of 94% during 5-fold cross-validation [15].

EfficientNet-B0 and a self-attention mechanism are included in a multi-stream network proposed by Rádli et al. [1] to improve pill detection capabilities. Local Binary Patterns (LBP) features are utilized in the method to enhance identification performance without requiring explicit training of printed or embossed patterns. Regarding pill recognition, the suggested multi-stream network with EfficientNet-B0 and a self-attention mechanism outperformed earlier models in terms of both Top-1 and Top-5 accuracy. The model also performed better than the YOLOv7 network in some test settings, demonstrating its effectiveness in a reference quality use-case [1].

Pallavi, et al. [16] investigated how deep learning and machine learning methods may be used to improve pill recognition in order to increase patient safety and medication compliance in the healthcare sector. By addressing the drawbacks of human judgment and traditional manual methods in pill identification—which can be laborious and error-prone, the study highlights the necessity of automated solutions to avoid medication mistakes and guarantee correct prescription drafting [16].

Pilania et al. [11] concentrated on the use of faster recurrent convolutional neural networks (FR-CNN), which are intended to extract information from drug pictures. This technique makes it possible to examine the medications' patterns and extract the deep information required for precise identification. In order to classify pharmaceuticals, the extracted characteristics are compared with ground truth labels. This allows the system to distinguish and identify different medications, which helps to solve the problem of chronic patients using the wrong medications [17].

Using deep learning techniques, S. Prabu presented a novel framework for real-time pill detection and identification. The suggested framework consists of three primary models: the text detection model finds the text information in the pill strip picture; the text recognition module identifies the localised text; and the first model uses the YOLOv5 to identify the pill strip. The suggested structure benefits considerably from the recognition module's ability to detect textual data like the name of the pill, its expiration date, its price, and so forth. The training and test photos in this study are taken from different angles and with different amounts of light. Even with less-than-ideal image quality, the suggested framework achieved a high recognition accuracy [18].

For the purpose of quickly and easily identifying different types of medications, Sakthimohan et al. [19] suggested a trained model that primarily uses Keras and TensorFlow. The pill database, where the pill name is discovered, is accessed by the detected pill (object detection). Following the detection phase, the pill is identified using the pre-trained dataset. Additionally, the dataset contains the necessary precise information and use cases for the specific drug. The project involves gathering data for automated medical detection systems. The experimental findings confirm the effectiveness of the suggested approach [19].

A detection system based on the B/S research application for pill box recognition was implemented by Xiang et al. [20] using a convolutional recurrent neural network (CRNN) as the text recognition framework and DBNet as the text detection framework. They also proposed an end-to-end graphical text detection and recognition model. For the detection and recognition procedures, no previous picture preparation was necessary. The front-end display was updated with the back-end recognition result. This recognition technique makes the model application simpler and the preprocessing before picture detection less complicated than with standard methods. Testing on 100 pill boxes shows that the suggested approach outperforms the earlier CTPN + CRNN technique in terms of text localization and recognition accuracy. Compared to the conventional way, the suggested methodology was much more user-friendly and accurate in terms of both training and recognition procedures [20].

In order to reduce medication mistakes by correctly detecting prescription medications, Heo et al. [21] devised a deep learning-based approach. Our algorithm finds the tablets in the corresponding pill databases in the US and Korea based on pill photos. The system is built by the author using two steps: pill recognition and pill retrieval.

Specifically, the pill recognition stage comprises modules that independently identify the three characteristics of tablets and their imprints, then adjust the identified imprint to match the real data. To fix imprint characters, the suggested system uses a language model as an imprint correction module. Using similarity scores between the pill's attributes and those in the database, the author is able to identify the pill. According to the testing findings, the suggested approach can identify unknown tablets in two separate databases with top-1 accuracy scores of 85.6% and 74.5%. Additionally, by training just one image per pill, the algorithm attains 78.0% top-1 accuracy with consumer photographs [21].

The literature review highlights several advancements in pill recognition techniques using computer vision and deep learning; however, notable gaps and limitations still persist. Firstly, although models such as Fast R-CNN, YOLO-based frameworks, and customized CNNs have achieved high accuracy in pill classification, most of them largely focus on technical performance without addressing the crucial aspect of accessibility for elderly or visually impaired users [2, 11]. The absence of integrated audio feedback and intuitive user interaction in existing models limits their usability among the targeted vulnerable population.

Secondly, while accuracy metrics under controlled conditions are promising, real-world scenarios involving varying lighting, backgrounds, and noise levels are rarely addressed through dedicated preprocessing techniques. Studies often neglect the enhancement of image quality through contrast optimization or noise reduction, thereby compromising the robustness of pill recognition systems under diverse conditions [1, 3].

