Indian Journal of Modern Research and Reviews

This Journal is a member of the '*Committee on Publication Ethics*'

Online ISSN:2584-184X

Research Article

AI-Based Clinical Prescription Recognition and Medication Management Using Multimodal OCR and Transformer Models: A Survey

Prince Godwin D [1], Vishanth V [2], Dinesh P [3], Sundari V[4*]

[1-4] Department of CSE, Meenakshi Sundararajan Engineering College, Chennai, Tamil Nadu, India

Corresponding Author: *Sundari V

DOI: https://doi.org/10.5281/zenodo.19435929

Abstract

Medical prescriptions often contain handwritten and unstructured information, making accurate interpretation difficult and leading to potential medication errors. Existing healthcare digitisation systems primarily focus on basic text extraction and fail to incorporate intelligent clinical analysis, such as drug interaction detection and medication scheduling. This paper proposes an intelligent AI-powered Clinical Prescription Recognition and Medication Management System using multimodal Optical Character Recognition and transformer-based Natural Language Processing with Role-Based Access Control. The system preprocesses prescription images through enhancement and normalisation techniques, converting them into structured medical representations, including drug names, dosage, frequency, and duration, using multi-entity extraction. Transformer-based models accurately interpret prescription text, while an integrated interaction analysis module identifies potential drug conflicts using knowledge-driven rules. A scheduling engine generates personalised medication plans based on extracted instructions and clinical constraints. The system further provides secure role-based access, enabling patients to receive medication reminders and administrators to monitor prescription history, interaction alerts, and adherence patterns. The proposed framework improves accuracy, safety, and usability compared to traditional OCR-based approaches, enabling intelligent and automated healthcare workflow management.

Manuscript Information

- **ISSN No:** 2584-184X
- **Received:** 15-02-2026
- **Accepted:** 27-03-2026
- **Published:** 06-04-2026
- **MRR:**4(4); 2026: 40-45
- **©2026, All Rights Reserved**
- **Plagiarism Checked:** Yes
- **Peer Review Process:** Yes

How to Cite this Article

D P G, V V, P D, V S. AI-Based Clinical Prescription Recognition and Medication Management Using Multimodal OCR and Transformer Models: A Survey. Indian J Mod Res Rev. 2026;4(4):40-45.

Access this Article Online

www.multiarticlesjournal.com

KEYWORDS: Prescription Recognition, OCR, Healthcare AI, Multi-Entity Extraction, Drug Interaction Detection, Transformer Models, Medication Scheduling, Clinical NLP.

INTRODUCTION

Medical prescription processing has become a critical challenge in modern healthcare systems due to the widespread use of handwritten and unstructured prescription formats. Illegible handwriting, ambiguous abbreviations, and incomplete information significantly contribute to medication errors, posing serious risks to patient safety and increasing healthcare costs [13, 15]. Prescriptions contain valuable clinical information such as drug names, dosage, frequency, and duration; however, their unstructured nature makes large-scale manual interpretation inefficient, time-consuming, and prone to human error [2, 9].

To address these challenges, recent research emphasises the use of Artificial Intelligence (AI) and Machine Learning (ML) techniques for automated prescription recognition and clinical data extraction [1, 4]. Deep learning models, particularly transformer-based architectures and multimodal vision-language models, have demonstrated strong performance in extracting structured information from handwritten and printed medical documents [3].

[6, 12]. These approaches enable accurate recognition of complex prescription patterns and contextual understanding of medical terminology.

In addition to data extraction, advanced systems incorporate multi-entity recognition and intelligent healthcare analytics to enhance functionality. Multi-task learning frameworks are increasingly used to simultaneously extract prescription details, detect drug interactions, and generate medication schedules, improving overall system efficiency and clinical relevance [4, 7]. Knowledge-based models and graph-driven approaches further support safe medication practices by identifying potential adverse drug combinations [5, 15].

