Global Journal of Engineering Innovations & Interdisciplinary Research

Original Article

Semantic Understanding And Contextual Explanation of Medical Documents

Ch Mahender Reddy[1], K Sai Varun Reddy[2], M Uday Sreechand[2], TN Sai Ram[2], T Ajith Karthikeya[2]

[1]Assistant Professor, Department of CSE, GITAM University, Hyderabad, India
[2]Undergraduate, Department of CSE, GITAM University, Hyderabad, India

Correspondence

Ch Mahender Reddy

Assistant Professor, GITAM University, Hyderabad, India

Abstract

This study introduces an AI-driven system designed to efficiently extract and summarize medical documents using Optical Character Recognition (OCR) and Natural Language Processing (NLP). By leveraging TesseractOCR, the system accurately retrieves textual data from diagnostic reports. The extracted content is then processed through a GPT-based API to generate brief, patient-friendly summaries in 2-3 lines. By simplifying complex medical terms, the system enhances patient understanding, supports informed decision-making, and improves communication between healthcare professionals and patients. This method seeks to bridge the gap between technical medical reports and patient awareness, ultimately contributing to better healthcare outcomes. The implementation is conducted on Google Colab, ensuring cloud-based execution for improved scalability and accessibility.

- Received Date: 09 Jan 2025
- Accepted Date: 21 Feb 2025
- Publication Date: 12 Mar 2025

Keywords

Medical Document Processing, Optical Character Recognition (ocr), Natural Language Processing (NLP), Paddle OCR, GPT API, Health Informatics, Patient-Centered Healthcare, Diagnostic Report Summarization

Introduction

Medical reports, including CT scan and MRI findings, play a crucial role in diagnosing health conditions. However, these reports often contain complex medical terms, making it challenging for patients to fully understand their health status. With advancements in Artificial Intelligence (AI), technologies like Optical Character Recognition (OCR) and Natural Language Processing (NLP) provide a promising solution to bridge this communication gap.

Traditionally, interpreting medical reports requires expertise, which may not always be readily available to patients. While some automated systems exist, they either struggle with accurate text extraction or fail to generate simplified summaries suitable for non-medical users. This research aims to overcome these challenges by developing an AI-driven system capable of efficiently extracting and summarizing medical reports in a way that enhances patient comprehension.

The objective of this study is to develop an AI-powered system that extracts and summarizes medical documents to enhance patient understanding. This system is designed to evaluate and compare different OCR models, such as PaddleOCR and Tesseract, to identify the most effective text extraction method for medical reports. Additionally, it incorporates advanced filtering techniques to improve text accuracy and readability while leveraging AI-based language models, including GPT, Google Pro, and LLaMA, to generate concise, patient-friendly summaries of medical findings. To ensure scalability and accessibility, the system is deployed in a cloud-based environment using Google Colab, facilitating better healthcare communication and informed decision-making.

The methodology follows a structured workflow integrating Optical Character Recognition (OCR) and Natural Language Processing (NLP) with advanced Large Language Models (LLMs) to extract, process, and summarize medical reports. It consists of four key stages: User Input, Preprocessing, Processing, and Post-processing. The process begins with the User Input stage, where users upload medical reports by either capturing an image using a camera or directly uploading a file. Since these reports often contain a mix of textual data and scanned medical documents, specialized extraction techniques are required to accurately retrieve relevant information for further processing.

In the Pre-processing stage, the uploaded image undergoes a series of enhancements to optimize OCR accuracy. The system applies cropping and resizing to eliminate unnecessary white spaces and standardize image dimensions. Converting the image to grayscale reduces its complexity, improving OCR performance, while noise reduction techniques help remove distortions, ensuring better text readability. To extract text efficiently from the processed image, OCR models such as PaddleOCR and Tesseract are evaluated and implemented.

Citation: Kiran Kumar M, Maruthi G, Sri Vardhan Reddy V, Praveen G, Aravind M. An Optimized Framework For Brain Tumor Detection And Classification Using Deep Learning And Texture Analysis. GJEIIR. 2025;5(1):25.

