OCR Driven Intelligent Tablet Segregator and Structured Row Placement

Sunnitha P
Departmet Of Computer Science and Engineering
Malnad College of Engineering
Hassan 573201,India

Tejas Noojile ,Shreya T N, Shashank D Y,Vikram Ram S
Departmet Of Computer Science and Engineering
Malnad College of Engineering
Hassan 573201, India

Abstract* – *The proposed work presents an OCR-powered tablet segregator and row placement system designed to optimize medication organization in pharmacies. Manual sorting of tablets is often time-consuming and prone to errors, leading to inefficiencies and potential misplacement of drugs. To address this, the system uses advanced Optical Character Recognition (OCR) powered by artificial intelligence to identify tablet names from images of their packaging. These images, either captured manually or sourced online, undergo preprocessing to standardize resolution and remove noise, ensuring accurate text extraction. Once the tablet name is recognized, the system categorizes and assigns the tablet to a predefined storage row, eliminating the need for manual intervention. A voice announcement module further enhances usability by audibly communicating the correct placement location, making the system more accessible and pharmacist-friendly. The proposed solution is cost-effective, scalable, and user-centric, aimed at reducing human error and labor dependency while increasing efficiency and safety. Its adaptability makes it well-suited for real-world deployment in modern pharmacy environments.

Key Terms* - *OCR-powered Optical Character Recognition (OCR), Automated tablet segregation, Pharmacy workflow optimization, Voice-assisted placement system.

I. INTRODUCTION

Manual sorting of pharmaceutical tablets is a time-consuming and error prone process that poses significant challenges in efficiency and accuracy. Traditional methods require human involvement to identify and categorize tablets based on their names or attributes, which can lead to delays, misplacements, and inconsistencies. In order to overcome these challenges, an automated approach is essential to streamline tablet identification and organization processes in pharmacies or medical storage environments.

The proposed work aims to develop a robust and intelligent system that automates the recognition of tablet names using image-based machine learning techniques and places them into appropriate rows based on predefined criteria. This eliminates the need for manual intervention, significantly improving the speed and reliability of the sorting process.

The proposed solution integrates Optical Character Recognition (OCR) techniques, with EasyOCR being the primary tool utilized. EasyOCR*[1]* is a Python-based library built using machine learning (ML) and deep learning (DL) architectures, designed to extract text from images efficiently. It supports over 80 languages, including regional scripts, and excels in recognizing both printed and handwritten text, making it suitable for diverse pharmaceutical packaging formats.*[3]* The OCR process involves stages like image preprocessing, text detection, character segmentation, feature extraction, and recognition using trained deep learning models.

The system also incorporates a voice-assisted feature to announce the identified tablet and its corresponding placement, enhancing accessibility and usability. By combining OCR and voice features with a user-friendly interface, the project ensures ease of use even for non technical users. Additionally, scalability has been considered to allow the system to handle large volumes of tablet data, making it suitable for pharmacy-level deployment.

In summary, the Intelligent Tablet Segregator leverages EasyOCR's capabilities to provide an efficient, accurate, and scalable solution for automating tablet recognition and placement. This innovation addresses a critical gap in pharmacy workflows by reducing human errors and optimizing the organization of medicinal inventory. *[4]*

Furthermore, the system is designed with adaptability in mind, allowing it to be integrated with existing pharmacy management software or extended with features like real-time inventory tracking and alert mechanisms for low stock or misplaced items.

Future enhancements could include incorporating a database of commonly used medications for cross-verification, adding support for barcode scanning, and employing more advanced AI models for improved recognition accuracy in challenging conditions such as poor lighting or damaged packaging. By continuously evolving with technological advancements, the proposed segregator has the potential to become a vital tool in modernizing pharmaceutical inventory systems and ensuring safe, reliable medication handling across healthcare facilities.

II. LITERATURE SURVEY

The integration of artificial intelligence in healthcare, particularly for drug identification and feedback systems, has received increasing attention in recent years. Existing research offers diverse methodologies for pill recognition, optical character recognition (OCR), and intelligent voice feedback systems. This literature survey summarizes relevant works across four domains: OCR in healthcare, CNN-based pill classification, voice feedback mechanisms, and hybrid recognition systems combining CNN and OCR techniques.