Recent advancements also highlight the importance of explainable AI techniques and secure data management in healthcare applications. Ensuring transparency, interpretability, and data privacy is essential for building trust and enabling real-world deployment in clinical environments [8, 14]. Role-based access control and secure storage mechanisms play a crucial role in protecting sensitive patient information while enabling efficient system usage.

The rest of the paper is organised as follows: Section II revisits the fundamental concepts of AI-based prescription analysis. Section III discusses deep learning techniques for prescription recognition and medication management. Section IV presents a comparative analysis of existing approaches. Section V outlines key challenges and research gaps, and Section VI concludes the paper with future research directions.

Revisiting Basics

AI-based prescription recognition systems primarily rely on machine learning and deep learning models for extracting meaningful information from structured and unstructured medical data. Early approaches utilised traditional Optical Character Recognition (OCR) techniques and classical machine learning algorithms for text extraction; although effective for printed text, these methods often struggle with handwritten prescriptions, complex layouts, and ambiguous medical abbreviations [2, 9]. In contrast, deep learning architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformer-based models enable automatic feature extraction and contextual understanding, making them suitable for complex clinical text interpretation tasks [1, 12].

Recent advancements incorporate multimodal learning and vision-language models to process both textual and visual features from prescription images. These approaches improve recognition accuracy by capturing spatial layout, handwriting patterns, and contextual dependencies within medical documents [3, 6]. Additionally, knowledge-driven systems and graph-based models enhance clinical intelligence by enabling drug interaction detection and semantic understanding of prescription components [5, 15]. Multi-task and multi-entity learning strategies further improve system capability by simultaneously extracting medication details, identifying drug interactions, and generating dosage schedules within a unified framework, increasing computational efficiency and contextual accuracy [4, 7].

Feature preprocessing techniques such as image enhancement, noise reduction, tokenisation, stopword removal, and named entity recognition are applied to transform raw prescription inputs into machine-readable structured formats suitable for deep learning models [3, 4]. These preprocessing steps are critical for improving model generalisation across diverse handwriting styles and prescription formats.

For medication analysis and decision support, extracted textual features are combined with structured medical knowledge such as drug databases, interaction rules, and dosage guidelines to enhance system reliability and clinical safety [5, 8]. Such integration enables accurate detection of potential adverse drug interactions and supports intelligent medication scheduling, thereby improving patient adherence and reducing healthcare risks.

AI-Based Prescription Recognition and Medication Management Using Deep Learning

Several researchers have proposed AI-based approaches for automated medical document analysis using machine learning and deep learning techniques. Recent studies focus on improving prescription digitisation by leveraging transformer-based architectures and multimodal learning frameworks. Brown *et al*. [1] demonstrated that transformer-based models significantly improve clinical text extraction accuracy compared to traditional OCR systems. Similarly, Nguyen *et al*. [9] proposed an end-to-end deep learning framework for prescription digitisation, highlighting the importance of contextual understanding in medical text processing. However, system performance remains highly dependent on data quality, handwriting variability, and domain-specific training.

Wang *et al*. [3] introduced multimodal AI frameworks that integrate visual and textual features for healthcare document analysis, achieving improved robustness in handling handwritten and complex prescription layouts. Vision transformer-based OCR systems further enhance recognition accuracy by capturing spatial dependencies and layout structures within medical documents [6]. Additionally, knowledge-driven approaches and graph-based models enable semantic understanding of prescriptions and improve downstream clinical tasks such as drug interaction detection [5].

A. Multi-Entity Extraction from Prescriptions

Multi-entity extraction from prescription images is a critical component of intelligent healthcare systems, as prescriptions contain multiple interconnected entities such as drug names, dosage, frequency, and duration. Transformer-based NLP architectures enable contextual understanding of medical text and improve extraction accuracy by capturing semantic relationships between entities [4].

