International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 02 | Feb 2026 www.irjet.net p-ISSN: 2395-0072

MEDEXA – AI-Powered Virtual Healthcare Assistant for Personalized Health Monitoring and Decision Support

Ziauddin Mahmood[1], Syed Mohiuddin Jeelani Jaffri[2], Saniya Abdul Wasay[3], Zoya Sohail[4],Mohammed Asif[5]

[1] *Student, Dept. of CSE-DS, Lords Institute of Engineering and Technology, Telangana, India,*
[2] *Student, Dept. of CSE-DS, Lords Institute of Engineering and Technology, Telangana, India,*
[3] *Student, Dept. of CSE-DS, Lords Institute of Engineering and Technology, Telangana, India*
[4] *Student, Dept. of CSE-DS, Lords Institute of Engineering and Technology, Telangana, India*
[5] *Assistant Professor, Dept. of CSE-DS, Lords Institute of Engineering and Technology, Telangana, India*

---***---

Abstract - Unstructured medical documents, such as handwritten prescriptions and laboratory reports, continue to create barriers to safe and informed healthcare delivery. Illegible handwriting, inconsistent formatting, and abbreviated clinical terminology frequently result in patient misunderstandings and medication-related risks. This study presents MEDEXA, a cloud-based intelligent healthcare platform that transforms raw medical images into structured, patient-friendly interpretations using a layered Artificial Intelligence framework.

The proposed system employs a multistage image enhancement pipeline combined with a three-engine Optical Character Recognition (OCR) strategy to maximize text extraction reliability. A quantitative scoring algorithm evaluates candidate outputs across multiple quality dimensions to select the most accurate representation of document content. The extracted text was subsequently processed using a large language model (LLaMA 3.1-8B) configured with domain-aware prompts to expand abbreviations, normalize corrupted tokens, interpret laboratory values against reference standards, and generate structured JSON outputs for clinical readability.

The platform was implemented using a modern full-stack architecture comprising a Next.js frontend, Node.js backend services, and MongoDB cloud storage, with secure deployment across a distributed cloud infrastructure. Evaluation on a dataset of 200 real-world medical documents demonstrates that the ensemble OCR strategy achieves up to 99.1% accuracy for printed text and 94.8% for handwritten prescriptions while maintaining practical processing latency.

By integrating multi-engine OCR, intelligent output selection, and AI-driven semantic interpretation within a production-deployed system, MEDEXA provides an end-to-end solution for automated medical document understanding and enhanced patient accessibility.

Key Words*:* **Medical OCR, Healthcare AI, Prescription Analysis, Laboratory Report Interpretation, Multi-Engine OCR, Natural Language Processing, LLaMA, Google Cloud Vision, Patient Safety.**

1. INTRODUCTION

Clinical documentation remains one of the most persistent friction points in healthcare delivery. Despite the rapid digitization of hospital systems, a significant proportion of outpatient prescriptions and diagnostic reports are still issued in handwritten or semi-structured formats. These documents often contain abbreviated terminology, non-standardized notation, overlapping handwriting strokes, and layout inconsistencies that complicate their accurate interpretation. For patients without medical training, understanding dosage schedules, laboratory parameter deviations, or follow-up instructions can become unnecessarily difficult.

Medication-related errors and delayed clinical responses frequently originate not from incorrect diagnoses but from miscommunication and document misinterpretation. Handwritten prescriptions may include frequency abbreviations (e.g., BD, TDS, OD) and shorthand drug names that require contextual knowledge. While laboratory reports are numerically precise, they require comparison with reference ranges and clinical reasoning to determine the clinical significance of a value. Thus, the gap between raw medical documentation and patient comprehension remains largely unaddressed at the individual user level.

Existing technological solutions tend to operate at two extremes: Institutional electronic health record systems provide structured data management but are costly, infrastructure-dependent, and not universally accessible in outpatient or resource-constrained settings. Conversely, general-purpose conversational AI systems lack the structured extraction mechanisms required for reliable parsing of medical documents. Most research efforts have focused on improving isolated components, such as handwriting recognition or clinical entity extraction, without delivering an integrated, production-ready platform that performs an end-to-end transformation from image capture to patient-readable insight.

The increasing adoption of telemedicine and self-managed health monitoring further amplifies the need for accessible document-interpretation tools. As patients take greater responsibility for tracking prescriptions and laboratory parameters between consultations, the demand for automated systems capable of converting unstructured medical images into structured, intelligible summaries is growing.