The Processing stage utilizes AI-driven techniques to extract meaningful insights from medical reports. If medical images are included, feature extraction is performed to identify key patterns and potential abnormalities. The extracted text then undergoes NLP-based structuring, which helps recognize essential medical terms, interpret context, and enhance coherence. Additionally, machine learning models trained on medical data classify the report based on factors such as disease type, severity, or recommended treatments, making the information more structured and accessible.

The Post-processing stage ensures that the extracted and processed information is both accurate and easy to comprehend. Advanced Large Language Models (LLMs), such as Gemini Pro, ChatGPT-4.0, and LLaMA, cross-check the medical findings against established knowledge bases to enhance credibility. The system then generates a concise, patient-friendly summary in 2-3 lines, simplifying complex medical terminology for better accessibility. This final output presents key medical insights in a clear and understandable format for the user.

By utilizing cloud-based execution through Google Colab, the system ensures scalability, computational efficiency, and accessibility. This end-to-end approach not only streamlines the processing of medical documents but also improves patient comprehension, ultimately supporting better healthcare decision-making.

Literature Survey

Research on medical text summarization and simplification using Natural Language Processing (NLP) has gained significant attention due to the complexity of medical terminology and the need for improved patient comprehension. Various studies have employed Optical Character Recognition (OCR) to extract text from medical reports, enabling further processing for analysis. Summarization techniques, including extractive methods like BERT-based models and abstractive approaches using Seq2Seq architectures, have been effective in generating concise summaries while retaining key medical insights[4]. Additionally, medical text simplification has been explored through rule-based and transformer-based models to make complex content more accessible to non-experts. Recent advancements integrate OCR, summarization, and simplification into automated systems, significantly enhancing the readability of medical documents. This research aligns with the increasing demand for AI-driven healthcare solutions that bridge the gap between medical professionals and patients, ultimately improving the accessibility of medical information.

Several studies have investigated the application of Natural Language Processing (NLP) techniques for extracting, summarizing, and simplifying medical texts to enhance patient understanding. Optical Character Recognition (OCR) is frequently utilized to extract text from medical reports, allowing further processing through summarization techniques. Both extractive approaches, such as BERT-based models, and abstractive methods, including Seq2Seq architectures, have been employed to generate concise summaries while preserving essential medical information. Additionally, medical text simplification has been explored through rule-based and transformer-based models to make complex terminology more comprehensible. Recent advancements, such as the work of have introduced semantic relevance-based neural networks to improve the readability and accuracy of medical summaries. Moreover, frameworks like Impression GPT have demonstrated the effectiveness of large language models in generating structured medical impressions without requiring additional fine-tuning[5]. The integration of these techniques into a unified system enhances the accessibility of medical documents, fostering better communication between healthcare providers and patients.

The different phases of Optical Character Recognition (OCR) have been analyzed, with a particular focus on the challenges associated with recognizing handwritten text. While OCR technology enables machines to extract information from both printed and handwritten documents, handwritten text poses difficulties due to variations in individual writing styles. Key OCR stages include preprocessing, feature extraction, recognition, and post-processing, each playing a crucial role in ensuring accuracy. Preprocessing is especially important as it helps reduce noise, enhance contrast, and separate text from images, improving recognition efficiency[6]. Several factors, such as scan quality, document type, and linguistic complexity, influence OCR performance. Additionally, recognizing scripts with intricate character structures, such as Chinese and Arabic, presents further challenges. However, with ongoing advancements in machine learning, OCR technology continues to improve in accuracy and expand its applications across various domains.

Various techniques for extracting text from images using the Tesseract-OCR engine have been explored, with a focus on overcoming challenges related to scanned documents. Factors such as font variations, text orientation, and background complexity often impact recognition accuracy[7]. To address these issues, preprocessing methods like noise reduction, binarization, and contrast enhancement are essential. Additionally, segmentation techniques, including text line separation and page frame detection, help optimize text extraction efficiency. Research has demonstrated that integrating Tesseract-OCR with Python-based image processing tools can significantly enhance recognition performance. Studies have also reviewed previous OCR advancements, including methods like connected component analysis and texture-based segmentation. Furthermore, AI and machine learning continue to refine OCR accuracy, particularly for complex scripts and handwritten text. Future research is expected to focus on multilingual recognition and real-time OCR applications, expanding its practical use in document digitization, language translation, and accessibility tools.