Recent studies emphasise multi-task learning frameworks that simultaneously extract multiple prescription attributes and support clinical decision-making processes [7]. These models improve efficiency by sharing representations across related tasks and reducing redundant computations. However, performance may degrade in cases of ambiguous handwriting, incomplete prescriptions, or non-standard abbreviations.

Text preprocessing strategies such as image enhancement, tokenisation, stop-word removal, lemmatisation, and named entity recognition are applied to improve model generalisation across diverse prescription formats [3, 4]. Additionally, data augmentation and normalisation techniques play a crucial role in reducing errors caused by variability in handwriting and image quality.

Recent advancements also explore attention mechanisms and contextual embedding strategies to enhance extraction accuracy in complex prescription scenarios. By leveraging contextual dependencies, transformer-based models can better interpret relationships between medications and instructions, improving overall system reliability [12].

B. Drug Interaction Detection and Medication Scheduling

Drug interaction detection is a critical aspect of patient safety and clinical decision support. Chen *et al.* [5] utilised NLP and knowledge graph-based approaches to identify potential adverse drug interactions from extracted prescription data. Similarly, graph-driven models improve interaction detection by analysing relationships between drugs, dosages, and patient-specific factors.

Medication scheduling systems generate personalised dosage plans based on extracted prescription details. Gupta *et al.* [7] demonstrated that AI-based scheduling systems significantly improve medication adherence and reduce human error. These systems incorporate rule-based logic and clinical constraints to ensure safe and effective treatment planning.

Despite strong performance, challenges such as incomplete prescription data, varying medical terminologies, and a lack of standardised formats remain significant barriers in real-world deployment. Additionally, ensuring accuracy and reliability in high-risk healthcare environments is essential for clinical adoption.

C. Explainable AI and Secure Healthcare Systems

Beyond prescription extraction and interaction analysis, recent research emphasises explainable AI (XAI) and secure data management in healthcare applications. Explainable models improve transparency by providing insights into how predictions are made, enabling better trust and adoption in clinical environments [8].

Secure healthcare systems incorporate role-based access control and encrypted data management to protect sensitive patient information. Das *et al.* [14] highlighted the importance of secure architectures in AI-driven healthcare systems, ensuring data privacy and compliance with regulatory standards.

Recent advancements also explore large language models and uncertainty-aware systems to improve prediction confidence and generalisation across diverse datasets [12]. Additional strategies such as ensemble modelling, attention mechanisms, and multimodal learning further enhance system robustness.

Despite these improvements, challenges such as data privacy concerns, model interpretability, and scalability continue to highlight the need for more reliable and adaptive AI-driven prescription recognition and medication management systems.

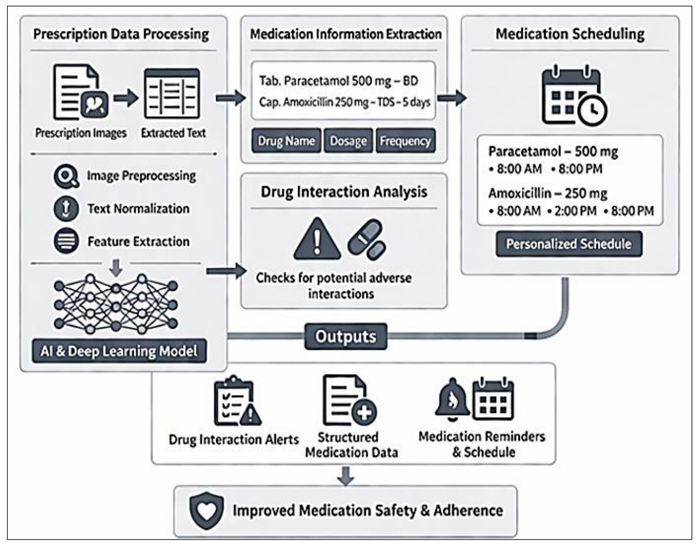

Fig 1: AI-Based Clinical Prescription Recognition and Medication Management Workflow

Comparative Analsis of Existing Approaches

AI-driven prescription recognition systems rely on machine learning models, medical datasets, and unstructured prescription images for information extraction and clinical decision support. While these approaches significantly improve automation and accuracy, they exhibit variations in methodology, data handling, interpretability, and scalability. A systematic comparison of existing studies reveals differences in modelling strategies, feature engineering techniques, and real-world deployment feasibility. Evaluating these approaches is essential to identify strengths, limitations, and research gaps in AI-based prescription recognition and medication management systems [1, 12].