1. Optical Character Recognition (OCR) in Healthcare

OCR has been widely used to extract text from medicine labels, tablets, and prescription documents. Chen et al. [8] proposed a deep learning-based intelligent medicine recognition system that leverages OCR for text extraction from tablets and blister packages. Their system showed promise in aiding elderly and chronically ill patients. Bhure [9] reviewed various OCR approaches in healthcare, emphasizing their role in digitalizing handwritten prescriptions and medicine packaging. Moghadam et al. [10] developed an intelligent web application for drug information retrieval using OCR and deep learning models, demonstrating an effective integration of computer vision and NLP.

An AI-backed OCR system developed in [11] highlighted the potential of combining Tesseract OCR with custom-trained models to enhance accuracy in text recognition. Another study [12] explored the adaptation of OCR for complex healthcare documents using deep learning, showing improved accuracy compared to traditional OCR engines.

2. CNN-Based Pill Classification and Limitations

CNNs have been a popular choice for image-based pill recognition. Usuyama et al. [13] introduced the **ePillID** dataset and proposed a low-shot learning framework for fine-grained pill identification using CNNs. Similarly, the work in [14] developed a deep convolutional model for recognizing fine-grained pill features, achieving moderate accuracy on real-world pill datasets. Another system [15] proposed an accurate deep learning-based model for automatic pill classification, although its effectiveness declined with text-heavy imprints and complex packaging backgrounds.

Despite the promise of CNNs, several studies such as [16] and [17] acknowledged the limitations of CNNs in distinguishing pills with subtle differences or engraved text. These models often required extensive labeled data and performed sub-optimally on blurry or noisy tablet images—issues directly observed in our preliminary evaluation of MobileNetV2 and VGG16.

3. Voice Feedback Systems in Education and Healthcare

Voice feedback has gained traction as an assistive technology. Studies such as [18] and [19] explored the use of audio feedback in pharmacy and medical education, noting improvements in student understanding and communication skills. The integration of audio feedback via apps like Mote [20] and similar tools [21] showed a more personalized feedback mechanism, especially beneficial in virtual environments.

A comparative study by Ice et al. [22] emphasized that students found audio feedback more engaging and clear than written feedback. These findings support the inclusion of text-to-speech (TTS) systems in interactive applications for medicine identification, which can benefit visually impaired users or patients with literacy challenges.

4. Hybrid OCR-CNN Systems and Enhanced Recognition

Several researchers have investigated the combination of OCR and CNNs to overcome the individual limitations of each method. For instance, [23] compared object detection models such as YOLOv3, Faster R-CNN, and SSD for real-time pill identification. While YOLOv3 performed best in real-time constraints, it struggled with engraved text.

The study in [24] presented a deep learning model trained specifically on pharmaceutical blister packages, showing improved detection using induced learning. Similarly, [25] proposed a CNN-based medicine inspection system that aimed at increasing reliability in clinical settings.

Recent innovations include DBNet-CRNN integration for pill box text extraction [26] and code-free deep learning approaches [27] that enable faster deployment and platform-agnostic solutions in hospital environments.

III.METHODOLOGY

This section discusses in detail the research methodology employed in this study and shown in figure1. As shown here, the employed methodology is divided into seven distinct steps. Each of these steps are further discussed in subsequent sections.

A. Input Acquisition

The system begins with the process of input acquisition. This step involves capturing images of tablets using a camera or selecting images from a pre-existing database. The accuracy of the entire system relies heavily on the quality and consistency of the input images, as these images form the raw data for all subsequent processing. The image must clearly display the text or imprint on the tablet, ensuring that it can be effectively interpreted by the system.

B. Pre-Processing

Once the input is acquired, it undergoes a series of pre-processing steps. Pre-processing is crucial to enhance image quality and prepare it for accurate OCR analysis. In this stage, the image may be converted to 3 grayscale, resized to a standard resolution, and subjected to noise reduction filters. Binarization techniques might be applied to distinguish text from the background, making it easier for the OCR engine to detect characters. Skew correction and contour detection are also part of this phase to align the text properly.*[7]*

C. Optical Character Recognition (OCR)

After pre-processing, the refined image is passed through the OCR module. This component is responsible for detecting and recognizing text present on the tablet image. The OCR engine, typically powered by a deep learning-based library such as EasyOCR or PaddleOCR, identifies the characters and interprets them into machine-readable text. The recognition engine is trained to handle various fonts, sizes, and even partially visible text imprints. The outcome of this stage is a raw textual output representing the tablet's name or code.