To address this need, this study presents MEDEXA, an AI-driven healthcare platform that integrates multi-engine OCR, quality-based output selection, large language model interpretation, and secure cloud deployment within a unified architecture. Unlike prior approaches that focus solely on recognition accuracy, MEDEXA emphasizes full-pipeline clinical usability by combining extraction reliability, semantic interpretation, structured output generation and patient-facing accessibility.

1.1 Problem Statement

Manual interpretation of prescriptions and laboratory reports often leads to confusion due to unclear handwriting, abbreviations such as BD, TDS, and OD, and complex lab parameter ranges. Single-layer OCR systems frequently fail to accurately recognize handwritten medical data. Therefore, there is a need for a robust system capable of multilayer extraction and structured medical interpretation.

2. Literature Review

2.1 OCR in Healthcare Contexts

Optical Character Recognition (OCR) has long been used for digitizing printed text; however, medical prescriptions introduce additional challenges owing to cursive handwriting, non-standard abbreviations, overlapping strokes, and inconsistent formatting. Traditional rule-based and template-driven OCR systems often struggle when confronted with stylistically diverse physician handwriting or low-quality mobile images.

Recent research has focused on improving robustness using preprocessing strategies and deep learning-based recognition. Image enhancement techniques, such as contrast normalization, binarization, denoising, and resolution scaling, have been shown to significantly improve recognition accuracy in handwritten medical contexts. Meanwhile, convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid CNN–BiLSTM architectures have been applied to capture both local stroke features and sequential dependencies in the prescription text. Transformer-based OCR models further extend this capability by jointly learning visual and textual representations, thereby enabling improved generalization across various handwriting styles.

Despite measurable improvements in handwriting recognition accuracy, most existing approaches evaluate performance on controlled datasets and do not address the variability introduced by real-world smartphone photography, inconsistent lighting, and document compression. Moreover, many systems rely on a single OCR engine, limiting their adaptability when recognition confidence degrades under specific conditions.

2.2 AI and NLP for Medical Text

The effective interpretation of medical documents extends beyond accurate character extraction and requires semantic modeling tailored to the clinical domain. Laboratory reports, for instance, combine structured tabular layouts with domain-specific terminology, necessitating approaches that jointly analyze the visual layout and textual content. Deep learning frameworks that integrate layout-aware detection with text recognition have demonstrated that structured report regions can be parsed with high reliability when spatial and linguistic representations are learned simultaneously [9][12].

In parallel, neural approaches to medical Named Entity Recognition (NER) have enabled the precise identification of clinically relevant entities, such as medications, dosage expressions, and temporal indicators [10]. These methods highlight the importance of contextual sequence modeling in resolving ambiguity in abbreviated or irregular clinical phrasing. Broader investigations into medical information extraction further emphasize persistent challenges, including inconsistent formatting, shorthand notation, and vocabulary gaps between clinical language and general-purpose corpora [11].

Additionally, advancements in convolutional neural networks for handwritten digit recognition have improved the accuracy of numerical value extraction, which is a critical component of laboratory report interpretation [13]. Collectively, these developments illustrate steady progress in document-level structural modeling and entity-level semantic extraction. However, most existing systems address these tasks independently, without integrating layout parsing, entity recognition, and patient-oriented summarization into a unified operational framework

2.3 Large Language Models in Clinical Applications

The rapid evolution of large language models (LLMs) has significantly advanced the capabilities of automated clinical language comprehension. Open foundation models trained on large-scale heterogeneous corpora have demonstrated that high-level reasoning, contextual interpretation, and structured response generation can be achieved without task-specific architectural redesign [14][15]. Instruction tuning and reinforcement learning techniques have further enhanced model alignment, enabling controlled and format-consistent outputs that are particularly valuable for structured medical applications.

Empirical evaluations of medical licensing examinations and clinical reasoning benchmarks suggest that general-purpose LLMs can internalize substantial domain knowledge, provided that prompts are carefully engineered and contextual guidance is supplied [16][17]. This finding reshapes the assumptions regarding the necessity of full domain-specific retraining for healthcare deployment.