Researchers categorise existing systems into traditional OCR-based methods, deep learning-based multimodal systems, and hybrid multi-task architectures. Each category differs in extraction accuracy, handling of handwritten text, interpretability, and computational complexity. Similar to other AI-driven healthcare systems, the choice of model architecture directly influences system reliability and applicability in clinical environments.

Some of the major comparative dimensions of AI-based prescription recognition systems are described as follows:

i **Traditional OCR vs. Deep Learning Models:**
Early approaches primarily relied on rule-based OCR and classical machine learning techniques for text extraction [2, 9]. While these methods are computationally efficient and easy to implement, they struggle with handwritten prescriptions and complex layouts. In contrast, deep learning architectures such as CNNs, transformers, and vision-language models automatically learn hierarchical representations from visual and textual data, resulting in significantly improved extraction accuracy and robustness [3, 6].

ii **Single-Task vs. Multi-Task Learning Frameworks:**
Many traditional systems perform tasks such as text extraction, drug identification, and scheduling independently, leading to redundant computation and limited contextual understanding. Multi-task learning frameworks address this limitation by simultaneously extracting prescription details, detecting drug interactions, and generating medication schedules within a unified architecture, improving efficiency and performance [4, 7].

iii **Handling of Data Variability and Noise:**
Prescription datasets often contain variations in handwriting styles, image quality, abbreviations, and incomplete information. Traditional models struggle to generalise across such variability. Techniques such as data augmentation, normalisation, and multimodal learning have been proposed to improve robustness and reduce recognition errors [3, 10].

iv **Multimodal and Context-Aware Modelling:**
Recent systems integrate both visual and textual features to enhance the understanding of prescription content. Vision transformers and multimodal AI frameworks capture spatial layout and contextual relationships, outperforming traditional text-only models [3, 6].

v **Explainability and Interpretability:**
Interpretability is critical in healthcare applications where incorrect predictions can lead to serious consequences. While traditional models offer higher transparency, deep learning systems often act as black boxes. To address this limitation, explainable AI techniques are integrated to provide insights into model predictions and improve trustworthiness [8, 12].

vi **Drug Interaction Analysis and Knowledge Integration:**
Some systems focus only on text extraction, whereas advanced frameworks incorporate clinical knowledge bases and graph-based models for drug interaction detection. These hybrid systems provide better decision support by combining extracted data with medical knowledge, improving patient safety [5, 15].

vii **Scalability and Real-World Deployment:**
Scalability varies across approaches. Traditional OCR systems require less computational power but lack accuracy, while deep learning and multimodal frameworks demand higher computational resources. Deployment feasibility depends on infrastructure, integration with healthcare systems, and data privacy considerations [3, 14].

Other comparative aspects include dataset diversity, evaluation metrics, multilingual support, and robustness to incomplete or ambiguous prescriptions. Although significant progress has been made in AI-based prescription recognition, challenges such as inconsistent data quality, limited interpretability, and a lack of fully integrated multi-task systems highlight the need for more robust, scalable, and clinically reliable solutions.

Challenges and Research Gaps in AI-Based Traffic Accident Analysis. The increasing adoption of deep learning-based frameworks for prescription recognition and medication management introduces significant challenges related to data quality, interpretability, scalability, and real-world applicability. Modern NLP architectures such as transformer-based models and multimodal vision-language systems operate as complex black-box models, making it difficult to interpret the reasoning behind predictions such as drug extraction, dosage identification, and interaction detection [3, 6, 12]. In real-world healthcare environments, this lack of transparency creates trust barriers between AI systems, clinicians, and patients. From a clinical perspective, system outputs must be explainable, reliable, and reproducible to support safe medical decision-making, making interpretability a critical requirement rather than an optional feature.

