JOURNAL OF DYNAMICS AND CONTROL

VOLUME 8 ISSUE 11

A MULTI-MODAL AI SYSTEM FOR AUTOMATED DETECTION OF PNEUMONIA, SKIN DISEASES, AND BONE FRACTURES

Aakash Walavalkar[1], Swati Vaishnav[2], Jay Ajmera[2], Kallind Soni[2], Atharva Rode[2], Andrila Sarkar[3]

[1]Triumph AI, Mumbai, India
[2]SVKM's NMIMS - MPSTME, Mumbai, India
[3]University of Engineering and Management, Jaipur, India

A MULTI-MODAL AI SYSTEM FOR AUTOMATED DETECTION OF PNEUMONIA, SKIN DISEASES, AND BONE FRACTURES

Aakash Walavalkar[1], Swati Vaishnav[2], Jay Ajmera[2], Kallind Soni[2], Atharva Rode[2], Andrila Sarkar[3]

[1]*Triumph AI, Mumbai, India*
[2]*SVKM's NMIMS - MPSTME, Mumbai, India*
[3]*University of Engineering and Management, Jaipur, India*
aakash.muskurahat@gmail.com, swati.vaishnav@nmims.edu, ajmerajay.ja@gmail.com, kallindsoni@gmail.com, rodeatharva3@gmail.com,sarkarmimi9051@gmail.com

Abstract: *This paper presents a multi-modal AI-driven solution designed to aid healthcare professionals in diagnosing pneumonia, skin diseases, and bone fractures. Leveraging deep learning techniques, the system utilizes three distinct YOLOv8 models for image segmentation, classification, and object detection. The integration of large language models (LLMs) such as Google Gemini enhances diagnostic accuracy. The system includes two workflows: new diagnosis and second opinion, allowing patients and medical ex- perts to upload images and obtain refined results. Combining advanced AI, secure cloud storage, and an intuitive front end, the solution significantly improves diagnostic speed and accuracy.*

1 Introduction

The global prevalence of conditions like pneumonia, skin diseases, and bone fractures presents a substantial challenge for healthcare systems. These diseases often require timely and precise diagnosis to ensure optimal treatment and patient outcomes. However, traditional diagnostic methods are frequently slow and reliant on the availability of specialized medical professionals. This limitation is particularly pronounced in high-demand or resource-limited healthcare environments, where delays can have severe consequences. In this paper, we propose a novel, automated, multi-modal AI system that leverages cutting-edge deep learning models and large language models (LLMs) to assist in the diagnosis of these conditions. The system utilizes YOLOv8 models for image processing, with Neo4j as a graph database to enhance the diagnostic capabilities through semantic search and LLM grounding.

2 Literature Review

Recent advances in artificial intelligence (AI), particularly in computer vision, have revolutionized diagnostic tools in healthcare. Object detection models such as YOLO (You Only Look Once) have been effectively applied in medical imaging for the detection of abnormalities. CNN-based systems have demonstrated high accuracy in identifying conditions like pneumonia and fractures. However, despite these advancements, current models are limited by their lack of interpretability and integration with larger knowledge bases. Graph databases like Neo4j and large language models (LLMs) can provide additional layers of context and relational data that are crucial in refining diagnostic results. This paper builds on the foundational work in object detection, combining it with graph-based semantics to provide a more robust diagnostic tool.

3 Methodology

This section outlines the detailed methodology used in the system, describing the user journey, model interactions, and result processing steps. The system's components include model training, workflow design, and Neo4j integration to enhance the reliability of results.

3.1 Workflow Design

The system offers two workflows, designed to address different diagnostic needs:

- **New Diagnosis**: This workflow is intended for patients seeking an initial diagnosis. Users create an account and upload relevant images, which are then analyzed by the system.
- **Second Opinion**: Patients or healthcare professionals seeking a secondary diagnosis can use this workflow. It requests prior medical history, allowing the AI model to incorporate contextual information.

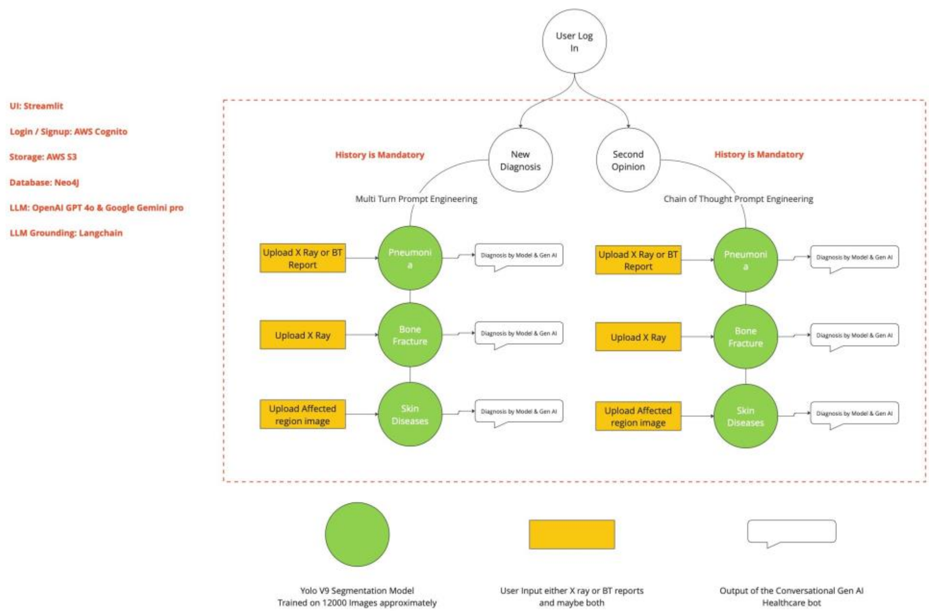


Figure 1: Workflow Diagram for Diagnosing Pneumonia, Skin Diseases, and Bone Fractures.