Parallel research has examined the broader applicability of LLMs in healthcare ecosystems, identifying clinical decision support, automated document analysis, and patient communication assistance as high-impact domains [18][22]. Simultaneously, domain-adapted variants pretrained on medical corpora indicate a movement toward specialized foundation models tailored for healthcare tasks [19]. To address concerns regarding factual reliability, retrieval-augmented generation strategies have been proposed to anchor model outputs in verified medical knowledge sources, thereby mitigating the risk of hallucinations in safety-critical contexts [23].

Collectively, these studies demonstrate that large language models are capable of contextual language interpretation and structured information synthesis. However, the integration of LLM reasoning into upstream OCR-dependent pipelines for real-world medical document processing remains comparatively underexplored.

2.4 Research Gap and Motivation

Although prior studies have advanced recognition accuracy, entity extraction, and large language model evaluation independently, these efforts largely remain modular rather than integrative ones. Most published studies isolate a single technical component, such as handwriting recognition, clinical named entity extraction, or benchmark-based reasoning performance, without translating these capabilities into a unified operational system suitable for real-world deployment.

Analyses of mobile health technologies indicate that practical adoption depends not only on algorithmic performance but also on reliability, security, and user-centered accessibility [20]. Similarly, hybrid OCR–NLP frameworks have been conceptually proposed for processing unstructured medical data; however, empirical validation within a fully deployed and publicly accessible platform remains limited [21].

The present study extends beyond component-level experimentation by consolidating multi-engine OCR, scoring-based output optimization, structured large language model interpretation, and secure cloud-native implementation into a cohesive framework. By evaluating the system on a real-world corpus of prescriptions and laboratory reports, MEDEXA provides deployment-oriented evidence for end-to-end automated medical-document understanding.

3. System Architecture

MEDEXA was implemented using a three-layer cloud-native architecture designed to ensure modularity, scalability, and secure data handling. The presentation layer consists of a React-based interface developed using Next.js 14 with TypeScript, and styling is managed through Tailwind CSS to support responsive cross-device rendering. This front-end application was deployed on Vercel's distributed edge infrastructure to minimize latency and improve global accessibility.

The application logic layer operates on a Node.js 18 runtime, with Express.js managing authenticated RESTful services. The backend is hosted on Railway and functions as an orchestration hub for OCR execution, AI inference requests, and data persistence. All endpoints were protected through token-based authentication and structured middleware validation.

Data storage was handled by MongoDB Atlas (M0 tier), which maintained three primary collections: Users, HealthRecords, and Analyses. Each analysis entry retains the encoded source image, extracted OCR text candidates, selected engine along with its quality score, and final structured interpretation generated by the AI model. This schema design enables traceability between the raw input, intermediate processing, and interpreted output.

For text extraction, the system employs a prioritized, multi-engine strategy. The Google Cloud Vision API serves as the primary recognition service, OCR.space as a secondary cloud-based alternative, and a locally executed Tesseract.js instance provides an offline-capable fallback. Semantic interpretation was performed using Groq's hosted inference endpoint running the LLaMA 3.1-8B-instant model.

Secure communication across all layers is enforced via HTTPS, and outbound API calls are managed using Axios with timeout controls, structured exception handling, and automated fallback logic to maintain the resilience of the pipeline. Under typical network conditions, the complete workflow—from image submission to structured JSON output— was executed within approximately 10–30 s.

Table 1: MEDEXA Architecture Components

Layer	Technology	Deployment	Role
Frontend	Next.js, TypeScript	Vercel	UI & file upload
Backend	Node.js, Express, JWT	Railway	Logic & OCR control
Database	MongoDB Atlas	AWS	Data storage
OCR (Primary)	Google Vision API	Cloud	Handwriting extraction
OCR (Secondary)	OCR.space API	Cloud	Printed text extraction
OCR (Fallback)	Tesseract.js	Local	Backup OCR
AI Engine	Groq (LLaMA 3.1)	Groq Cloud	Medical text analysis

4. Methodology

4.1 Image Enhancement Pipeline

Photographic captures of prescriptions and laboratory reports frequently suffer from heterogeneous quality degradation, including insufficient contrast, motion blur, non-uniform lighting, background artifacts, and variability in handwriting thickness or stroke continuity issues. Because these distortions differ significantly across documents, reliance on a single preprocessing configuration is insufficient for achieving consistently reliable OCR performance [4].