Another major challenge involves model bias and fairness. Prescription datasets may vary across hospitals, regions, and practitioners, leading to uneven representation of handwriting styles, medical terminologies, and drug usage patterns. This can result in biased predictions and reduced system reliability in diverse clinical settings. Ensuring fairness requires diverse training datasets, standardised annotation practices, and robust evaluation metrics.

One of the primary challenges in prescription recognition systems is the heterogeneity and inconsistency of prescription data. Medical prescriptions often contain illegible handwriting, abbreviations, incomplete instructions, and non-standard formats [2, 9]. Unstructured textual data further increases complexity, as similar medications or instructions may be expressed differently across prescriptions. Although NLP-based models improve automated extraction, their performance heavily depends on data quality, annotation accuracy, and domain-specific training. Handling multilingual prescriptions and adapting models across different healthcare systems remains an open research problem.

In real-world healthcare applications, AI-generated prescription outputs must be accurate, secure, and compliant with medical regulations. Systems must ensure data integrity, traceability, and auditability through secure storage, encrypted communication, and reliable logging mechanisms [14].

Protecting sensitive patient information such as medical history, prescribed drugs, and dosage details is essential for maintaining privacy and regulatory compliance.

Data imbalance and variability also present significant challenges. Certain drugs, dosage patterns, or prescription formats may appear less frequently, leading to skewed model learning and reduced performance on rare cases. While techniques such as data augmentation and transfer learning can mitigate these issues, maintaining consistent accuracy across diverse inputs remains difficult. Additionally, evaluation metrics such as accuracy alone may not reflect true system performance, emphasising the need for more comprehensive metrics.

Another research gap lies in the limited integration of multi-task learning frameworks that simultaneously perform text extraction, drug interaction detection, and medication scheduling. Many existing systems address these tasks independently, leading to inefficiencies and reduced contextual understanding [4, 7]. Developing unified architectures that combine visual, textual, and clinical knowledge remains an important direction for future research.

Explainability and trustworthiness also remain ongoing challenges. Although techniques such as attention visualisation and explainable AI methods provide insights into model decisions, they may increase computational complexity and still fail to fully capture clinical reasoning [8, 12]. Furthermore, the lack of standardised benchmarks for evaluating interpretability makes it difficult to compare models effectively.

Scalability and deployment feasibility further limit large-scale adoption. Deep learning models require significant computational resources, particularly when processing large volumes of prescription images. Integration with existing healthcare systems, data privacy regulations, and infrastructure constraints poses additional challenges. Cross-institution adaptability is also limited, as models trained on specific datasets may not generalise well to different hospitals or regions.

Although significant progress has been made in AI-based prescription recognition and medication management, several research gaps remain, including standardised datasets, unified multi-task architectures, improved handling of heterogeneous data, and stronger interpretability frameworks. Addressing these challenges is essential for developing robust, scalable, and trustworthy AI-driven healthcare systems capable of improving patient safety and clinical efficiency.

CONCLUSION

The effectiveness and reliability of AI-driven prescription recognition and medication management systems have become increasingly important with the rapid digitisation of healthcare and the growing need for accurate clinical data processing. Deep learning and Natural Language Processing (NLP) technologies, while highly effective in extracting prescription information and enabling clinical decision support, are not inherently flawless and depend on robust model design, high-quality data, and systematic validation processes. Similar to clinical decision-support systems that rely on probabilistic reasoning rather than absolute certainty, AI-powered prescription analysis frameworks must be strengthened through integrated multi-task architectures, balanced data handling strategies, explainable modelling techniques, and scalable deployment mechanisms.