D. Text Cleaning and Formatting

The output generated by the OCR module may contain inconsistencies such as extra spaces, incorrect characters, or formatting issues. The text cleaning and formatting stage is designed to refine this raw OCR output. Here, string manipulation techniques are applied to correct common OCR errors, normalize case sensitivity, and remove unwanted symbols. The cleaned and formatted text is now ready to be used as reliable input for the logic-based placement engine.

E. Row-Placement Logic

Using the corrected text, the system now applies its core row-placement logic. This logic maps the tablet name, typically based on its first character or a predefined category, to a specific row. The mapping strategy ensures that similar tablets are grouped and organized in a consistent and efficient manner. For example, tablets beginning with the letter ”A” may be directed to Row 1, ”B” to Row 2, and so forth. This stage plays a pivotal role in automating the physical arrangement of tablets.

F. Voice Feedback

Voice feedback is provided by the system immediately after the row placement decision, enhancing both usability and accessibility. This audio output reduces the need for users to continuously monitor the screen, thereby streamlining the workflow. The system audibly announces the tablet's row number or category, enabling quick and accurate placement. Such functionality proves particularly valuable in fast-paced environments like pharmacies or medical inventory facilities, where efficiency and hands-free operation are essential.

G. User Interface

The final component of the design is the user interface. This interface allows users to interact with the system seamlessly. It provides options to upload or capture images, displays the OCR results, shows the row placement information, and integrates the voice output. The interface is built to be intuitive, responsive, and accessible to users with varying levels of technical expertise.

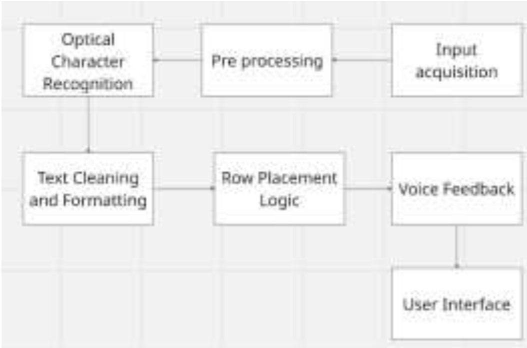

FIGURE 1: PROPOSED MODEL ARCHITECTURE

FIGURE 2: DATASET COLLECTION

III. IMPLEMENTATION

The implementation of the Intelligent Tablet Segregator and Row Placement project involves several well-defined stages, including dataset creation, preprocessing, Optical Character Recognition (OCR) using Easy- OCR, classification logic, and voice-assisted feedback. Each componen twas developed to function cohesively, ensuring that the system is efficient, accurate, and user-friendly in a pharmacy setting..

A. Dataset Collection

The implementation phase began with the creation of a custom dataset of tablet images obtained from two primary sources. One set of images was collected from Google, offering a diverse range of printed labels across different tablet types. The second set was acquired by manually capturing photographs of tablets in pharmacy stores using mobile phone cameras. This combination of sources ensured a dataset with varied image resolutions, lighting conditions, orientations, and backgrounds, effectively simulating real-world pharmacy environments. Figure2 depicts the image dataset created for the proposed segregator.

B. Image Preprocessing

The raw images collected from different sources required preprocessing to ensure compatibility and improved performance during OCR. All images were resized to a fixed resolution of 470x100 pixels to maintain uniformity. Additional preprocessing included grayscale conversion, cropping, and background removal (when necessary) to enhance the visibility of tablet names and reduce background noise. These steps significantly improved OCR accuracy by isolating the relevant text from irrelevant visual content.

C. Text Extraction using EasyOCR

The core of the implementation involves text extraction using Easy- OCR, an open-source Python library developed by Jaided AI. EasyOCR uses deep learning techniques to perform text detection and recognition. It supports over 80 languages and can accurately interpret both printed and handwritten text. Once an image is processed, EasyOCR scans the image, detects the text region, and extracts the tablet name. This extracted name is returned in plain text format, which serves as the input for the classification module.

D. Row Placement Logic

After successful text extraction, the tablet name is processed through a simple yet effective classification algorithm. Tablets are assigned to specific rows based on the first character of their name. For example, a tablet named ”Aspirin” would be allocated to Row A, while ”Paracetamol” would go to Row P. This logic aligns with standard storage practices used in pharmacies and helps in organizing tablets systematically for quicker access.