Another prominent limitation is the computational complexity of existing architectures, which restricts their deployment on resource-constrained mobile devices. Many deep learning models require high-end hardware for real-time performance, making them impractical for widespread use among non-technical users [10]. Furthermore, real-time integration involving smartphone-based image acquisition, cloud-server communication via REST API, and immediate voice feedback remains underexplored in prior research efforts.

In addressing these challenges, the proposed (EAPR-Net) contributes a contrast-enhanced preprocessing module to handle environmental variations, a lightweight CNN architecture optimized for mobile devices, and a fully integrated real-time system providing voice-based assistance. This user-centered approach bridges the gap between high-performance recognition models and practical, accessible healthcare applications for the visually impaired and senior citizens.

3. METHODOLOGIES

The overall systematic design of the proposed voice-assisted drug pill recognition system is illustrated in Figure 1. It outlines the key components, including image acquisition, contrast-enhanced preprocessing, EAPR-Net-based pill classification, REST API communication, and voice-based feedback delivery. A detailed explanation of each module is provided in the subsequent subsections.

3.1. Data Collection

Images have been taken from commercially accessible pharmaceutical datasets or from existing pharmaceutical databases, such as the FDA database. The name of the pill, the dose, the manufacturer, and other important details are noted on each image. This study's dataset was obtained from Kaggle [20]. A total of 692 annotated photos make up the collection. Rotations, flips, and brightness variations are examples of data augmentation techniques used here to expand the diversity of training samples and improve the model's generalization. After data augmentation, the entire dataset comprises 900 photos.

3.2. Image Processing

Additionally, before inserting the photos into the model, it ensures that they are of uniform quality and that any noise or extraneous background information has been eliminated. To standardize image quality, contrast enhancement is performed, particularly for photos taken in different lighting conditions. The Pillow library in Python is used for enhancing contrast. A series of contrast levels (e.g., 1.5, 2.0, 3.0, 4.0) are defined. Contrast enhancement techniques have been proven highly effective for improving recognition accuracy under variable lighting conditions, as demonstrated in recent studies Tian et al. [22]. These values represent the multiplier by which the contrast will be adjusted. It was found that a contrast level of 1.5 provides the most balanced enhancement. It improves the distinction between different shades of color in the image while maintaining natural details and avoiding harsh transitions.

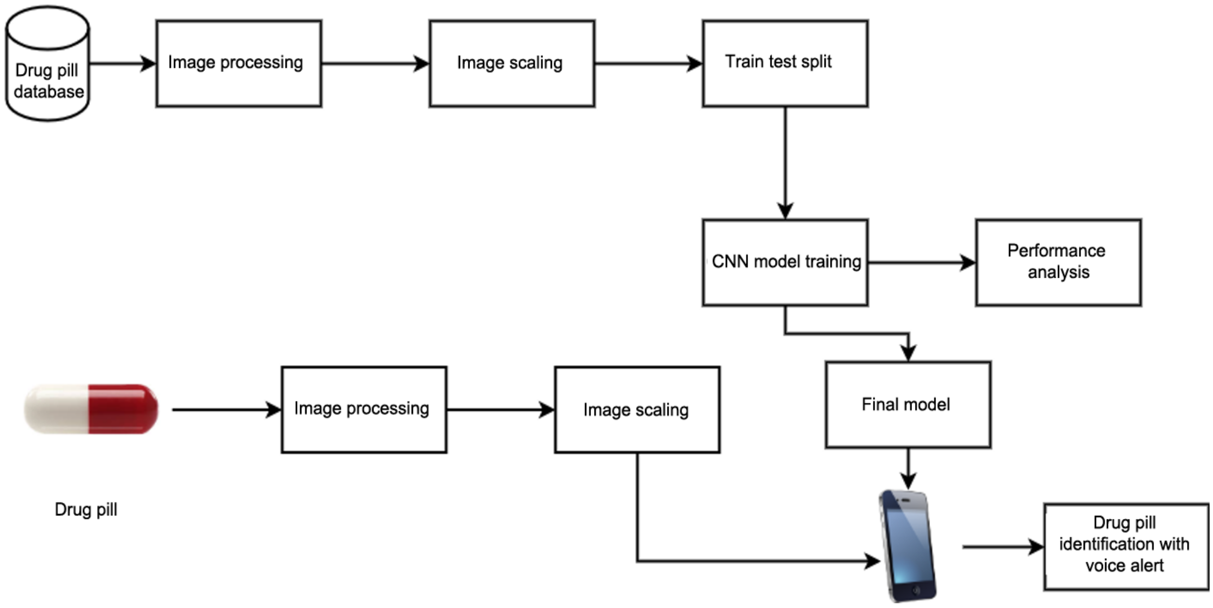

Figure 1. Architecture of proposed system.