The process of extracting text from images using the Tesseract OCR engine plays a crucial role in converting image-based text into machine-readable formats. However, challenges such as variations in font styles, text orientation, and background complexity can affect recognition accuracy. To address these issues, preprocessing techniques like noise reduction and contrast enhancement are applied to improve text clarity. Additionally, segmentation techniques help isolate text regions for more precise recognition[8]. Research has also explored integrating Tesseract OCR with modern web technologies, utilizing React JS for the front-end and Flask for the back-end to develop an efficient text extraction system. Furthermore, OCR technology has proven valuable across industries such as banking, healthcare, and legal documentation, where automated text extraction enhances accessibility and operational efficiency. Looking ahead, advancements in AI and deep learning are expected to further improve OCR accuracy, particularly for recognizing complex scripts and multilingual text.

Extracting text from images is an essential technique used

across various industries for document digitization, data retrieval, and automated processing. This process converts image-based text into a machine-readable format, making it easier to store, search, and analyze[9]. However, challenges such as variations in text size, orientation, alignment issues, and noisy backgrounds can complicate accurate recognition. Optical Character Recognition (OCR) technology, particularly Google's Tesseract OCR engine, has greatly improved text extraction accuracy. As an open-source tool, Tesseract supports multiple languages and scripts, leveraging machine learning to enhance recognition over time. It is designed to handle distortions like rotation and skewing, making it suitable for applications such as document scanning, indexing, and archiving. With continuous advancements in OCR, text extraction methods are becoming more reliable, enabling businesses and organizations to efficiently manage and utilize textual data from images.

Proposed System

The proposed system is designed to transform complex medical reports into concise, easy-to-understand summaries, improving patient comprehension and aiding clinical decision-making. The architecture follows a structured workflow consisting of three key stages: pre-processing, processing, and post-processing, each contributing to the efficient extraction and summarization of medical information.

Pre-processing

Pre-processing enhances the quality of medical reports before analysis, ensuring better accuracy in text extraction. Users upload scanned images or PDFs of medical documents, which then undergo various enhancements. Techniques such as cropping, resizing, binarization, noise reduction, and grayscale conversion are applied to improve image clarity, remove artifacts, and correct skewed text. These refinements enhance the performance of the Optical Character Recognition (OCR) system, ensuring accurate extraction of medical text.

Processing

The processing stage involves two key components: OCR and Natural Language Processing (NLP) to extract, interpret, and structure the medical data.

- **OCR System:** The OCR module detects and extracts text from the pre-processed medical reports, segmenting it into lines and words. Machine learning models, such as Convolutional Neural Networks (CNNs), are used for character recognition, improving accuracy. Additionally, a post-processing step corrects misrecognized words using contextual analysis, ensuring more precise results. The extracted text is then passed to the NLP module for further refinement.
- **NLP System:** The NLP module enhances the extracted text through entity recognition, semantic analysis, and classification techniques. It identifies critical medical elements such as symptoms, diagnoses, medications, and test results. By applying deep learning models, it structures this information and establishes meaningful relationships between medical terms, preparing the data for the summarization process.

Post-processing and Summarization

The post-processing stage focuses on refining and validating the extracted medical information to generate a clear, concise, and accurate summary. The structured text from the NLP module is processed through the Summarization Module, which employs both extractive and abstractive summarization techniques:

- **Extractive Summarization:** Identifies and retains the most relevant sentences using techniques like TF-IDF and TextRank, ensuring key medical insights are preserved.
- **Abstractive Summarization:** Utilizes deep learning models such as BERT and GPT to generate natural, human-like summaries that encapsulate the essential meaning of the medical report.

To enhance reliability, the generated summary is cross-verified against a medical knowledge base. Any inconsistencies or errors are flagged for correction, ensuring accuracy and credibility. The final output is a concise 2-3 line summary highlighting the most critical findings of the medical report.