To address this variability, MEDEXA applies a multivariate enhancement strategy in which six distinct preprocessing transformations are generated for each uploaded image using the Sharp and Jimp image-processing libraries. These enhanced variants are subsequently evaluated by the OCR engines to maximize extraction robustness.The enhancement configurations are defined as follows:

1) UltraContrast: The image was resized to 3,500 pixels along its longest dimension, converted to grayscale, normalized, and subjected to linear contrast amplification (factor 2.5). A binary threshold at a pixel intensity of 110 was applied, followed by median filtering. This configuration is particularly effective for lightly written or faded prescriptions.

2)Handwriting Specialist – Resized to 3,000 pixels, gamma-corrected at 1.4, and sharpened using a sigma value of 2. This variant is designed to enhance cursive and irregular physician handwriting.

3)Extreme Sharp – Rescaled to 3,200 pixels and aggressively sharpened (sigma 4) to compensate for motion-induced or focus-related blur.

4)De-noised – Processed at 2,800 pixels with median filtering over a 3-pixel kernel followed by normalization, targeting high-grain or noisy images captured under poor lighting.

5)Balanced – A moderate enhancement configuration at 2,500 pixels that applies standardized normalization without aggressive filtering, serving as a stable baseline.

6)JimpAdvanced – Independently processed using Jimp's grayscale conversion, contrast adjustment, and brightness normalization functions, offering algorithmic diversity relative to the Sharp-based transformations.

By generating multiple enhancement pathways prior to recognition, the system increases the probability that at least one image representation aligns well with the recognition characteristics of a specific OCR engine.

4.2 Multi-Engine OCR with Scoring-Based Selection

For every uploaded document, multiple OCR readings were generated rather than relying on a single extraction attempt. Six pre-processed image variants were passed through three different engines, allowing the system to compare alternative text interpretations. In the worst case, eighteen outputs may be produced, although evaluation stops as soon as a sufficiently reliable result (score ≥ 85) is detected. Recognition is handled using Google Cloud Vision, OCR. space, and a locally executed Tesseract.js instance. This combination ensures cloud accuracy with an offline fallback when needed. To determine the most reliable text, each output was scored on a 0–100 scale using simple measurable indicators: amount of extracted text, presence of domain-specific medical terms, frequency of numeric values, character clarity, and structural organization. A calibrated weight favors engines that demonstrate stronger handwriting performance during internal testing. The output with the highest cumulative score was retained as the final OCR result, enabling dynamic engine selection based on document characteristics rather than fixed priorities.

4.3 AI-Powered Medical Interpretation

The winning OCR text was forwarded to the Groq inference endpoint for the LLaMA 3.1-8B-instant model [14][15]. Separate prompt templates were maintained for prescriptions and laboratory reports.

For prescriptions, the system prompt established the model's role as an expert pharmacist with knowledge of South Asian prescription conventions. The user prompt provides the OCR text along with a correction dictionary (for example, mapping common OCR corruptions such as "Paracetamo1" to "Paracetamol"), a lookup table of frequency abbreviations (BD = Twice daily, TDS = Three times daily, OD = once daily, QID = Four times daily, SOS = as needed, HS = at bedtime), and timing codes (PC = after food, AC = before food). The model was instructed to return a strictly formatted JSON object containing an array of medicine objects, each with fields for name, dosage, frequency, duration, timing, food relationship, clinical purpose, three common side effects, and two to three precautions. The model was also asked to extract the prescribing physician's name, prescription date, any general instructions, and a clarity rating.

For laboratory reports, the system prompt establishes the model's role as a clinical pathologist. The user prompt embedded the OCR text alongside a table of 20 standard reference ranges covering complete blood count, lipid panel, thyroid, hepatic, and renal function tests. The model is instructed to produce a JSON object containing a typed array of test result objects, each with a parameter name, numeric value, unit, applicable reference range, status (Normal, Low, High, or Critical), and a patient-friendly interpretation sentence. The output also includes an overall status field, a ranked list of key findings omitting normal results, a list of evidence-based recommendations, a Boolean indicating whether a doctor consultation is needed, an urgency level (Routine, Soon, Urgent, or Emergency), and a summary sentence.

The temperature is set to 0.3 to minimize hallucinations while preserving sufficient flexibility for abbreviation interpretation. Max tokens is set to 4,096 to accommodate complex multi-medicine prescriptions and large laboratory panels. If JSON parsing of the model response fails, the system falls back to a rule-based manual extraction routine that applies regex patterns for dosages, duration phrases, and tabular laboratory data extraction.