3.3. Image Scaling

CNN model input dimensions are constant when all pictures are scaled to the same size, which is essential for efficient training. The resizing of images to a standard resolution, such as 224x224 pixels, is common for CNNs used in image recognition. The model's capacity to capture adequate detail and computational efficiency are balanced in the selection of this dimension.

3.4. Normalization

Pixel values are normalized (often between 0 and 1) after resizing in order to guarantee consistency across pictures and enhance model convergence during training. By lowering the variance in the input data, this normalization aids the CNN model in processing pictures more efficiently.

3.5. Train-Test Split

This stage is crucial for assessing the model's generalizability to new data, which is necessary for practical uses. Training and test sets make up approximately 70% and 30% of the dataset, respectively. The model is trained on the training set, and its performance is assessed on the test set.

3.6. EAPR-Net Architecture

The core classification model EAPR-Net is a lightweight Convolutional Neural Network (CNN) specifically designed for efficient mobile deployment and high recognition accuracy. The network comprises sequential convolutional, pooling, regularization, and fully connected layers optimized to capture essential pill features such as color, shape, and imprint while minimizing computational overhead.

The architecture of EAPR-Net includes:

1. Input Layer

- The model takes input photos that have been resized to 224×224×3, where 3 represents the RGB color channels, and 224 is the height and width.

2. Convolutional Layers

- There are three convolutional blocks in the design. Progressively, more complex features are extracted from the input images by each block.
- First Block:
- A 2D convolutional layer (Conv2D) is used, which has 32 3x3 filters.
- To provide non-linearity, the Rectified Linear Unit (ReLU) activation function is employed.
- Using a 2x22 max pooling procedure (MaxPooling2D), the spatial dimensions are halved.
- Second Block:
- A Conv2D layer including 64 3x33 filters.
- After applying a ReLU activation function, a second 2x22 max pooling operation is performed.
- Third Block:
- A Convolutional 2D layer with 128 3x3 filters.
- After applying a ReLU activation function, 2×2 9max pooling is performed.
- To lower the chance of overfitting, 30% of the neurons are randomly deactivated using a dropout rate of 0.3.

3. Flattening Layer

- The Flatten layer prepares the data for dense layers by flattening the convolutional layers' output into a one-dimensional vector.

4. Fully Connected Layers

- To capture high-level feature interactions, a dense layer comprising 128 neurons is used, followed by the ReLU activation function.
- To improve generalization and reduce overfitting, 50% of the neurons are deactivated during training using a Dropout layer set to 0.5.

5. Output Layer

- The output layer has as many neurons as there are classes in the dataset.
- To ensure that the total of the probabilities for every class equals 1, the output is converted into probabilities using the Softmax activation function.

6. Regularization Techniques

- Dropout layers: introduced to avoid overfitting after dense and convolutional layers.
- Data Augmentation: used to improve the generalization of the model in the preprocessing stage. Zooming, width and height changes, random rotations, and horizontal flips are examples of augmentation techniques.

7. Training Configuration

- Loss Function: For situations involving multi-class classification, categorical cross-entropy works well.
- Optimizer: The Adam optimizer is renowned for its effectiveness when working with large datasets and adaptive learning rates.
- Evaluation Metrics: During training, the model is assessed using accuracy and loss.

8. Training Procedure

EAPR-Net was trained over 20 epochs with a batch size of 16, utilizing the augmented dataset. Performance was validated using a separate testing set to avoid overfitting and ensure the network's generalization capabilities. The detailed layer configuration is summarized in Table 1.

Table 1. Summary of EAPR-Net Layers.

Layer type	Filters/Units	Kernel size	Activation	Pooling size	Dropout
Conv2D + ReLU	32	3×3	ReLU	2×2	-
Conv2D + ReLU	64	3×3	ReLU	2×2	-
Conv2D + ReLU	128	3×3	ReLU	2×2	0.3
Flatten	-	-	-	-	-
Dense + ReLU	128	-	ReLU	-	0.5
Dense + Softmax	Num classes	-	Softmax	-	-

