



Research Article

## AI-Powered Bone Fracture Detection and Advisory System: A Deep Learning Method Using GPT-Based Health Advice

Hafiz Muhammad Usman Zahoor<sup>1</sup> , Furqan Ahmad<sup>2\*</sup> , Hafsa Ahmad<sup>3</sup> , Fiaz Hussain<sup>4</sup> ,  
Jawad Jameel<sup>5</sup>  and Muhammad Owais Ishaq<sup>6</sup> 

<sup>1,4,5,6</sup>Department of Software Engineering, The Islamia University of Bahawalpur, Bahawalpur, Pakistan

<sup>2\*</sup>State Key Laboratory of Multiphase Flow in Power Engineering, School of Energy and Power Engineering, Xi'an Jiaotong University, China

<sup>2\*</sup>Department of Mechanical Engineering, University of Engineering and Technology Lahore (Narowal Campus), Pakistan

<sup>3</sup>Department of Nursing, University of Health and Sciences Lahore, Lahore, Pakistan

### Article Information

### Abstract

#### Article History

Received: 9 February 2026

Revised: 3 March 2026

Accepted: 21 March 2026

Published online: 1 April 2026

#### Keywords

Bone Fractures

Healthcare

CNN

Orthopedic Diagnosis

Fracture Detection

Deep Learning

#### Correspondence

[hafizfurqanahmad538@gmail.com](mailto:hafizfurqanahmad538@gmail.com)

#### ORCID


Hafiz Muhammad Usman Zahoor   
<https://orcid.org/0009-0007-0582-1173>

Furqan Ahmad   
<https://orcid.org/0009-0003-3148-0887>

Hafsa Ahmad   
<https://orcid.org/0009-0001-6901-5345>

Fiaz Hussain   
<https://orcid.org/0009-0008-4902-4245>

Jawad Jameel   
<https://orcid.org/0009-0007-5387-7876>

Muhammad Owais Ishaq   
<https://orcid.org/0009-0002-5898-2169>

One of the most common musculoskeletal conditions that poses major challenges to healthcare systems across the globe is bone fractures. Traditional diagnostic methods are often prone to variability and human errors, although timely and accurate diagnosis is essential to avoid complications. The AI-Fracture Detection and Advisory System proposed in this research combines Convolutional Neural Networks (CNNs) for deep learning-based fracture detection with a Generative Pretrained Transformer (GPT)-based advisory module for health recommendations. CNNs utilize several convolutional layers activated by ReLU for scanning X-ray images. Pooling layers and fully connected layers with a Softmax function are used for classification. With 0.92 accuracy, 0.94 sensitivity, and 0.88 specificity, the model achieves 90% classification accuracy. The GPT-based advisory component provides recommendations based on fracture location, severity, and patient-specific information. It is built on a React.js-based front-end real-time interface, while image uploading and processing are handled by a scalable Flask/FastAPI backend. The system is trained on a balanced dataset of 10,580 X-ray images using preprocessing techniques such as pixel normalization, random rotation, flipping, brightness adjustment, and resizing to  $180 \times 180$  pixels. By suggesting appropriate treatment options, such as surgery or immobilization based on fracture type, the system offers computer-aided decision support. Existing and future research aims to incorporate 3D imaging modalities and adopt a multiclass classification approach for detecting various types of fractures to further enhance the clinical applicability and robustness of AI-based diagnostic systems.

© 2026 Centre for Research and Innovation (CRI). This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## I. INTRODUCTION

Bone fractures affect millions of patients annually, making them the most common type of musculoskeletal injuries. Fractures are one of the top reasons for emergency department visits, particularly among individuals involved in high-impact sports and the elderly, according to the World Health Organization (WHO) [1]. To prevent complications

such as malunion, delayed healing, and loss of function, fractures must be identified correctly and promptly [2]. Many traditional fracture diagnoses remain dependent on manual radiologist interpretation, which is prone to fatigue due to workload and differences in skill levels [3]. This human judgment-dependent factor leads to serious issues such as misclassification, non-standard clinical decision-making, and delays in diagnosis. Misdiagnosis is more prevalent in areas

where there are fewer radiology specialists, making the use of automated diagnostic systems necessary [4].

As medical imaging has increasingly incorporated artificial intelligence (AI), deep learning models—most notably Convolutional Neural Networks (CNNs)—have achieved expert-level performance in image-based diagnosis [5]. CNN-based models have been effectively used for multiple medical imaging tasks, such as fracture classification, pneumonia diagnosis, and tumor detection [6]. These models have significantly improved fracture detection precision, with ResNet models reaching 90% precision in wrist fracture detection [7]. Despite their success, current AI-based fracture diagnosis models remain of limited value to clinicians because they often adopt a binary classification approach that distinguishes only between “fractured” and “non-fractured” cases without providing interpretability or additional clinical information [8].