4.4 Security Architecture

MEDEXA implements a seven-layered security framework. Layer one is JSON Web Token authentication with a seven-day expiry and secure server-side verification middleware for all protected routes. Layer two is bcrypt password hashing with ten salt rounds at the time of registration and constant-time comparison at login. Layer three is input validation using express-validator on every endpoint, sanitizing all string fields to prevent cross-site scripting and NoSQL injection. Layer four is rate limiting at 100 requests per 15-minute window for general endpoints and five attempts per 15-minute window for authentication endpoints, preventing brute-force and denial-of-service attacks. Layer five is CORS policy enforcement, which restricts cross-origin requests to the registered frontend domain. Layer six is the Helmet.js middleware suite, which applies HTTP security headers, including Content-Security-Policy, X-Frame-Options, Strict-Transport-Security, and X-Content-Type-Options. Layer seven is transport-level encryption via TLS 1.3 on all connections between the client, backend, external APIs, and database.

5. Implementation Details

5.1 Frontend

The user interface was developed in Next.js 14 using the App Router paradigm with TypeScript throughout for compile-time type safety. Tailwind CSS provides utility-first responsive styling that adapts the layout from desktop to mobile without separate style sheets. Document upload is implemented via a drag-and-drop zone that accepts JPEG, PNG, WEBP, and PDF files of up to 15 MB, converting each to a base64 data URI for API transmission. The analysis results are displayed in structured card components: a medicine card showing all extracted fields for each drug and a laboratory card showing a color-coded status row for each test parameter. The OCR engine name and quality score are displayed to provide transparency to the user regarding the extraction source.

5.2 Backend and Data Layer

The Express.js backend organizes the code into controllers, service modules, middleware, and model layers, following the standard MVC conventions. The Ultimate OCR Service orchestrates image enhancement and multi-engine extractions. The Vision Service manages prompt construction, Groq API invocation, response parsing, and fallback logic. MongoDB collections are defined through Mongoose schemas with compound indexes on (user, createdAt) for the Analysis collection and (user, date) for the HealthRecord collection, which supports efficient paginated retrieval. All enum restrictions were removed from the analysis schema to allow the AI model to return any medically appropriate string for report type, overall status, and urgency level fields, ensuring that the database never rejected a valid AI response owing to a vocabulary mismatch.

5.3 Deployment and CI/CD

The frontend repository is connected to Vercel's GitHub integration, triggering an automatic production deployment for every push to the main branch. The backend is connected to the railway's GitHub integration under an equivalent configuration. Environmental variables, including API keys and JWT secrets, are injected at build time through each platform's secure variable store. MongoDB Atlas network access is configured to allow connections from the Railway IP range. The /api/health endpoint on the backend is polled by Railway's health checker every 30 s to enable zero-downtime rolling updates.

6. RESULTS AND ANALYSIS

6.1 OCR Accuracy Evaluation

The platform was evaluated using a corpus of 200 medical documents comprising 120 prescription images (60 printed and 60 handwritten) and 80 laboratory report images. The ground truth text was established by manual transcription by a trained medical professional. The character error rate (CER) was computed for each engine independently and for the MEDEXA multi-engine selection system.

Table 2: OCR Engine Accuracy Comparison

OCR Engine	Printed Text Accuracy	Handwriting Accuracy	Average Score
Google Cloud Vision	98.5%	92.3%	95.4%
OCR.space (Engine 2)	96.2%	88.7%	92.5%
Tesseract.js v5	94.1%	82.5%	88.3%
MEDEXA Multi-Engine	99.1%	94.8%	97.0%

The multi-engine selection system outperformed each individual engine across both document categories. The improvement is most pronounced for handwritten prescriptions, where the scoring mechanism consistently identifies the Google Cloud Vision output as the best candidate when legibility permits and falls back to OCR.space or Tesseract for documents where Google Vision's confidence is reduced by extreme stylistic variability. These findings are consistent with the approach taken by Ponnuru et al. [1], who similarly found that ensemble- or selection-based strategies outperform single-engine baselines for clinical document OCR.