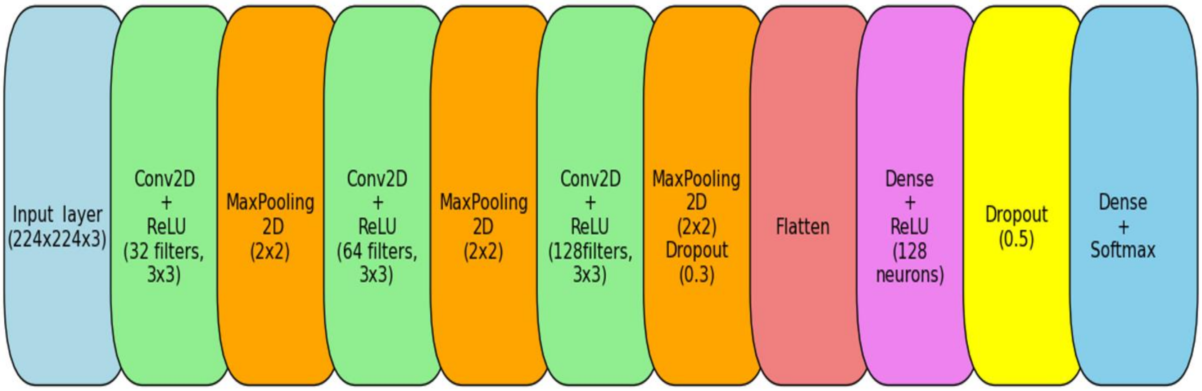

Figure 2. EAPR-net architecture.

Using contemporary CNN approaches to accurately classify prescription pill pictures while reducing overfitting through dropout layers and data augmentation, this architecture is selected for its simplicity and efficacy. Designing lightweight CNN architectures optimized for mobile applications has been recognized as a critical factor for efficient on-device computation [22]. Deploying deep learning models efficiently on edge devices such as smartphones remains a significant research challenge, especially for real-time medical image analysis tasks [23]. Figure 2 shows the visually represented design of the EAPR-Net architecture.

4. RESULTS AND DISCUSSION

This section presents the evaluation results of the proposed EAPR-Net on the prepared pill image dataset. The dataset, after augmentation, comprised 900 pill images spanning 14 distinct categories. Initially, the dataset underwent a variety of image enhancement techniques, including 3×3 box filtering, contrast enhancement with a multiplier of 1.5, sharpness adjustment with a strength of 1.5, and edge contour enhancement using the Canny method. Among these, contrast-enhanced images provided the best visual clarity for distinguishing pills, significantly improving model learning.

The performance of EAPR-Net across these different enhancement methods is summarized in Table 2.

Table 2. Image enhancement technique impact on accuracy.

Image enhancement techniques	Train accuracy (%)	Test accuracy (%)
Box filter	93%	76%
Contrast	100%	98%
Sharpness	100%	98%
Contour	94%	90%

The impact of various image enhancement methods on model performance is shown in Table 2.

Since it has been noted that CNN models with a contrast enhancement technique of 1.5 level offer the greatest accuracy in their class, the author has employed the same model to examine the findings in more detail. 30% of the dataset is used for testing, while 70% is used for training. The details of the train-test split are provided in Table 3.

Table 3. Dataset description.

Model	Train dataset	Test dataset
EAPR Net	620	280

While the model's testing accuracy of 98% demonstrated strong generalization to new, unknown data, its training dataset accuracy of 100.00% demonstrated the model's capacity to learn patterns from the data. The classification report of EAPR-Net across all 14 classes is summarized in Table 4.

Table 4. Classification report.

Class	Precision	Recall	F1-Score	Support
Amoxicillin 500 mg	1.00	1.00	1.00	11
Atomoxetine 25 MG	1.00	1.00	1.00	12
Calcitriol 0.00025 mg	1.00	1.00	1.00	11
Oseltamivir 45 MG	1.00	0.82	0.90	11
Ramipril 5 mg	1.00	1.00	1.00	08
apixaban 2.5 MG	1.00	1.00	1.00	11
aprepitant 80 MG	1.00	1.00	1.00	13
benzonatate 100 MG	1.00	1.00	1.00	12
carvedilol 3.125 MG	0.85	1.00	0.92	11
celecoxib 200 MG	1.00	0.90	0.95	10
duloxetine 30 MG	1.00	1.00	1.00	11
eltrombopag 25 mg	1.00	1.00	1.00	10
montelukast 10 MG	0.91	1.00	0.95	10
mycophenolate mofetil 250 mg	1.00	1.00	1.00	12
pantoprazole 40 mg	0.91	1.00	0.95	10
pitavastatin 1 mg	1.00	1.00	1.00	09
prasugrel 10 MG	1.00	1.00	1.00	11
saxagliptin 5 mg	1.00	1.00	1.00	09
sitagliptin 50 MG	1.00	0.91	0.95	11
Tadalafil 5 mg	1.00	1.00	1.00	10
Accuracy			0.98	213
Macro avg	0.98	0.98	0.98	213
Weighted avg	0.98	0.98	0.98	213

- Precision: evaluates the precision of each class's (pill type) positive predictions. An accuracy of 1.00 indicates that every forecast for that particular pill type is accurate.
- Recall: shows how well the model can recognize every instance of a certain type of drug. A recall of 1.00 indicates that every genuine case is detected by the model.
- F1-Score: The model reflects the balance of accuracy through the harmonic means of precision and recall.
- Support: The number of cases for every kind of medication.

