A Comprehensive Literature Survey on Automated Prescription Digitization: OCR Techniques, Deep Learning Approaches, and Medical Information Retrieval Systems

Swapnil Ananda Badave[1], Sarvesh Sanjay Dhule[2,] Shreyash Purushottam Ghanekar[3]
Kanad Prasad Kulkarni[4], Mrs. Sumitra Jakhete[5]
[1,2,3,4]*Department of Information Technology Pune Institute of Computer Technology Pune, Maharashtra, India*
[5]*Assistant Professor Department of Information Technology Pune Institute of Computer Technology Pune, Maharashtra*

***Abstract*—Handwritten medical prescriptions remain a persistent challenge in healthcare systems due to illegible handwriting, inconsistent prescription formats, and frequent use of non- standard abbreviations. These issues often lead to medication errors, delays in dispensing, and increased workload for pharma- cists and healthcare professionals. Recent advances in computer vision, optical character recognition (OCR), deep learning, and biomedical natural language processing have enabled the development of automated systems capable of converting handwritten prescriptions into structured digital records.**

This paper presents a comprehensive literature survey of automated prescription digitization systems, focusing on OCR techniques, deep learning–based handwriting recognition models, and medical information extraction frameworks. The study analyzes traditional OCR engines such as Tesseract, modern deep learning models including CNN-RNN architectures and transformers, object detection models such as YOLOv5 for region localization, and biomedical language models such as BioBERT for entity recognition.

Furthermore, this review discusses common architectural approaches found in recent studies for integrating image pre- processing, text detection, OCR, entity extraction, and drug knowledge base validation into unified healthcare automation pipelines. The survey highlights current challenges including handwriting variability, dataset limitations, OCR errors, and the lack of standardized evaluation benchmarks. The insights provided in this paper aim to guide researchers and developers in understanding the current state-of-the-art and identifying open gaps in prescription digitization technologies.

***Index Terms*—Prescription digitization, Literature Survey, OCR, Deep Learning, YOLOv5, BioBERT, Natural Language Processing, Healthcare automation**

I. INTRODUCTION

Handwritten prescriptions remain prevalent across health- care systems worldwide, particularly in outpatient clinics and smaller hospitals in developing regions [1]. While this traditional approach offers convenience for clinicians during rapid consultations, it presents numerous critical challenges that compromise patient safety and healthcare efficiency [2]. The fundamental problems include inconsistent medical abbreviations (e.g., “bd” versus “bid” for twice daily), ambiguous drug name representations mixing brand and generic nomenclature, and highly variable handwriting styles that differ significantly across medical practitioners [3].

Research has demonstrated that medication errors attributed to illegible handwriting affect approximately 7,000 to 9,000 patients annually, with costs reaching billions of dollars in additional healthcare expenditures [1]. Pharmacists frequently expend substantial time and cognitive resources attempting to decipher ambiguous prescriptions, which introduces delays in treatment initiation and

elevates the risk of dispensing incorrect medications or dosages [2]. A survey conducted in Egypt revealed that 96% of respondents supported the development of software applications capable of translating handwritten medical prescriptions into structured digital text [1].

Automating the conversion of prescription images into structured, validated, and machine-readable data promises transformative benefits for modern healthcare ecosystems [3]. These benefits include: (1) substantial reduction in manual transcription effort by pharmacy personnel, (2) accelerated

medication dispensing workflows enabling faster patient ser- vice, (3) integration of clinical decision support systems for real-time contraindication and drug interaction checking, and (4) enhanced patient safety through elimination of interpretation errors [4], [5].

However, constructing such systems presents significant technical challenges. Solutions must demonstrate robustness against extreme handwriting variability, accommodate diverse image quality conditions including blur and non-uniform illumination, handle domain-specific medical notations and abbreviations, and integrate with existing healthcare information systems [7], [8].

The objective of this paper is to synthesize best practices identified in existing literature to provide a comprehensive overview of the current state-of-the-art. By examining how recent studies tackle preprocessing, detection, OCR, biomed- ical entity recognition, and knowledge-based validation, this survey aims to serve as a foundational reference for future research in healthcare automation.