Output and User Feedback Mechanism

The final medical summary is presented to the user in an easily

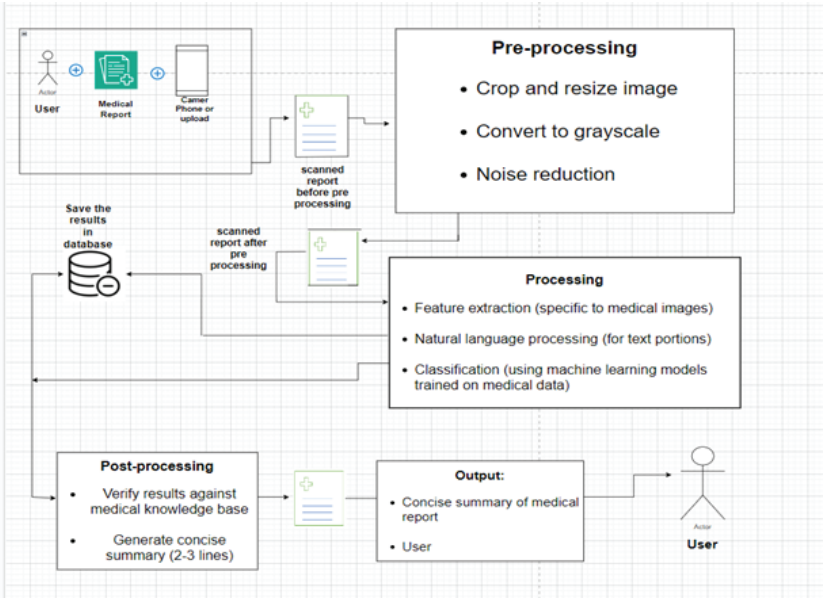


***Figure 1**. Methodology of medical report*

digestible format, ensuring that the extracted insights are clear and actionable. A user feedback mechanism allows healthcare professionals to review the summary and suggest refinements if necessary. This iterative process helps the system improve continuously by learning from expert feedback, enhancing its precision over time.

System Scalability and Integration

The proposed system is designed for scalability and seamless integration with Electronic Health Record (EHR) systems, hospital databases, and cloud-based medical platforms. This adaptability ensures that medical professionals can efficiently access structured and concise patient data, streamlining clinical workflows.

By combining OCR for text extraction, NLP for semantic analysis, and AI-powered summarization, this system effectively converts complex medical documents into clear, concise, and actionable insights. This structured approach not only enhances accessibility but also improves decision-making in healthcare, benefiting both professionals and patients.

Discussion

In this study, we evaluated three Optical Character Recognition (OCR) tools—Tesseract OCR, EasyOCR, and PaddleOCR—to assess their effectiveness in extracting text from MRI, CT, and X-ray scans. Each OCR tool was analyzed based on accuracy, ease of use, and performance in handling medical images. The extracted text was then summarized into concise 2-3 line outputs using GPT-4o, BART, and Gemini Pro, making medical reports more accessible and understandable.

Comparison of OCR Models

Tesseract OCR

Tesseract OCR is a widely used open-source tool for text extraction from structured documents. While it performs well with printed and clear text, it struggles with noisy, distorted, or unstructured medical scans. Although preprocessing techniques like binarization and noise reduction improve accuracy, its performance remains limited for complex medical images. Additionally, Tesseract is relatively slow compared to deep-learning-based OCR models and requires significant manual tuning to be effective in medical applications.

EasyOCR

EasyOCR offers better accuracy and flexibility than Tesseract, supporting over 80 languages and handling both printed and handwritten text. It is user-friendly and requires minimal preprocessing, making it easier to integrate into various applications. However, while EasyOCR performs better than Tesseract, it still falls short in extracting detailed medical information from highly structured images such as X-rays, MRIs, and CT scans. Its general-purpose design limits its effectiveness for complex medical documents.

PaddleOCR

PaddleOCR is specifically optimized for handling structured and complex data, making it the best choice for medical imaging applications. It supports over 100 languages and delivers state-of-the-art accuracy, even in challenging environments like noisy medical scans. Unlike Tesseract and EasyOCR, PaddleOCR effectively extracts critical medical text, adapting well to different orientations and embedded text within scans. While it requires a more complex setup, its superior accuracy and adaptability make it the preferred OCR tool for medical document processing.