3.2 User Journey and Interaction Flow

The user journey is structured to allow easy access to both workflows:

1. **Account Creation and Login**: Users create an account using AWS Cognito for secure authentication. After logging in, they can choose between "New Diagnosis" and "Second Opinion."
2. **Workflow Selection and Image Upload**: Users select the condition (pneumonia, skin disease, or bone fracture) and upload the appropriate medical images (X-rays for pneumonia and fractures; affected skin area images for skin disease).
3. **Model Processing and Prediction**: Uploaded images are stored on Amazon S3 and processed by the respective YOLOv8 models:
    - **Pneumonia Detection**: The YOLOv8 segmentation model detects pneumonia severity.
    - **Skin Diseases Classification**: The model classifies the skin condition.

- **Bone Fracture Detection**: The model identifies fracture locations and severity.

4. **Refinement with LLM and Neo4j Grounding**: The results are further refined using Google Gemini's LLM. For bone fracture cases, the LLM's responses are grounded through Neo4j, which enhances accuracy by referencing relational data.
5. **Result Display and Follow-up Interaction**: Refined results are presented to the user via a chat interface. Users can ask follow-up questions, and the system provides additional insights based on prior interactions.

3.3 Model Training

The AI models used for diagnosis were trained on extensive datasets to ensure accuracy and reliability:

- **Pneumonia Detection**: Utilizes a segmentation model based on YOLOv8 to detect signs of pneumonia in chest X-rays.
- **Skin Diseases Classification**: A classification model trained to identify multiple skin conditions with a dataset of over 15,000 images.
- **Bone Fracture Detection**: An object detection model capable of identifying bone fractures, further enhanced with semantic data from Neo4j.

3.4 Neo4j Integration

The Neo4j Graph Database plays a crucial role in grounding the LLM's output. For bone fracture diagnosis, Neo4j organizes data into nodes and relationships that represent factual information. Using semantic search through cosine similarity, the LLM can retrieve refined answers by referencing relational data in Neo4j, which improves the accuracy of bone fracture diagnostics.

4 Implementation

The system implementation involves the integration of frontend and backend components to create a seamless diagnostic experience for users.

4.1 Frontend Interface

The frontend is developed using Streamlit, a Python-based framework that facilitates rapid prototyping of data applications. This user interface allows patients and healthcare providers to upload images, interact with diagnostic results, and access the chatbot for further queries.

4.2 Authentication and Data Security

User authentication is managed through AWS Cognito, ensuring a secure sign-in process. Uploaded images and user data are stored in Amazon S3, which provides scalable storage with robust data security protocols.

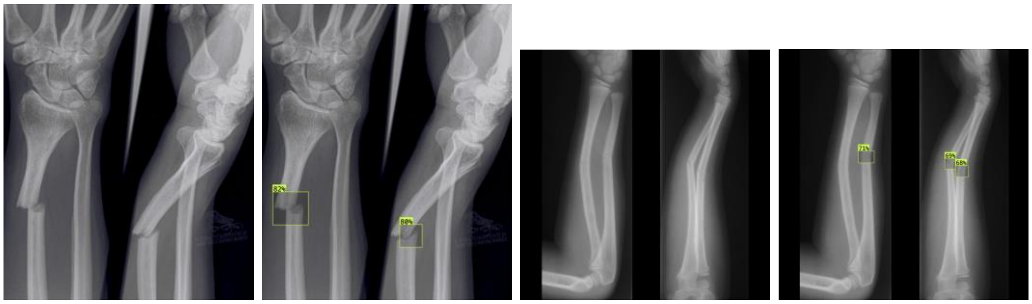
4.3 LLM Integration and Grounding

Google Gemini's LLM is used to process and refine the diagnostic output. The LLM's responses are grounded through Neo4j, especially for bone fracture cases, ensuring the diagnostic output is accurate and supported by relational data.

5 Case Studies

This section demonstrates the effectiveness of our system in detecting fractures, skin diseases, and pneumonia.

5.1 Fracture Detection



(a) Test Image 1 (b) Prediction 1 (c) Test Image 2 (d) Prediction 2

Figure 2: Fracture Detection with YOLOv8 and Neo4j Grounding

5.2 Skin Disease Detection

(a) Skin Disease 1 (b) Prediction 1 (c) Skin Disease 2 (d) Prediction 2

Figure 3: Skin Disease Detection with YOLOv8

5.3 Pneumonia Detection

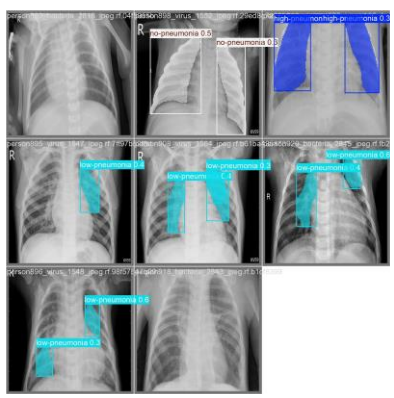


Figure 4: Pneumonia Detection using Segmentation Model

6 Conclusion

This paper presents an AI-driven multi-modal system for the automated detection of pneumonia, skin diseases, and bone fractures. The combination of YOLOv8 models, large language models, and Neo4j grounding provides reliable diagnostics and enhances the system's interpretability and accuracy. Future work will focus on expanding the model's diagnostic capabilities and refining the user interface for a smoother experience.

7 References

1. A. M. Rashid et al., ”Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning,” *Measurement*, vol. 152, 2020.

2. M. K. Jain and P. Jain, ”Bone Fracture Detection and Classification using Deep Learning Approach,” in *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, 2020.