It is therefore essential to develop systems with strong technical foundations, transparent decision-making pipelines, and standardised evaluation protocols to minimise errors arising from handwriting variability, incomplete prescriptions, contextual ambiguity, and model limitations. The discussions presented in this work highlight multiple dimensions of performance variation and research challenges across prescription extraction models, drug interaction detection systems, medication scheduling frameworks, and explainability mechanisms.

At the same time, this paper outlines emerging directions such as unified multi-task learning architectures, multimodal vision-language models, knowledge graph integration, and interpretable AI frameworks that can enhance system reliability and clinical usability. Strengthening these aspects will play a crucial role in advancing intelligent, data-driven, and secure healthcare systems capable of improving patient safety, reducing medication errors, and enabling efficient clinical workflows.

REFERENCES

1. Brown J, *et al*. Transformer-based medical document understanding for clinical text extraction. IEEE Access. 2024;12:45231–45245.
2. Kumar A, Singh R. Deep learning approaches for handwritten prescription recognition. Pattern Recognition Letters. 2024;182:45–52.
3. Wang L, *et al*. Multimodal AI for healthcare document analysis using vision-language models. Information Fusion. 2025;105.
4. Patel S, *et al*. Clinical NLP for medication extraction using transformer models. Journal of Biomedical Informatics. 2025;145.
5. Chen M, *et al*. Drug interaction detection using NLP and knowledge graphs. IEEE Journal of Biomedical and Health Informatics. 2025;29(2).
6. Zhao Y, *et al*. Vision transformers for medical OCR systems. IEEE Transactions on Artificial Intelligence. 2024.
7. Gupta R, *et al*. AI-based medication scheduling and adherence systems. Artificial Intelligence in Medicine. 2025;152.
8. Lee H, *et al*. Explainable AI in healthcare systems: Methods and applications. IEEE Access. 2024;12.
9. Nguyen T, *et al*. End-to-end deep learning framework for prescription digitisation. Computers in Biology and Medicine. 2024;168.
10. Sharma K, *et al*. Handwritten text recognition using CNN-transformer hybrid models. Expert Systems with Applications. 2025.

11. Roy P, *et al*. Multilingual medical OCR using large language models. ACM Transactions on Intelligent Systems and Technology. 2025.
12. Li X, *et al*. Large language models for clinical information extraction. Nature Digital Medicine. 2025.
13. World Health Organisation. Global patient safety report: Medication errors and digital health. Geneva: World Health Organisation; 2026.
14. Das S, *et al*. Secure healthcare data management using role-based access control and AI. IEEE Transactions on Cloud Computing. 2024.
15. Park J, *et al*. Knowledge graph-based drug interaction prediction. IEEE Transactions on Knowledge and Data Engineering. 2025.
16. Hou Q, *et al*. Deep transfer learning for medical document recognition. Computers and Electrical Engineering. 2025.
17. Zhang Y, *et al*. Vision-language models for document AI in healthcare. IEEE Transactions on Multimedia. 2025.
18. Verma N, *et al*. Robust OCR pipelines for noisy document images. IEEE Access. 2024.
19. Mehta A, *et al*. Clinical decision support using AI and NLP. Artificial Intelligence Review. 2025.
20. Kim D, *et al*. Transformer-based sequence labelling for medical entity recognition. IEEE Access. 2024.
21. Iyer S, *et al*. Deep learning for healthcare data extraction and analytics. Future Generation Computer Systems. 2025.
22. Hassan M, *et al*. Privacy-preserving AI in healthcare systems. IEEE Transactions on Dependable and Secure Computing. 2025.

About the corresponding author

Sundari Veerappan is an Assistant Professor in the Department of Computer Science and Engineering at Meenakshi Sundararajan Engineering College (MSEC). She holds a B.E. from Periyar Maniammai College and a PG degree from Anna University. With over 15 years of teaching experience, her interests include Machine Learning, Deep Learning, and Computer Vision.