Feature	Tesseract OCR	Easy OCR	Paddle Ocr
Accuracy	Good for simple, clean text	High accuracy for multi-lingual text	High accuracy for structured data
Ease of Use	Requires preprocessing for best results	Easy to integrate and use out-of-box	Slightly more complex setup
Speed	Moderate	Fast	Fast
Support for Languages	Extensive	Over 80 languages	Over 100 languages
Output Reliability	Struggles with handwritten or noisy images	Handles handwritten and complex images better	Performs well with both clean and complex inputs
Use Case Suitability	Basic text extraction	Medical imaging and structured data	Medical imaging and structured data
Community Support	Strong	Moderate	Growing

Figure 2. *Comparison between the OCR models*

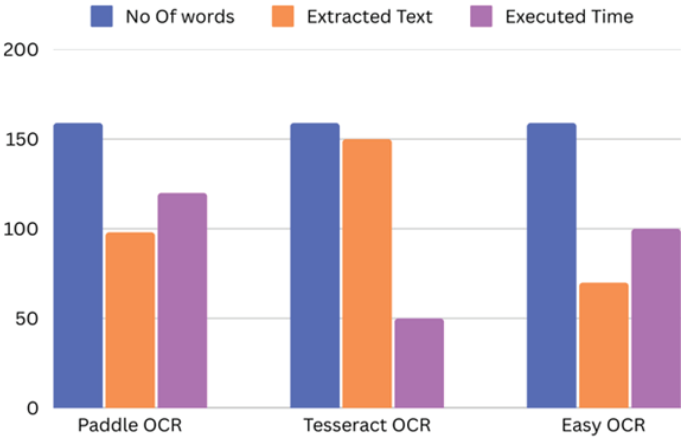


Figure 3. *Graphs of OCR models*

Summarization

Summarization is the process of condensing large amounts of text while retaining essential information. In AI, summarization techniques are widely used to improve efficiency in medical reports.

GPT-4o

GPT-4o, an advanced AI model developed by OpenAI, excels in processing and generating human-like content across multiple modalities, including text, images, and audio. In this project, GPT-4o plays a crucial role in analyzing diverse inputs such as medical images and patient reports. By leveraging its multimodal capabilities, it generates comprehensive and insightful summaries that go beyond traditional text-based analysis. This allows doctors to quickly grasp essential information, saving time and enabling them to focus on critical aspects of patient care.

Gemini Pro

Gemini Pro, developed by Google AI, is a multimodal AI model with a unique ability to process both text and images. This capability makes it highly effective in medical image analysis. By integrating extracted text data with the original medical images, Gemini Pro gains a deeper understanding of medical findings. This results in more comprehensive and contextually rich summaries, enhancing diagnostic accuracy. By considering both textual information and visual cues, Gemini Pro provides

Feature	GPT-4	BART	Gemini Pro
Accuracy of summarization	High; captures context well	Good; excels in abstractive summarization	Excellent; designed for nuanced tasks
Ease of Integration	Simple, with APIs readily available	Requires moderate setup	Simple; designed for human-centric tasks
Language Proficiency	Excellent, supports multiple languages	Good, with focus on multi-tasking	Excellent, particularly for technical summaries
Human-Readable Format	Outstanding; prioritizes clarity	Very good; concise and coherent outputs	Outstanding; human-friendly tone
Performance in Medical Context	Reliable for general medical content	Moderate; needs domain adaptation	Excellent; specialized for complex medical texts
Speed of Response	Moderate	Fast	Fast
Use Case Suitability	General-purpose AI	Abstractive summarization and analysis	Summarizing and explaining medical content