Precision, Recall, and F1-Score metrics all demonstrate a consistently high level of performance across most classes, indicating robust generalization of the EAPR-Net model. The evolution of training accuracy and loss over the training epochs is depicted in Figure 3.

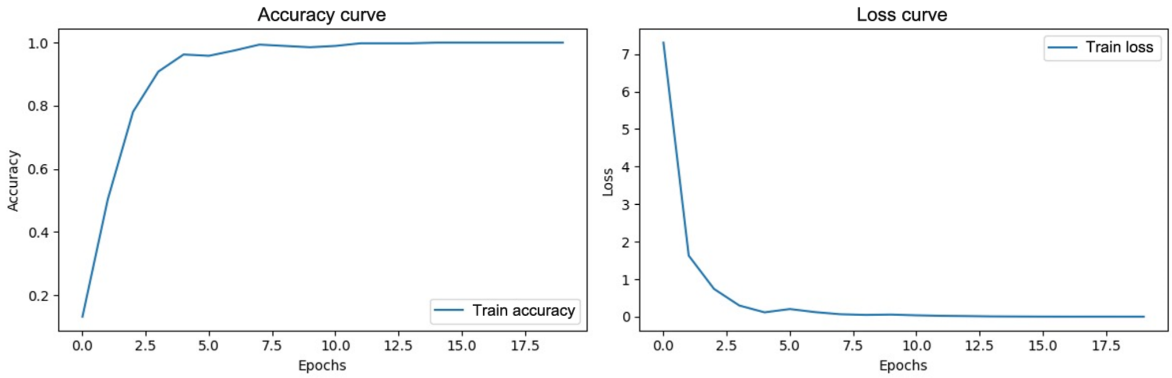

Figure 3. (a) Accuracy and (b) Loss curve for EAPR-Net.

The model's accuracy curve is displayed in Figure 3(a). The evolution of training accuracy across 20 epochs is depicted in this Model Accuracy figure. The training accuracy (blue line) begins poorly at first but gradually rises to about 95% by the tenth epoch. The model is successfully learning from the training data, as evidenced by this steady increase. Figure 3 (b) displays the model's loss curve. The model's ability to reduce mistakes is demonstrated by this Model Loss plot, which displays the training loss across 20 epochs. The training loss (blue line), indicating that the model successfully learns from the training data, begins high and gradually decreases until stabilizing at epoch 5. The confusion matrix, demonstrating class-wise prediction performance, is shown in Figure 4.