II. RELATED WORK AND COMPARATIVE ANALYSIS

Automated handwriting recognition and medical entity ex- traction represent active research domains with substantial academic and industrial interest. Historical approaches combined classical OCR engines such as Tesseract with heuristic preprocessing techniques [4], while contemporary research emphasizes deep learning methodologies including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and transformer-based architectures like TrOCR [7]. Region detection models including Mask R-CNN and YOLO variants have demonstrated effectiveness for locating text regions in complex document images [3]. Biomedical language models such as BioBERT and ClinicalBERT have substantially advanced named entity recognition capabilities in clinical text domains [8].

Despite these advances, critical gaps persist in the literature. Many published systems report performance metrics on small- scale or proprietary datasets without releasing reproducible implementations [1], [2]. Few comprehensive frameworks integrate robust drug knowledge verification mechanisms or provide clear pathways for deployment and clinical validation [3], [4]. Table I summarizes representative recent research, highlighting methodologies, limitations, and applicability do- mains.

III. COMMON ARCHITECTURAL PARADIGMS IN PRESCRIPTION DIGITIZATION

A review of the literature reveals that most successful prescription digitization systems employ modular architectures to handle the high complexity of the task. These architectures typically divide the workload across distinct layers:

Presentation Layer: Client-facing interfaces are commonly developed to handle prescription image uploads and display structured results. Recent studies often implement human-in- the-loop mechanisms at this layer, allowing pharmacists to review low-confidence OCR outputs and verify medication alternatives before dispensing.

Application Layer: This layer orchestrates the core processing pipelines. Research indicates a shift toward microservice architectures, where image preprocessing, text region detection, OCR execution, and named entity recognition are separated into distinct modules. This allows for independent scaling and easier integration of newer, more efficient deep learning models.

Data Layer: Comprehensive systems heavily rely on robust data layers. Beyond storing the digitized records, state-of- the-art frameworks integrate curated drug knowledge bases (containing active ingredients, dosages, and contraindications) to cross-validate the extracted entities.

The standard processing flow observed in contemporary research typically proceeds from image

acquisition and pre- processing to region detection, transcription (OCR), entity extraction (NER), and finally, semantic validation against a medical database.

IV. CORE COMPONENTS OF DIGITIZATION PIPELINES

This section describes the individual modules and technique most frequently utilized in recent literature to build prescription digitization pipelines.

A. Image Preprocessing Techniques

A robust preprocessing pipeline is universally cited as essential for improving OCR accuracy. The literature frequently highlights the following sequential operations:

- Color Normalization and Resizing: Converting images to grayscale using luminosity-preserving transformations

[1] and normalizing resolution to reduce computational load.

- Noise Reduction: Applying Gaussian blur kernels and Non-Local Means (NLM) denoising algorithms [3] to remove scanner artifacts.
- Skew Correction: Utilizing Hough transform-based line detection or projection profile analysis [4] to estimate document skew and correct orientation.
- Adaptive Thresholding: Employing local thresholding techniques like Sauvola or adaptive Otsu methods [1] to handle non-uniform illumination.
- Morphological Operations: Using opening and closing operations [3] to connect broken strokes and remove speckles.

B. Text Region Detection

Before transcribing text, locating the areas of interest is cru- cial. Object detection frameworks, particularly YOLO variants, are prevalent in recent studies due to their single-pass inference architecture and real-time processing capabilities [3], [9]. Researchers frequently fine-tune these models on annotated prescriptions to localize drug names, dosages, frequencies, and patient demographics. Non-maximum suppression (NMS) is standardly implemented to filter overlapping bounding boxes and minimize false positives.

TABLE I Comparative Analysis of Related Research on Handwritten Prescription Recognition

Paper (Year)	Methodology	Dataset/Evaluation	Accuracy	Key Limitations
Pavithiran et al. [1]	CNN+RNN+LSTM with fuzzy search	Custom multi-language dataset	89.5% with augmentation	Sensitive to handwriting variations; limited dataset size; regional language complexity
Kumar et al. [3]	YOLOv5 detection + DL recognition + Flask web app	Real-world prescriptions	Not reported	Dependent on image quality; limited scalability evaluation; lacks validation mechanism
Sanjeevaiah et al. [4]	CNN preprocessing + Tesser- act OCR	Small proprietary dataset	Moderate	Low performance for diverse handwriting; dataset-limited; no entity validation
Amritha et al. [5]	CNN + Tesseract + NER for mobile app	Mobile-captured images	Not reported	Focus limited to drug extraction; mobile scope constraints; no comprehensive interpretation
Lopez et al. [6]	CLEAR: Retrieval augmented pipelines with LLMs	Clinical text corpus	High	Requires high-quality NER; scalability concerns;

				computationally expensive
Usman et al. [7]	Mask R-CNN + TrOCR trans- former	Domain-specific handwriting dataset	Low CER	Requires substantial domain data; compute-intensive; limited real-world testing
Kanisshka et al. [8]	EasyOCR + BioBERT token classification	Annotated prescription images	Good entity F1	Sensitive to lighting/preprocessing; limited entity type coverage