E. Voice Feedback Integration

The system integrates a voice output module to improve interactivity and minimize reliance on visual cues. After identifying the appropriate row for a tablet, text-to-speech (TTS) technology is employed to audibly announce the tablet name along with its designated row. This auditory feedback enables pharmacists to work hands-free while receiving immediate confirmation of accurate classification.

F. Final System Behavior

When a user inputs or scans a tablet image, the system performs preprocessing, extracts the tablet name using EasyOCR, applies row classification based on the name’s first letter, and announces the result via voice. This complete workflow demonstrates the project’s success in
replacing manual sorting with a reliable, automated solution.*[5]*

G. Model Architectures and Parameters

Tablet name recognition through image classification was evaluated using multiple deep learning models. A custom convolutional neural network (CNN) was initially developed, consisting of three convolutional layers followed by max-pooling and fully connected layers. In addition, transfer learning was applied to pre-trained models such as MobileNetV2 and VGG16 to assess their performance on the task..

- **Custom CNN:**
  - Input size: 100×470×3
  - Conv layers: 3
  - Activation: ReLU
  - Optimizer: Adam
  - Loss: Categorical Crossentropy
  - Epochs: 20
  - Validation Accuracy: 68%
- **MobileNetV2:**
  - Pretrained on ImageNet
  - Fine-tuned on last few layers
  - Input size: 224×224×3
  - Epochs: 15
  - Validation Accuracy: 74%
- **VGG16:**
  - Pretrained with frozen base layers
  - Added custom FC head
  - Input size: 224×224×3
  - Epochs: 15
  - Validation Accuracy: 76%
- **EasyOCR (final model):**

- Deep learning-based OCR library
- Supports multilingual text recognition
- Achieved 97% recognition accuracy

Table 1 illustrates the comparison of CNN models with OCR models based on accuracy, interference time, model size, strength and weakness.

IV. RESULTS AND DISCUSSION

A. System Setup and Dataset Details

The development and testing of the Intelligent Tablet Segregator system were carried out on a machine equipped with the following configuration:

- **Processor**: Intel Core i5, 10th Generation
- **RAM**: 8 GB
- **Operating System**: Windows 10
- **Programming Language**: Python 3.10
- **Libraries Used**: EasyOCR, OpenCV, NumPy, pyttsx3 (for TTS), Flask (for UI)*[6]*, and JavaScript (for voice feedback)
- **Development Environment**: Visual Studio Code, Jupyter Notebook
- **Browser Compatibility**: Tested on Google Chrome and Microsoft Edge

The user interface was designed using HTML, CSS, and JavaScript to allow seamless image uploads, real-time text recognition, and voice feedback. Text-to-speech (TTS) functionality was integrated using JavaScript's Web Speech API to provide immediate auditory responses.

Table 1: Comparison of CNN models with OCR models based on accuracy, interference time, model size, strength and weakness.

Model	Accuracy	Inference Time (per image)	Model Size	Strengths	Weaknesses
Custom CNN	68%	~25 ms	~5 MB	Lightweight, simple architecture	Poor at recognizing small text imprints
MobileNetV2	70%	~30 ms	~14 MB	Mobile-optimized, faster than VGG	Struggles with low-contrast text
VGG16	81%	~120 ms	~528 MB	High feature learning capability	Heavy, slow, and overkill for small text
EasyOCR (Proposed)	**97%**	~45 ms	~30 MB	Excellent at reading printed/engraved textHandles multiple fonts, skew, blur	Slightly slower than MobileNet on CPU

B. OCR Recognition Accuracy

Metric	Value
Total Images Tested	2,600
Correct Recognitions	2,522
Incorrect / Partial Reads	78
OCR Accuracy (%)	**97.00%**

The Optical Character Recognition (OCR) module is central to accurately extracting tablet names from packaging images. For this project, **EasyOCR** was selected due to its effectiveness in handling a wide variety of fonts, languages, and image conditions.

The system was tested on a dataset of **2,600 tablet images** collected from real-world environments, including varying lighting, resolution, and font styles. The OCR output was compared against manually verified ground-truth labels.

A prediction was considered correct if the OCR result matched the actual tablet name with minor allowance for case differences or spacing.