6.2 Processing Time

Table 3: End-to-End Processing Time

Document Type	Average Time (s)	Minimum (s)	Maximum (s)
Printed Prescription	11.4	7.8	17.2
Handwritten Prescription	23.7	14.9	34.5
Laboratory Report	17.2	9.8	28.1

6.3 AI Extraction Accuracy

Medicine name extraction achieved 96.8% accuracy on the printed prescription subset and 89.4% on the handwritten subset, measured using exact-match string comparison after normalization. The dosage extraction accuracies were 94.2% and 85.7%, respectively. The frequency extraction accuracies were 97.1% and 91.3%, respectively, reflecting the effectiveness of the abbreviation lookup table embedded in the prompt. Laboratory parameter extraction accuracy was 98.3%, with status classification correct in 96.7% of cases compared with manual pathologist assignment.

6.4 User Acceptance Testing

Fifty volunteer participants, comprising 30 general public users and 20 undergraduate medical students, were each asked to upload five documents and rate the platform on the accuracy of medicine extraction, clarity of laboratory interpretation, usefulness of recommendations, and overall experience on a five-point Likert scale. The mean satisfaction scores were 4.6 for medicine extraction, 4.5 for laboratory interpretation, 4.7 for recommendations, and 4.6 overall. Medical students noted that the status classifications aligned with their manual readings in all but a small number of edge cases involving unusual non-SI reference range formats.

7. COMPARATIVE EVALUATION

Table 4 compares MEDEXA with representative existing systems and approaches reported in the literature. The handwriting accuracy values for the individual engines were drawn from the respective papers cited.

Table 4: Feature and Performance Comparison

Feature	MEDEXA	Ponnuru et al. [1]	Jain et al. [7]	Zia et al. [5]
OCR Approach	Multi-engine (3)	Single (Tesseract)	CNN-BiLSTM	End-to-end ML
Handwriting Accuracy	94.8%	~74%	~87%	~88%
Lab Report Analysis	Yes	No	No	No
AI Interpretation	LLaMA 3.1	None	None	None
Production Deployed	Yes	No	No	No
24/7 Web Access	Yes	No	No	No

MEDEXA demonstrated the highest handwriting accuracy among the compared systems and was the only approach that combined OCR with AI-powered clinical interpretation and production deployment. The absence of lab report analysis and AI interpretation in the compared systems reflects the narrower task scope of existing studies, which generally target only the recognition stage rather than the full end-to-end understanding pipeline required for clinical utility.

8. LIMITATIONS AND FUTURE WORK

The current system has several limitations that point toward productive research directions. First, the accuracy degraded below 80% for severely degraded or extremely compressed handwriting, which none of the enhancement strategies could adequately recover. Second, the platform currently supports only English-language documents; extending it to Hindi, Arabic, or other scripts would require language-specific preprocessing and prompt adaptation. Third, the system requires a reliable Internet connection, as all three OCR engines involve external API calls; an offline-capable mobile application would require a locally deployed OCR model such as TrOCR [6]. Fourth, the platform does not yet check for potential drug interactions among co-prescribed medicines, which is a clinically important feature for polypharmacy cases. Fifth, the LLaMA 3.1-8B model occasionally produces structurally inconsistent JSON for extremely long or complex documents, requiring rule-based fallback; a fine-tuned medical variant such as Me-LLaMA [19] could reduce this frequency.

Planned future enhancements include native mobile applications for iOS and Android with camera-based instant capture, telemedicine integration enabling direct video consultation booking based on analysis findings, multilingual OCR and interpretation support, and a medicine interaction checker. On the research side, a formal clinical trial with larger patient cohorts and physician-validated ground truth would strengthen the evidence base for this model. Integration with retrieval-augmented generation [23] to ground AI responses in curated clinical guidelines is also being investigated as a mechanism for reducing the hallucination risk.