***Figure 4**. Comparison of Summarization models*

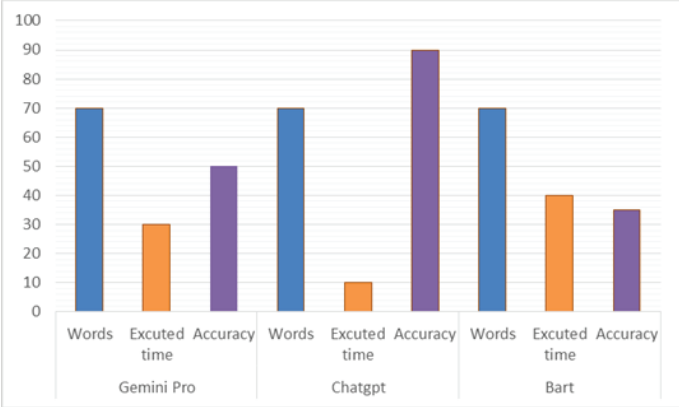


***Figure 5**. Bar Graph of Summarization Models*

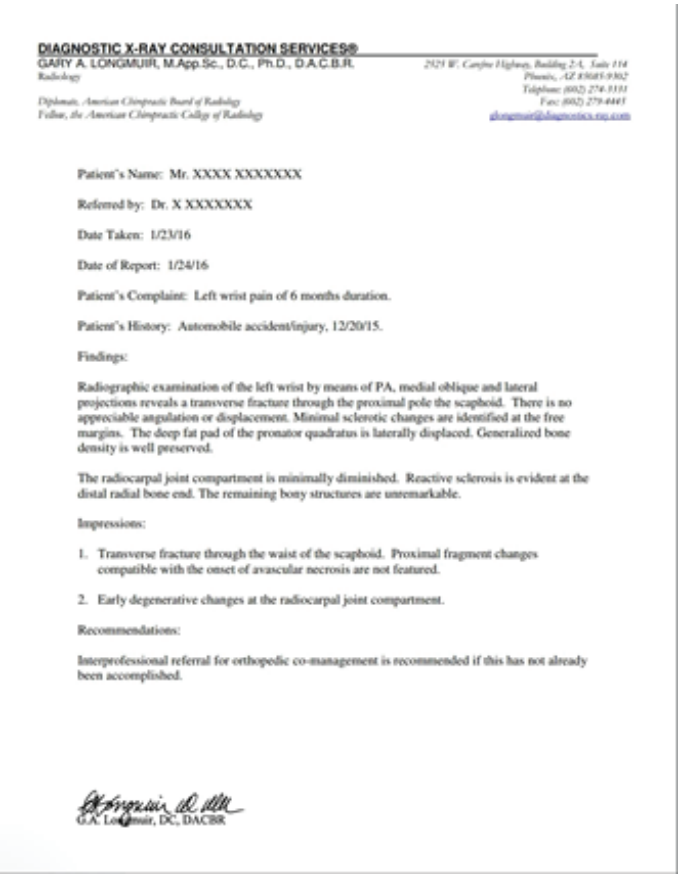
a more holistic interpretation of medical reports, improving clinical decision-making.

BART

BART, a transformer-based language model, specializes in text comprehension and summarization. In medical applications, BART is particularly useful for condensing lengthy medical reports into concise and informative summaries. Its ability to generate precise medical terminology ensures clarity and accuracy in documentation. By automating report summarization, BART reduces manual review time, minimizes errors, and enhances communication among healthcare providers. This contributes to improved decision-making and better patient outcomes.

Results

The proposed system streamlines the processing of medical reports by integrating Optical Character Recognition (OCR), Natural Language Processing (NLP), and summarization techniques. The OCR module efficiently extracts text from scanned reports, delivering high accuracy after applying preprocessing techniques such as noise reduction and grayscale conversion. Once extracted, the text is analyzed by the NLP module, which identifies key medical details, including symptoms, diagnoses, and prescribed medications, ensuring the information is well-structured and meaningful.To simplify lengthy medical reports, the summarization module condenses the extracted information while preserving crucial details. This reduces complexity, making



DIAGNOSTIC X-RAY CONSULTATION SERVICES®

GARY A. LONGMUIR, M.App.Sc., D.C., Ph.D., D.A.C.B.R.
Radiology

Diplomate, American Chiropractic Board of Radiology
Fellow, the American Chiropractic College of Radiology

2929 W. Camelback Highway, Building 2-A, Suite 114
Phoenix, AZ 85085-9302
Telephone: (602) 274-3333
Fax: (602) 279-4445
glongmuir@diagnostics-xray.com

Patient's Name: Mr. XXXX XXXXXXX

Referred by: Dr. X XXXXXXX

Date Taken: 1/23/16

Date of Report: 1/24/16

Patient's Complaint: Left wrist pain of 6 months duration.

Patient's History: Automobile accident/injury, 12/20/15.

Findings:

Radiographic examination of the left wrist by means of PA, medial oblique and lateral projections reveals a transverse fracture through the proximal pole the scaphoid. There is no appreciable angulation or displacement. Minimal sclerotic changes are identified at the free margins. The deep fat pad of the pronator quadratus is laterally displaced. Generalized bone density is well preserved.