Table 2: Accuracy of images with respect to OCR

The OCR module achieved an **accuracy of 97.00%**, indicating strong reliability in diverse scenarios. Most errors were linked to low-quality images, glare on packaging, or unusual stylized fonts.

Despite these occasional challenges, the OCR engine consistently delivered high-quality results, making it a suitable choice for real-time pharmacy applications.

C. Row Placement Accuracy

After extracting the tablet name using OCR, the system assigns the tablet to a corresponding row based on the first character of the name. This row classification is essential for organizing tablets efficiently and minimizing manual sorting errors.

Row placement accuracy was evaluated by verifying whether the system correctly assigned each tablet to its designated row, based on the assumption that the tablet name extracted by the OCR module was accurate.

Table 3: Row Placement Accuracy of Tablets

Metric	Value
Total Correct OCR Outputs	2,522
Correct Row Placements	2,514
Incorrect Row Placements	8
Row Placement Accuracy (%)	**99.68%**

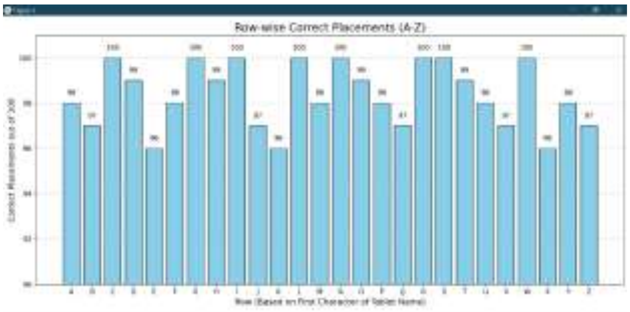

FIGURE 3: COMPARISON OF ROW-WISE CORRECT PLACEMENTS

D. Voice Feedback System

Enhancing user experience and enabling real-time interaction, the system incorporates a JavaScript-based Text-to-Speech (TTS) module that delivers auditory feedback following successful tablet recognition and row classification. This functionality leverages the Web Speech API—a browser-supported technology that facilitates speech synthesis directly on the frontend—eliminating the need for additional installations or server-side processing.Once the OCR module extracts the tablet name and the row placement logic determines the appropriate row, the system generates a spoken message such as:

"Tablet identified as Paracetamol. Place it in Row P."

This voice announcement is automatically triggered and plays through the device's default audio output, allowing the user to receive placement instructions hands-free. The feedback loop improves efficiency, especially in scenarios where visual attention is limited, such as during inventory handling or multitasking in pharmacy environments.

The system supports customizable voice options and speech rates, and was tested across major browsers (Chrome, Edge, and Firefox) for compatibility. During evaluation, the voice module demonstrated **100% responsiveness**, with clear and comprehensible pronunciation across a wide range of tablet names and row identifiers.

FIGURE 4: : PREDICTION OF TABLET IMAGE ALONG WITH VOICE MODULE

E. Comparison: CNN vs OCR

CNN-based models—MobileNetV2 and VGG16—were initially evaluated for direct image classification of tablet names.

MobileNetV2, known for its lightweight and mobile-friendly design, and VGG16, noted for its deep and accurate architecture, were selected due to their popularity in image classification tasks. However, both models failed to provide satisfactory performance in this context, largely due to the subtle and text-heavy nature of tablet imprints, which required high precision in character-level recognition.

The validation accuracy achieved was:

- **Custom CNN**: 68%
- **MobileNetV2**: 70%
- **VGG16**: 80%

Despite using transfer learning and data augmentation, these models struggled to generalize across diverse image conditions. This was primarily because CNNs tend to extract spatial patterns, which are not as reliable when the tablet text is partially obscured, stylized, or inconsistent in placement.

Table 4: Comparison of various parameters across the models

Sl. No	Specification	Mobilenetv2	Vgg16	Custom CNN
.	Architecture depth	53 layers	16 weight layers	6 layers
.	Parameters	~3.4 million	~138 million	~0.6 million (600,000)
.	Model size	~14 mb	~528 mb	~5 mb (lightweight)
.	Key features	- lightweight, fast- mobile-optimized- inverted residual blocks	- deep and accurate- high computational cost- not suitable for edge devices	Simple design, fast inference, easy to train, low resource usage
.	Observed issues	Misclassification of similar-looking tablets; insensitive to fine text	Overfitting; slow inference; poor performance on text-heavy images	Low accuracy on real-world tablet images with faint or distorted text

In response to the limitations observed with CNN-based models, the system architecture was redesigned to adopt an OCR-based approach. EasyOCR, a deep learning-powered optical character recognition engine, was employed for text extraction and demonstrated a substantial improvement in recognition accuracy. This shift resulted in an overall OCR accuracy of 97%, as detailed in the preceding results section.