C. OCR Transcription Strategies

The literature shows a transition from single-engine OCR to multi-engine and deep learning approaches. Many researchers leverage Tesseract for printed text [4] and EasyOCR for its strong handwritten and multilingual capabilities [8]. Some state-of-the-art proposals utilize a dual-engine or ensemble approach, scoring outputs based on confidence metrics and dictionary matching, to mitigate individual engine weaknesses and reduce error propagation [8].

D. Text Postprocessing and Named Entity Recognition (NER)

Following transcription, raw text requires normalization and semantic extraction. Standard postprocessing includes expanding common medical abbreviations using domain-specific lexicons [1] and employing fuzzy string matching to correct typical OCR errors [2].

For extraction, biomedical language models like BioBERT have become the standard [6], [8]. Pretrained on biomedical literature, these models are fine-tuned for token-level sequence labeling (e.g., IOB tagging) to categorize text into entities such as DRUG, DOSAGE, FREQUENCY, and ROUTE. The literature demonstrates that domain-adapted transformers significantly outperform generic models on these specific tasks.

E. Drug Knowledge Base Integration

Extracting text is often insufficient for clinical safety; the extracted data must be validated. Recent frameworks increasingly incorporate structured drug knowledge bases [11]. These databases allow systems to verify recognized drug names, check for active pharmaceutical ingredients, and even power recommendation engines that suggest composition-equivalent alternative medications based on cosine similarity of active ingredients.

V. DATASETS AND ANNOTATION STRATEGIES IN EXISTING

A persistent challenge noted across the literature is the lack of large-scale, public, annotated datasets for handwritten prescriptions. High-quality data is crucial for training robust models, leading researchers to adopt specific strategies:

- Proprietary Collection: Most studies rely on diverse prescription images collected from local clinics, strictly following de-identification and privacy protocols.
- Standardized Annotation: Researchers utilize tools like Labeling for object detection bounding boxes [3] and BRAT for token-level NER annotations, often employing multiple human annotators to ensure high inter-annotator agreement.
- Data Augmentation: To compensate for limited dataset sizes, synthetic augmentation techniques—such as affine transformations, blur injection, and lighting variation—are heavily utilized to simulate diverse handwriting styles and capture conditions [1], [10].

VI. STANDARD EVALUATION METRICS

The literature employs a multi-stage evaluation framework to assess both individual modules and end-to-end performance.

Detection Performance: Mean Average Precision at loU threshold 0.5 (mAP@0.5) is widely used to evaluate region
detection:

$$mAP@0.5 = \frac{1}{N}\sum_{i=1}^{N} AP_i \quad (1)$$

where N represents the number of classes and AP; denotes
Average Precision for class i [9].
OCR Accuracy: Character Error Rate (CER) is the standard
for measuring transcription quality:

$$CER = \frac{S + D + I}{N} \quad (2)$$

where S, D, and / represent substitutions,
deletions and insertions respectively, and N is the total number of characters in the ground truth [7].
NER Performance: Entity-level F1-score evaluates extraction capabilities:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

These metric balances precision and recall, which is crucial where false positives and false negatives both carry clinical risks [8].

B. End-to-End System Metrics
At the system level, researchers evaluate Extraction Accuracy (the percentage of prescriptions where all critical triples-drug, dosage, frequency-are correctly extracted) and Processing Latency to ensure the solution is viable for real time clinical workflows.