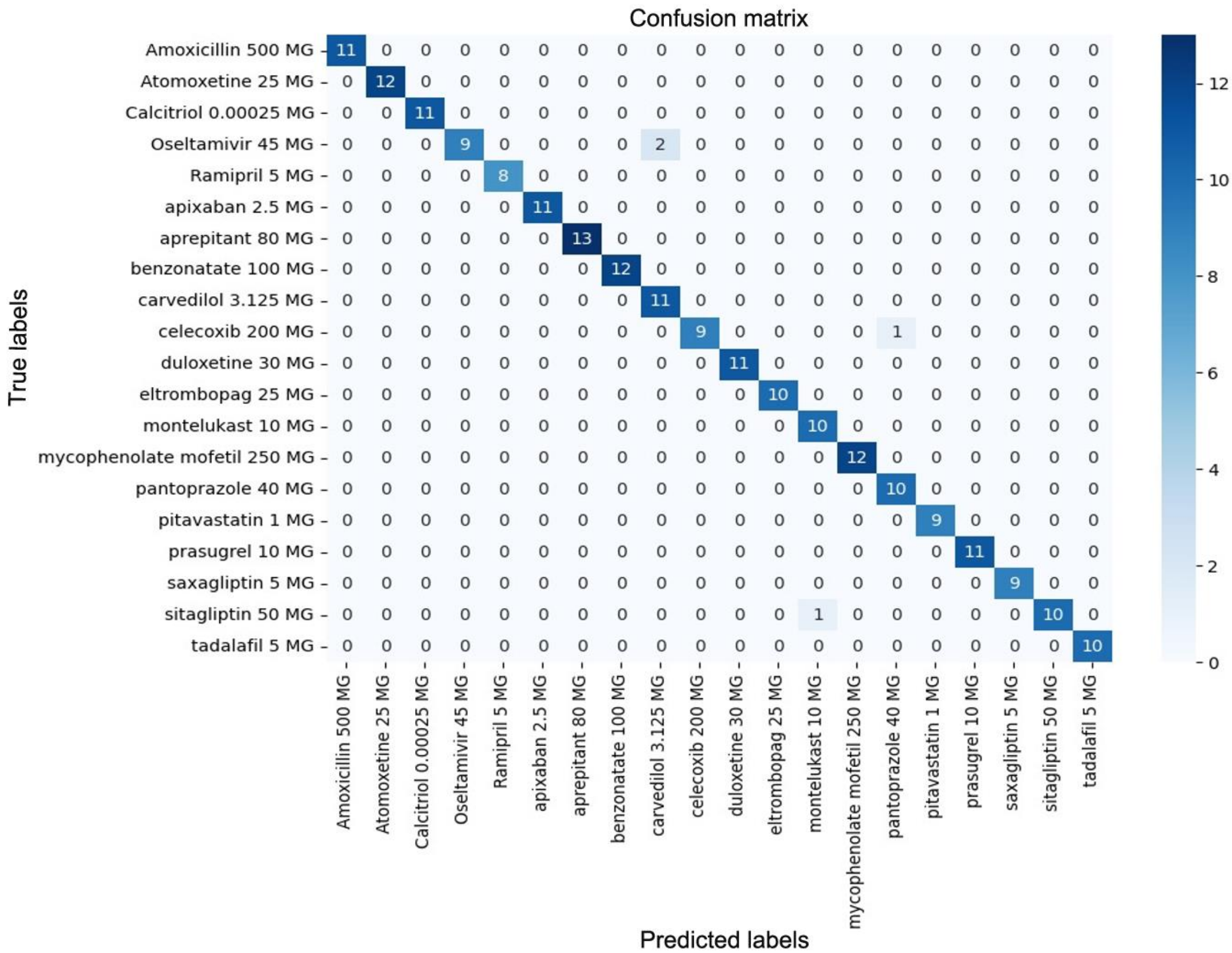

Figure 4. Confusion matrix of EAPR-net prediction.

4.1. Android Front End

To ensure real-world applicability, EAPR-Net was integrated with an Android application through a REST API. Users can capture pill images with their smartphone cameras, which are then classified by the EAPR-Net model hosted on a cloud server. The results, including the recognized pill name, are displayed and simultaneously announced through a voice output system, significantly enhancing accessibility for visually impaired users. The program is very accessible for visually impaired users since it shows the findings, including the medicine name and image, and offers voice aid to pronounce the recognized pill name. Voice-assisted mobile healthcare applications have been increasingly recognized as a vital tool for improving accessibility among visually impaired users [24]. By providing accurate, inclusive, and useful functionality for a variety of user demographics, this user-friendly design guarantees smooth interaction, lowers the possibility of medication mistakes, and improves healthcare assistance. An example of the Android application's output is presented in Figure 5.