The radiocarpal joint compartment is minimally diminished. Reactive sclerosis is evident at the distal radial bone end. The remaining bony structures are unremarkable.

Impressions:

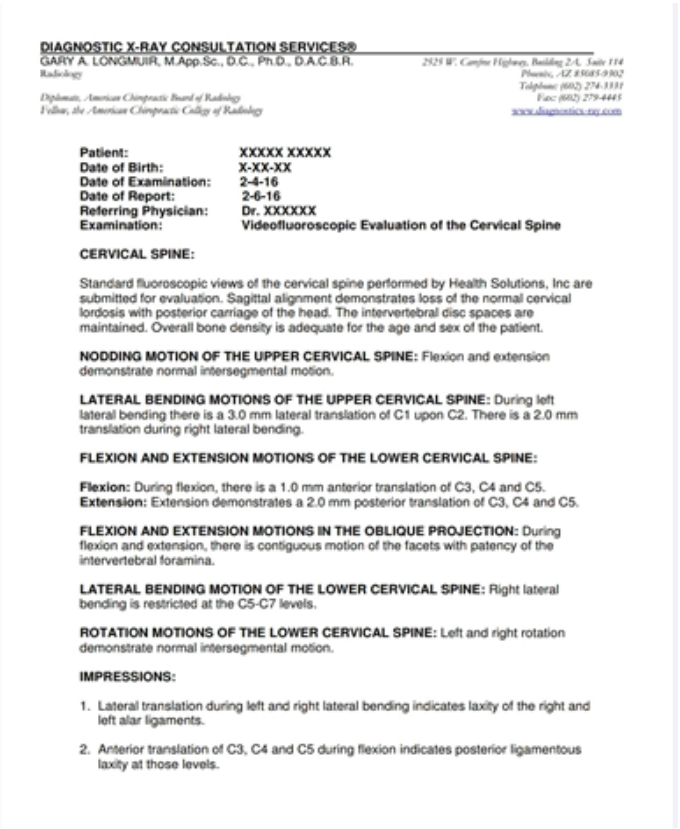
1. Transverse fracture through the waist of the scaphoid. Proximal fragment changes compatible with the onset of avascular necrosis are not featured.
2. Early degenerative changes at the radiocarpal joint compartment.

Recommendations:

Interprofessional referral for orthopedic co-management is recommended if this has not already been accomplished.

G.A. Longmuir, DC, DACBR

***Figure 6**. Input Image*



DIAGNOSTIC X-RAY CONSULTATION SERVICES®

GARY A. LONGMUIR, M.App.Sc., D.C., Ph.D., D.A.C.B.R.
Radiology

Diplomate, American Chiropractic Board of Radiology
Fellow, the American Chiropractic College of Radiology

2929 W. Camelback Highway, Building 2-A, Suite 114
Phoenix, AZ 85085-9302
Telephone: (602) 274-3333
Fax: (602) 279-4445
www.diagnostics-xray.com

Patient: XXXXX XXXXX
Date of Birth: X-XX-XX
Date of Examination: 2-4-16
Date of Report: 2-6-16
Referring Physician: Dr. XXXXXX
Examination: Videofluoroscopic Evaluation of the Cervical Spine

CERVICAL SPINE:

Standard fluoroscopic views of the cervical spine performed by Health Solutions, Inc are submitted for evaluation. Sagittal alignment demonstrates loss of the normal cervical lordosis with posterior carriage of the head. The intervertebral disc spaces are maintained. Overall bone density is adequate for the age and sex of the patient.

NODDING MOTION OF THE UPPER CERVICAL SPINE: Flexion and extension demonstrate normal intersegmental motion.

LATERAL BENDING MOTIONS OF THE UPPER CERVICAL SPINE: During left lateral bending there is a 3.0 mm lateral translation of C1 upon C2. There is a 2.0 mm translation during right lateral bending.

FLEXION AND EXTENSION MOTIONS OF THE LOWER CERVICAL SPINE:

Flexion: During flexion, there is a 1.0 mm anterior translation of C3, C4 and C5.
Extension: Extension demonstrates a 2.0 mm posterior translation of C3, C4 and C5.