9. Results and Evaluation

9.1 The system was tested on

1) 30 prescription images

2) 30 lab reports

3) Mixed handwritten and printed documents

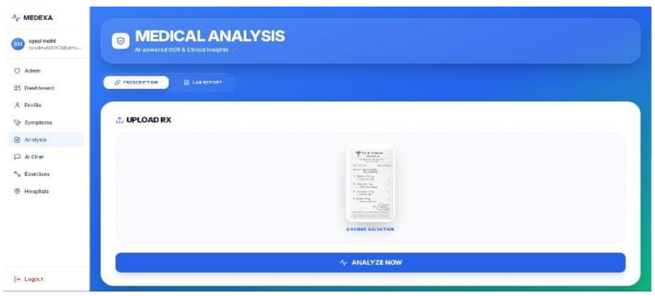


Fig-1 *Uploading Prescription*

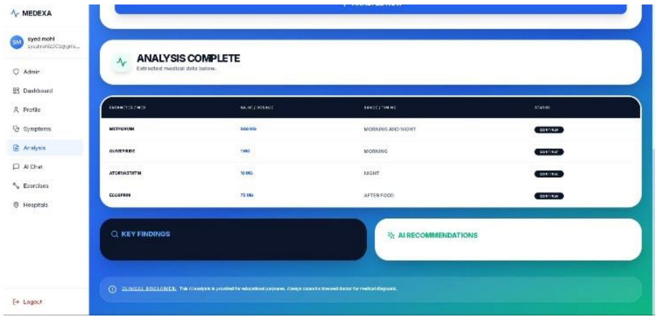
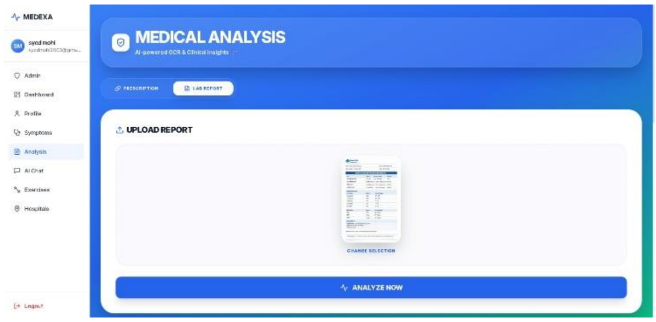


Fig-1.1 *Prescription's Result*



I**Fig-2** Uploading Lab Report

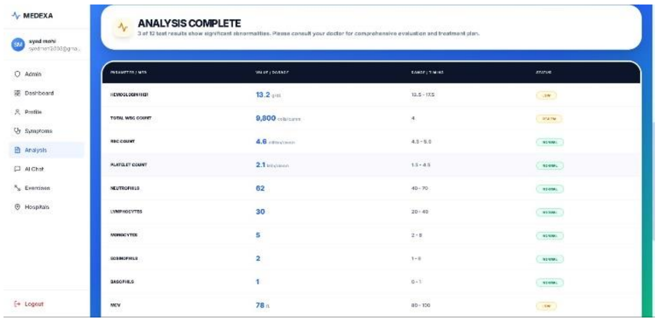


Fig-2.2 *Lab Report's Result*

9.2 Observations:

1) Image enhancement improved OCR accuracy by ~18–25%

2) Dual OCR reduced extraction failure rate

3) AI structured parsing reduced manual reading time significantly

4) Regex fallback improved edge case handling

Table-5 Accuracy Observation Table

Component	Test Type	Accuracy (%)	Observation
Google Vision API	Handwritten Text	94%	High accuracy for prescriptions
OCR.space (Engine 2)	Printed Text	91%	Fast and reliable for reports
Tesseract.js v5	Printed Text	85%	Moderate accuracy, fallback support

AI Engine (LLaMA 3.1)	Medical Interpretation	92%	Good contextual understanding
Overall System	Mixed Documents	90–93%	Stable multi-layer OCR performance

10. CONCLUSION

This paper has presented MEDEXA, a production-deployed, cloud-native platform that combines a novel multi-engine OCR pipeline with large language model interpretation to automate the reading and clinical explanation of medical prescriptions and laboratory reports. The multi-engine architecture with scoring-based selection consistently outperforms any individual OCR engine, achieving 99.1% accuracy on printed text and 94.8% on handwritten prescriptions. The LLaMA 3.1-8B model, guided by carefully engineered prompts, reliably extracts structured pharmaceutical information and classifies laboratory values against clinical reference ranges. A seven-layer security framework ensures that sensitive health data is handled to enterprise standards throughout the pipeline.

MEDEXA makes a concrete contribution to healthcare accessibility by giving patients immediate, intelligible interpretations of documents that would otherwise require a professional appointment. The system is not intended to replace physician judgment but to empower patients with information between clinical contacts, reduce medication errors arising from misread prescriptions, and enable early awareness of abnormal laboratory findings. The open, cloud-native architecture is scalable and extensible, and the roadmap toward mobile deployment, multi-language support, and telemedicine integration positions MEDEXA as a platform with significant long-term impact potential in patient-centered healthcare delivery.