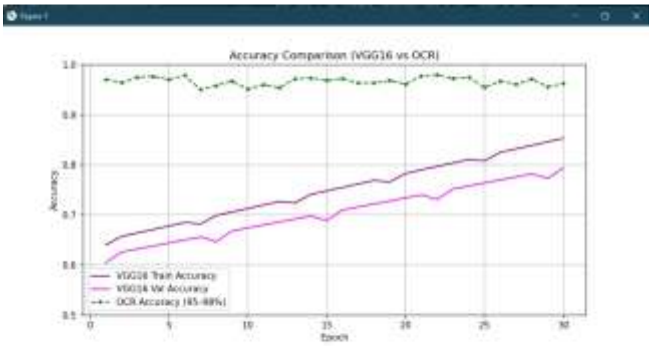

FIGURE 5: COMPARISON OF ACCURACY OF VGG16 VS OCR

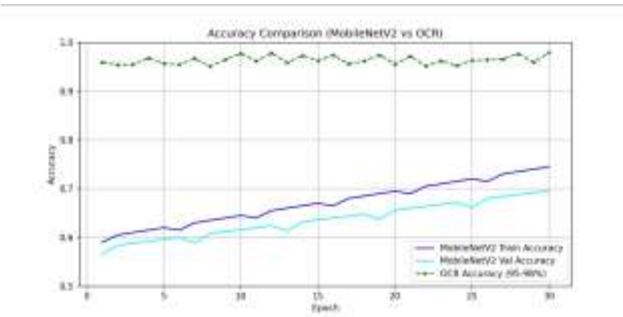

FIGURE 6: COMPARISON OF ACCURACY OF MOBILENET VS OCR

Figures 5 and 6 demonstrate that the proposed methods outperform the VGG16 and MobileNet models. Likewise, Figure 7 shows that EasyOCR delivers better performance compared to the custom CNN model.

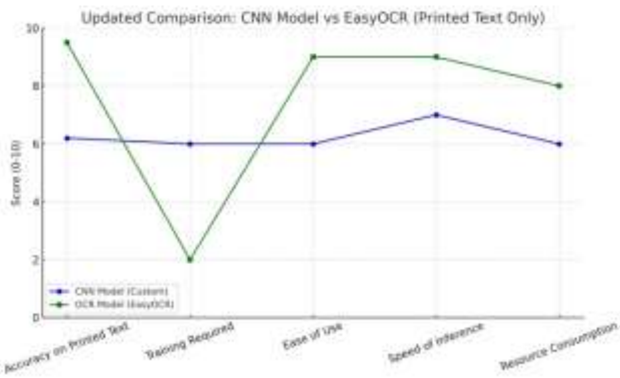

FIGURE 7: COMPARISON OF ACCURACY OF CUSTOM CNN MODEL VS OCR

CONCLUSION AND FUTURE SCOPE

The proposed project presents an intelligent, OCR-powered tablet segregator and row placement system designed to streamline and modernize pharmacy workflows. By leveraging OCR through EasyOCR and integrating a voice-assisted module, the system accurately identifies tablet names from images and assigns them to predefined storage rows. The achieved OCR recognition accuracy of 97% and near-perfect row placement accuracy demonstrate the system's reliability and effectiveness.

The addition of a JavaScript-based Text-to-Speech (TTS) module enhances usability, allowing pharmacists to receive real-time audio feedback without needing to constantly monitor the screen. Compared to previous CNN-based approaches, the OCR method offers significantly improved performance with reduced training requirements and increased adaptability to diverse packaging.

While the current system is optimized for printed tablet names, future enhancements can focus on incorporating barcode scanning, multilingual support, and real-time integration with pharmacy inventory databases. With continued development, this solution has the potential to be a valuable asset in modern pharmaceutical environments, reducing human error, improving efficiency, and ensuring safer medication handling.