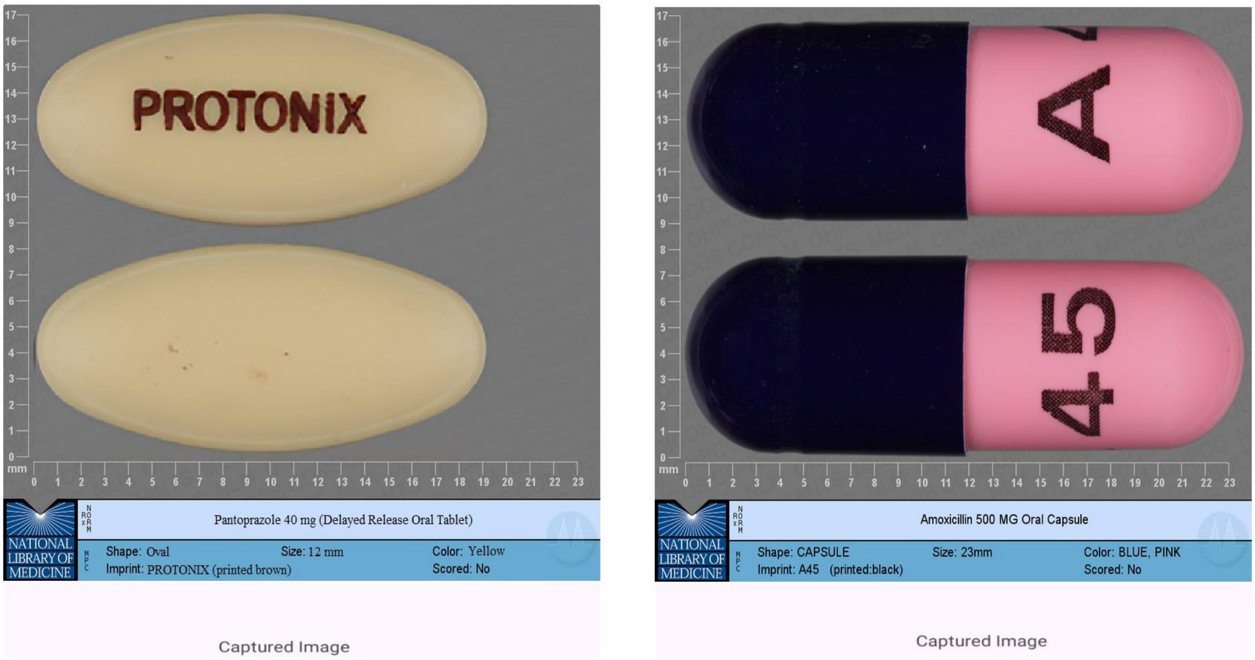

Figure 5. Android App results for drug pill identification using EAPR-Net.

5. CONCLUSION AND FUTURE SCOPE

This study proposed the Enhanced Accessibility Pill Recognition Network (EAPR-Net), a lightweight deep learning framework aimed at improving medication management for elderly and visually impaired users. By combining contrast-enhanced preprocessing, an efficient CNN architecture, and real-time voice-assisted feedback through an Android application, the system achieved 98% testing accuracy across 14 pill categories, demonstrating strong generalization and practical usability.

The contributions of EAPR-Net include robust contrast optimization for image preprocessing, a mobile-friendly CNN model design, and seamless integration of accessibility features such as voice output, which together enhance user independence and safety. The successful deployment of EAPR-Net through a smartphone application highlights its feasibility for real-world healthcare support.

Future work will focus on expanding the system's capabilities by adding multilingual voice support and offline functionality to serve a broader user base, especially in areas with limited connectivity. Incorporating additional healthcare features, such as dosage reminders and drug interaction alerts, and optimizing EAPR-Net further for edge AI processing will enhance its scalability and impact. With continued refinement, EAPR-Net holds significant potential to evolve into a comprehensive, accessible medication management platform for global healthcare ecosystems.

Funding: This study received no specific financial support.
Institutional Review Board Statement: Not applicable.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] R. Rádli, Z. Vörösházi, and L. Czúni, "Multi-stream pill recognition with attention," presented at the 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). https://doi.org/10.1109/IDAACS58523.2023.10348930, 2023, pp. 942-946.

[2] N. Boonthep, J. Chaiwongsai, B. Srisungsittisunti, and T. Udomsripaiboon, "Drug image classification with deep learning by using fast Region-based convolution neural network," presented at the 2024 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON). https://doi.org/10.1109/ECTIDAMTNCON60518.2024.10479992, 2024, pp. 551-555.

[3] E. Ponte, X. Amparo, K. Huang, and D. Kwak, "Automatic pill identification system based on deep learning and image preprocessing," presented at the 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE). https://doi.org/10.1109/CSCE60160.2023.00324, 2023, pp. 1-6.

[4] Y. Y. Shieh, F. Y. Tsai, A. Anavim, M. Shieh, M. D. Wang, and C.-M. C. Lin, "Mobile healthcare: The opportunities and challenges," *International Journal of Electronic Healthcare*, vol. 4, no. 2, pp. 208-219, 2008.

[5] J. Wang, S. Wang, and Y. Zhang, "Artificial intelligence for visually impaired," *Displays*, vol. 77, p. 102391, 2023. https://doi.org/10.1016/j.displa.2023.102391

[6] R. Rádli, Z. Vörösházi, and L. Czúni, "Metric-based pill recognition with the help of textual and visual cues," *IET Image Processing*, vol. 18, no. 14, pp. 4623-4638, 2024. https://doi.org/10.1049/ipr2.13273

[7] A. Roge, A. Sontakke, and N. P. Wyawahare, "Visual ease: Empowering the visually impaired with AI pill recognition," presented at the 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0. https://doi.org/10.1109/OTCON60325.2024.10687636, 2024, pp. 1-5.

[8] A. R. Ashraf, A. Somogyi-Végh, S. Merczel, N. Gyimesi, and A. Fittler, "Leveraging code-free deep learning for pill recognition in clinical settings: A multicenter, real-world study of performance across multiple platforms," *Artificial Intelligence in Medicine*, vol. 150, p. 102844, 2024. https://doi.org/10.1016/j.artmed.2024.102844

[9] A. Dhayanithi, S. Dalali, B. Madhavi, G. Surla, B. Marapelli, and A. K. Arigela, "Revolutionizing healthcare: A smart pill detection system empowered by deep learning for intelligent drug recognition," presented at the 2024 International Conference on Trends in Quantum Computing and Emerging Business Technologies. https://doi.org/10.1109/TQCEBT59414.2024.10545160, 2024, pp. 1-6.