VII. OPEN CHALLENGES AND FUTURE DIRECTIONS

Despite significant advancements, several challenges remain open in the field of prescription digitization:

- Extreme Handwriting Variability: The diversity in medical cursive continues to challenge even state-of-the- art deep learning models. Future research points toward utilizing active learning to continuously improve models based on human-in-the-loop corrections.
- Lack of Standardized Benchmarks: The reliance on small, proprietary datasets makes it difficult to compare different frameworks objectively. There is a pressing need for a large-scale, open-source benchmark dataset for medical prescriptions.
- Regional Abbreviations: Practitioner-specific and regional abbreviation usage complicates interpretation, re- quiring dynamic, easily updatable lexicons.
- Regulatory Compliance: Transitioning these research prototypes into clinically deployed medical devices re- quires rigorous validation studies to satisfy stringent healthcare regulations.

VIII. CONCLUSION

This manuscript has presented a comprehensive literature survey on automated hand Watten perception digitization, exploring the intersection of computer vision, OR, and natural language processing by synthesizing best practices and analyzing current methodologies-from YOLO-based region detection and ensemble OCR transcription to BioBERT- driven entity extraction-this survey highlights the significant progress made toward reducing medication errors and streaming pharmacy workflows.
However, the review also underscores notable gaps in current research, primarily concerning data set limitations, extreme handwriting variability, and the need for robust, knowledge-based validation mechanisms. Addressing these challenges through standardized benchmarks and integrated, modular pipelines represents the critical next step for researchers.
Ultimately, advancing these digitization technologies holds immense potential for enhancing patient safety and dem concretizing access to structured medical information in modem
healthcare ecosystem s.

REFERENCES

[1] G. Pavithiran, S. Padmanabhan, N. Divya, V. Aswathy, P. I. Jerusha, these metric balances precision and recall, which is crucial where false positives and false negatives both carry clinical risks [8].and B. Chandar, “Doctor’s Handwritten Prescription Recognition System in Multi-Language Using Deep Learning,” arXiv preprint arXiv:2210.11666, 2022.
[2] E. Hassan, H. Tarek, M. Hazem, S. Bahnacy, L. Shaheen, and W. H. Elashmwai, “Medical

Prescription Recognition using Machine Learning," in 2021 IEEE 11th Annual Computing and Communication Work- shop and Conference (CCWC), pp. 0973-0979, 2021.

[3] K. N. Kumar, G. S. Saatvik, S. L. Kumar, and A. Rajiv, "Automated Medical Prescription Recognition and Analysis System Using OCR and Deep Learning," International Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 5, May 2025.

[4] K. Sanjeevaiah, M. A. Lateef, M. A. Mufasil, M. Ganesh, and M. P. Kumar, "Medical Prescription Optical Character Recognition," Int. J. Eng. Technol. Manag. Sci., vol. 8, no. 3, pp. 246-253, 2024.

[5] A. Babu et al., "Handwritten Prescription Recognition using CNN and Tesseract with NER," International Journal of Engineering Research, 2025.

[6] I. Lopez et al., "Clinical Entity Augmented Retrieval for Clinical Information Extraction," npj Digital Medicine, 2024.

[7] U. Ali, S. Ranmbail, M. Nadeem, H. Ishfaq, M. U. Ramzan, and W. Ali, "Leveraging Deep Learning with Multi-Head Attention for Accurate Extraction of Medicine from Handwritten Prescriptions," arXiv preprint arXiv:2412.18199, 2024.

[8] K. U. P., U. Jayachandiran, and G. S. Ananya, "Intelligent Medical Prescription Analysis Using EasyOCR and NER," Int. J. Novel Res. Dev., vol. 9, no. 12, pp. 786-793, 2024.

[9] G. Jocher et al., "YOLOv5," GitHub repository, 2020. [Online]. Avail- able: https://github.com/ultralytics/yolov5

[10] X. Liu, G. Meng, S. Xiang, and C. Pan, "Handwritten Text Generation via Disentangled Representations," IEEE Signal Process. Lett., vol. 28, pp. 1838-1842, 2021.

[11] World Health Organization, "Medication Errors," Technical Report Se- ries, WHO, Geneva, 2023.

[12] S. Tabassum et al., "Recognition of Doctors Cursive Handwritten Medical Words by using Bidirectional LSTM and SRP Data Augmentation," in 2021 IEEE Technology Engineering Management Conference Europe (TEMSCON-EUR), pp. 1-6, 2021.

[13] Q. Chen, T. Gong, L. Li, C. L. Tan, and B. C. Pang, "Handwriting Recognition for Medical Prescriptions using a CNN-Bi-LSTM Model," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.