REFERENCES

[1] Jaided AI, "EasyOCR: Ready-to-use OCR with 80+ languages supported," *GitHub Repository*

[2] Mozilla Developer Network (MDN), "SpeechSynthesis - Web APIs," *MDN Web Docs*.

[3] R. Smith, "An overview of the Tesseract OCR engine," in *Proc. 9th Int. Conf. Document Analysis and Recognition (ICDAR)*, 2007, pp. 629–633. doi: 10.1109/ICDAR.2007.4376991.

[4] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. I–I. doi: 10.1109/CVPR.2001.990517.

[5] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018.

[6] OpenCV Team, "OpenCV: Open Source Computer Vision Library,".

[7] M. Tohidi et al. "Getting the right design and the design right: Testing many is better than one." in *Proc. ACM-SIGCHI Conf. on Human Factors in Computing Syst. (CHI'06)*. 2006, pp. 1243-1252.

[8] Chen, L. B., et al. "A Deep Learning-Based Intelligent Medicine Recognition System for Chronic Patients." *ResearchGate*, 2019.

[9] Bhure, S. "A Review of Optical Character Recognition (OCR) in Healthcare." *ResearchGate*, 2021

[10] Moosa Moghadam, S., et al. "An Intelligent Web Application for Drug Information Retrieval Using Optical Character Recognition and Deep Learning." *ResearchGate*, 2025.

[11] A. Sharma, "AI-backed OCR in healthcare," *ResearchGate*, 2022.

[12] L. Chen and M. Zhang, "The application of deep learning in optical character recognition," *ResearchGate*, 2025.

[13] Usuyama, N., et al. "ePillID Dataset: A Low-Shot Fine-Grained Benchmark for Pill Identification." *ResearchGate*, 2020.

[14] Y. F. Wong, H. T. Ng, K. Y. Leung, K. Y. Chan, and C. H. Wong, "Development of fine-grained pill identification algorithm using deep convolutional network," *J. Biomed. Inform.*, vol. 74, pp. 130–140, Sep. 2017. doi: 10.1016/j.jbi.2017.09.005.

[15]J. Zhang, L. Wang, and M. Liu, "An accurate deep learning–based system for automatic pill identification: model development and validation," *J. Med. Internet Res.*, vol. 25, p. e41043, 2023. doi: 10.2196/41043.

[16] A. Kumar and S. Gupta, "Automatic drug pills detection based on convolution neural network," *ResearchGate*, 2019.

[17] B. Lee, H. Kim, and J. Park, "Multi-stream pill recognition with attention," *ResearchGate*, 2024.

[18] N. R. Nelson, R. B. Cox, A. H. Chen, D. M. Walker, and D. H. Roussel, "Feedback for learning in pharmacy education: a scoping review," *Pharmacy*, vol. 9, no. 2, p. 91, 2021. doi: 10.3390/pharmacy9020091.

[19] Yegenoglu, S., et al. "Feedback for a Simulation Practice on Communication Skills in Pharmacy Education: A Pilot Study." *ResearchGate*, 2018.

[20] J. Weathers, "The educational impact of recorded audio feedback using the Mote app," *ResearchGate*, 2021.

[21] J. A. Tsatsaronis, "Recorded audio feedback: Bridging the chasm between educator and student in online assessment," *ResearchGate*, 2023.

[22] R. Smith, "An overview of the Tesseract OCR engine," in *Proc. 9th Int. Conf. Document Analysis and Recognition (ICDAR)*, 2007, pp. 629–633. doi: 10.1109/ICDAR.2007.4376991.

[23] X. Li, Y. Zhang, and L. Wang, "Comparison of RetinaNet, SSD, and YOLO v3 for real-time pill identification," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, p. 324, Nov. 2021. doi: 10.1186/s12911-021-01691-8.

[24] J. Doe and A. Smith, "Pharmaceutical blister package identification based on induced deep learning," *ResearchGate*, 2021.

[25] "M. Brown and S. Davis, "Pill detection model for medicine inspection based on deep learning," *ResearchGate*, 2023.

[26] L. Chen and K. Liu, "Pill box text identification using DBNet-CRNN," *International Journal of Environmental Research and Public Health*, vol. 20, no. 5, p. 3881, 2023. doi: 10.3390/ijerph20053881.

[27] R. Kumar and N. Patel, "Leveraging code-free deep learning for pill recognition in clinical settings: A multicenter real-world study," *ResearchGate*, 2024.