[10] D. M. Tenneti, D. A. Kittur, H. R. T. Shankar, and R. Kumar P, "Advanced object detection for capsules and tablets identification through deep learning," in *2023 7th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*. https://doi.org/10.1109/CSITSS60515.2023.10334092, 2023: IEEE, pp. 1-6.

[11] U. Pilania, M. Kumar, and S. Singh, "Leveraging convolutional neural networks for improved drug identification," presented at the 2023 6th International Conference on Contemporary Computing and Informatics (IC3I). https://doi.org/10.1109/IC3I59117.2023.10398164, 2023, pp. 671-676.

[12] D. Zhu, Z. Huang, K. Yung, and A. W. H. Ip, "Drug recognition detection based on deep learning and improved YOLOv8," *Journal of Organizational and End User Computing*, vol. 36, no. 1, pp. 1-21, 2024. https://doi.org/10.4018/JOEUC.359770

[13] S. Kim, E.-Y. Park, J.-S. Kim, and S.-Y. Ihm, "Combination pattern method using deep learning for pill classification," *Applied Sciences*, vol. 14, no. 19, p. 9065, 2024. https://doi.org/10.3390/app14199065

[14] K. Al-Hussaeni, I. Karamitsos, E. Adewumi, and R. M. Amawi, "Cnn-based pill image recognition for retrieval systems," *Applied Sciences*, vol. 13, no. 8, p. 5050, 2023. https://doi.org/10.3390/app13085050

[15] T. Siripraiwan, N. Foythong, P. Wattanasiri, T. Nilanon, and N. Chanlen, "PharmaDetect: Drug identification and guidance through visual recognition," in *2024 IEEE International Conference on Cybernetics and Innovations (ICCI)*. https://doi.org/10.1109/ICCI60780.2024.10532475, 2024: IEEE, pp. 1-6.

[16] M. O. Pallavi, Channabasava, and G. M. Deeksha, "*Smart pill detection using machine learning models*," QTanalytics Publications. https://doi.org/10.48001/978-81-966500-0-1-7, 2024, pp. 108–117.

[17] R. Painam, T. Jaiswal, and S. Jena, "Artificial intelligence-based medicine recognition system using faster recurrent convolutional neural networks," *International Journal of Scientific Methods in Intelligent Engineering Networks*, vol. 1, pp. 46–56, 2023.

[18] S. Prabu and A. S. K. Joseph, "*Real-time pill detection and recognition framework based on a deep learning algorithm*," *in* Advancement, Opportunities, and Practices in Telehealth Technology, IGI Global. https://doi.org/10.4018/978-1-6684-5231-8.ch007, 2022.

[19] M. Sakthimohan, G. Elizabeth Rani, D. Amuthaguka, S. Akshaya, A. C. Uthaman, and S. Sridhar, "Detection and identification of pills using machine learning models," in *Proceedings of the 2nd IEEE International Conference on Vision Toward Emerging Trends in Communication and Networking Technology (ViTECoN). IEEE*. https://doi.org/10.1109/VITECON58111.2023.10157873, 2023, pp. 1-6.

[20] L. Xiang, H. Wen, and M. Zhao, "Pill box text identification using DBNet-CRNN," *International Journal of Environmental Research and Public Health*, vol. 20, no. 5, p. 3881, 2023. https://doi.org/10.3390/ijerph20053881

[21] J. Heo, Y. Kang, S. Lee, D.-H. Jeong, and K.-M. Kim, "An accurate deep learning–based system for automatic pill identification: Model development and validation," *Journal of Medical Internet Research*, vol. 25, p. e41043, 2023. https://doi.org/10.2196/41043

[22] A. Hosseini *et al.*, "A mobile application based on efficient lightweight CNN model for classification of B-ALL cancer from non-cancerous cells: A design and implementation study," *Informatics in Medicine Unlocked*, vol. 39, p. 101244, 2023. https://doi.org/10.1016/j.imu.2023.101244

[23] M. Li, Y. Jiang, Y. Zhang, and H. Zhu, "Medical image analysis using deep learning algorithms," *Frontiers in Public Health*, vol. 11, p. 1273253, 2023. https://doi.org/10.3389/fpubh.2023.1273253

[24] A. E. Chung, A. C. Griffin, D. Selezneva, and D. Gotz, "Health and fitness apps for hands-free voice-activated assistants: content analysis," *JMIR Mhealth and Uhealth*, vol. 6, no. 9, p. e9705, 2018. https://doi.org/10.2196/mhealth.9705