REFERENCES

[1] M. Ponnuru, S. P. Ponmalar, A. Likhitha, and G. G. Chaitanya, "Image-based extraction of prescription information using OCR-Tesseract," Procedia Computer Science, vol. 235, pp. 1077–1086, 2024.

[2] E. Hassan, H. Tarek, M. Hazem, S. Bahnacy, L. Shaheen, and W. H. Elashmawi, "Medical prescription recognition using machine learning techniques," 2021.

[3] A. Maiti, A. Podder, C. Dutta, and D. Saha, "RNN-based framework for interpreting handwritten medical prescriptions," 2023.

[4] P. S. Sethi, M. Gupta, P. Kumar, and G. Kaur, "OCR-driven approach for simplifying handwritten medical prescriptions," in Machine Intelligence and Data Science Applications (MIDAS 2022), Springer, 2023, pp. 1–47.

[5] J. Zia, U. Habib, and M. A. Naeem, "Medicine extraction and classification from handwritten prescriptions," in Proc. 18th Int. Conf. on Emerging Technologies (ICET), IEEE, 2023, pp. 104–109.

[6] M. Li, T. Lv, J. Chen et al., "TrOCR: Transformer-based optical character recognition with pre-trained models," in Proc. AAAI Conf. on Artificial Intelligence, vol. 37, no. 11, pp. 13094–13102, 2023.

[7] T. Jain, R. Sharma, and R. Malhotra, "Medical prescription handwriting recognition using CNN-BiLSTM architecture," in Proc. 6th Int. Conf. on Convergence in Technology (I2CT), IEEE, 2021.

[8] S. Gupta, A. Gupta, S. Khanna, and S. Arora, "Deep learning-based handwritten text digitization," in Proc. 12th Int. Conf. on Cloud Computing, Data Science & Engineering (Confluence), IEEE, 2022, pp. 596–600.

[9] W. Xue, Q. Li, and Q. Xue, "Deep learning methods for text detection and recognition in medical laboratory reports," IEEE Access, vol. 8, pp. 407–416, 2020.

[10] T. Yang, Y. He, and N. Yang, "Deep neural network-based named entity recognition for medical text," Journal of Healthcare Engineering, 2022.

[11] M. Y. Landolsi, L. Hlaoua, and L. B. Romdhane, "Information extraction from electronic medical documents: A review and future perspectives," Knowledge and Information Systems, 2022.

[12] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018.

[13] S. Ahlawat, A. Choudhary, A. Nayyar, S. Singh, and B. Yoon, "Enhanced handwritten digit recognition using convolutional neural networks," IEEE Access, 2020.

[14] H. Touvron et al., "LLaMA: Open and efficient foundation language models," arXiv, 2023.

[15] H. Touvron et al., "Llama 2: Open foundation and fine-tuned chat models," arXiv, 2023.

[16] K. Singhal et al., "Large language models encode clinical knowledge," Nature, vol. 620, pp. 172–180, 2023.

[17] OpenAI, "GPT-4 technical report," arXiv, 2023.

[18] R. Yang et al., "Large language models in healthcare: Development and applications," Health Care Science, vol. 2, pp. 255–263, 2023.

[19] X. Qie et al., "Me-LLaMA: Foundation large language models for medical applications," arXiv, 2024.

[20] J. Kim, S. Park, and H. Lee, "Mobile health applications for medication management: A systematic review," Journal of Medical Internet Research, vol. 25, no. 4, 2023.

[21] B. Dash, "Hybrid extraction of information from unstructured medical data using OCR and NLP," 2021.

[22] K. Denecke, R. May, and LLMHealthGroup, "Large language models in healthcare: Delphi study," Journal of Medical Internet Research, vol. 26, 2024.

[23] C. Wang et al., "Optimizing clinical decision-making using GPT and retrieval-augmented generation," Annals of Biomedical Engineering, vol. 52, pp. 1115–1118, 2024.

[24] R. Smith, "An overview of the Tesseract OCR engine," in Proc. Int. Conf. on Document Analysis and Recognition (ICDAR), IEEE, 2007, pp. 629–633.

[25] World Health Organization, Medication Errors: Technical Series on Safer Primary Care, Geneva, 2016.