FLEXION AND EXTENSION MOTIONS IN THE OBLIQUE PROJECTION: During flexion and extension, there is contiguous motion of the facets with patency of the intervertebral foramina.

LATERAL BENDING MOTION OF THE LOWER CERVICAL SPINE: Right lateral bending is restricted at the C5-C7 levels.

ROTATION MOTIONS OF THE LOWER CERVICAL SPINE: Left and right rotation demonstrate normal intersegmental motion.

IMPRESSIONS:

1. Lateral translation during left and right lateral bending indicates laxity of the right and left alar ligaments.
2. Anterior translation of C3, C4 and C5 during flexion indicates posterior ligamentous laxity at those levels.

***Figure 7**. Input Image*

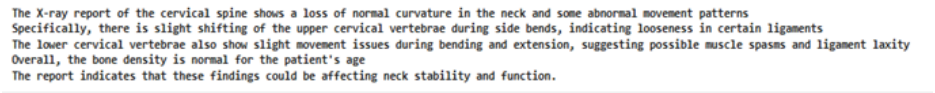


***Figure 8**. Output - Summarization of Input image*

reports easier to understand without losing essential insights.

Performance evaluations indicate that the system generates precise and reliable summaries, significantly minimizing the need for manual interpretation. Secure database storage ensures easy retrieval of reports, while the user feedback mechanism allows healthcare professionals to refine and improve summaries, enhancing the system's reliability over time.

Conclusion

Integrating AI-driven summarization into medical report analysis has the potential to transform healthcare by enhancing efficiency, accuracy, and accessibility. This project utilizes advanced models such as GPT-4o, Gemini Pro, and BART to simplify complex medical data, generating concise and meaningful summaries for healthcare professionals. These AI models go beyond text processing by incorporating insights from medical images, leading to a more holistic understanding of patient reports.

By automating the summarization process, this approach significantly reduces the time required for manual review, minimizes errors, and supports better clinical decision-making. Additionally, it bridges the communication gap between doctors and patients by translating complex medical terminology into easily understandable language. As AI technology continues to evolve, its role in medical data processing will expand, streamlining workflows, improving patient outcomes, and allowing healthcare professionals to dedicate more time to patient care rather than administrative tasks.

References

1. N. Kanwal and G. Rizzo, "Attention-based clinical note summarization," Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing (SAC '22), pp. 813–820, 2022.
2. M. Nishio, T. Matsunaga, H. Matsuo, M. Nogami, Y. Kurata, K. Fujimoto, O. Sugiyama, T. Akashi, S. Aoki, and T. Murakami, "Fully automatic summarization of radiology reports using natural language processing with large language models," Intelligent Medicine, vol. 6, 2024.
3. C. Nitsch and U. Kebschull, "Summarization and Simplification of Medical Articles using Natural Language Processing," Proceedings of IEEE Conference, 2024.
4. A. Patel, S. P. Pujari, and P. K. Atrey, "Summarization and Simplification of Medical Articles using Natural Language Processing," in Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2023.
5. C. Ma, "ImpressionGPT: A Large Language Model Enhanced Framework for Radiology Report Impression Generation," In arXiv preprint, vol. 2304, no. 08448, 2023
6. K. Karthick, K. B. Ravindrakumar, R. Francis, and S. Ilankannan, "Steps Involved in Text Recognition and Recent Research in OCR; A Study," International Journal of Recent Technology and Engineering (IJRTE), vol. 8, no. 1, pp. 3095–3100, May 2019
7. S. K. Garai, S. Ghoshal, O. Paul, N. Biswas, U. Dey, and S. Mondal, "A Novel Method for Image to Text Extraction Using Tesseract-OCR," American Journal of Electronics & Communication, vol. III, no. 2, pp. 8-11, 2022.
8. Shinde, A., Singh, P., Patil, J., Singh, J., & Baraskar, T. (2021). Text Extraction from Images using Tesseract. International Research Journal of Engineering and Technology (IRJET), 08(07), 295-301.
9. Kumar, S., Sharma, N. K., Sharma, M., & Agrawal, N. (2024). Text Extraction from Images Using Tesseract. In P. S. Rathore, S. Ahuja, S. R. Burri, A. Khunteta, A. Baliyan, & A. Kumar (Eds.), Advances in Computational Techniques for Image Processing.