The lack of decision-support systems that provide clinical information beyond image classification is a significant but often overlooked drawback in AI-assisted fracture diagnosis. Physicians are left to interpret raw AI predictions independently because current AI models are unable to provide contextual details such as the type of fracture, severity level, or suggested medical procedures [9]. While radiologists and orthopedic specialists require more than simple classification results to make informed treatment decisions, the clinical utilization of deep learning–driven diagnostic models have been limited due to the absence of a combined AI-enabled advisory system [10]. The proposed AI-supported Bone Fracture Detection and Advisory System address this critical shortfall by combining deep learning–based classification with a GPT-orchestrated medical advisory system to provide accurate detection and automated decision-making support [11].

There are two major flaws in existing deep learning–based fracture detection systems. First, because CNN-based models are black-box classifiers that output only classification labels and not severity levels or available treatment options, they are neither context-aware nor easily interpretable [12]. In complex cases where treatment varies based on fracture morphology, patient age, and comorbidities, clinicians require more than simple classification explanations. Second, no fracture detection model with an AI-based decision-support system currently exists to assist doctors in decision-making after diagnosis. An AI-led advisory system can fill this knowledge gap by providing evidence-based treatment guidelines, particularly in low-resource settings where access to trained radiologists is limited [13]. For instance, consider a rural clinic with limited radiology expertise that receives an X-ray of a suspected fracture. An AI model detects the fracture but provides no further guidance [14]. Because the clinician may have limited experience with fracture treatment algorithms, they may struggle to decide between immobilization, referral for surgery, or conservative management [15]. An integrated advisory system can support decision-making by suggesting appropriate next steps based

on AI-generated evidence, thereby optimizing clinical outcomes and reducing unnecessary referrals [16].

This article introduces a novel AI-based Bone Fracture Detection and Advisory System that combines deep learning–based image classification with an AI-driven medical advisory model to overcome these drawbacks. The system uses a GPT-style natural language processing (NLP) model to provide advisory support following fracture detection, enabling higher interpretability and clinical utility compared to standard CNN-based classifiers [17]. By providing site- and severity-aware advisory guidance, the system bridges the gap between AI-assisted diagnosis and real clinical decision-making.

A key aspect of this study is the development of an interactive real-time platform built with React.js, designed for integration into hospital operational workflows. The system is easily accessible and suitable for real-world use, allowing medical practitioners to upload an X-ray image, receive AI-driven predictions, and access advisory recommendations in real time. This framework represents a flexible and expandable AI system for broader clinical applications, as it can be extended to detect other medical conditions beyond bone fractures, such as soft tissue injuries, osteoporosis, and joint dislocations.

Although GPT-based models have been applied for medical advisory tasks and CNNs have been used for fracture diagnosis, no prior study has integrated these two approaches into a single diagnostic and advisory system. Existing models such as [18], [19], and [20] focus mainly on text classification or advisory tasks but do not provide an end-to-end AI-driven clinical decision-support system. This work bridges that gap by presenting a comprehensive framework for AI-based medical imaging and advisory support. The main contribution is an advisory system that enhances clinical usability and diagnostic precision by combining deep learning and GPT-based technologies. By incorporating contextual information such as fracture severity and recommended treatment protocols, our approach improves interpretability and reduces the workload of radiologists compared to earlier models. We also propose a real-time AI clinical deployment system using React.js, which can be easily integrated into telemedicine applications, clinics, and hospitals. Beyond fractures, the scalable architecture also provides a generalizable AI-powered diagnostic platform for detecting conditions such as osteoporosis, joint dislocations, and soft tissue injuries. By integrating deep learning with NLP-enabled advisory capabilities, this work advances the state of the art in medical imaging and holds promise for more advanced, context-aware, and clinically useful AI-powered healthcare solutions.

## II. METHODOLOGY

The Bone Fracture Detection and Advisory System is designed for fracture detection and medical consultation using a multi-layered architecture that combines a deep learning model, a backend server, and an online interface.

The overall system structure and data flow emphasize the relationship between the advisory system, image processing workflow, deep learning module, and report production unit. The process is initiated when a user uploads an X-ray image, which subsequently undergoes a preprocessing stage involving quality checks and image enhancement. The refined image is then analyzed by a detection module to identify the presence of a fracture.

Upon positive identification, the system stores the relevant data and based on the user’s request, proceeds to an advisory phase. This phase generates tailored medical advice, including treatment protocols and pharmaceutical suggestions. The pipeline culminates in the compilation of a comprehensive diagnostic report, which is ultimately presented to the user, thereby providing an end-to-end automated diagnostic and advisory solution. The flowchart is shown in Figure 1. Figure 2 provides a thorough illustration of the communication between several modules. The backend server consists of two main layers: the API layer, which controls data interaction with the web interface, and the processing layer, which oversees preprocessing and model

inference. A Convolutional Neural Network (CNN) created in TensorFlow is used to construct the fracture detection component. Convolution, ReLU activation, and pooling layers are used by the model to extract features. Fully connected layers and a Softmax classifier are then used to evaluate whether a fracture is present. The results are saved in a MongoDB database after detection is completed, ensuring secure access and future retrieval. If the system detects a fracture, the output is sent to the AI-powered advisory module, which uses a GPT-based language model to provide treatment suggestions. Before being included in the final report, these recommendations are reviewed and validated by a doctor. If no fracture is found, the system immediately notifies the user through the user interface (UI). Figure 3 illustrates user roles and system interactions, highlighting three main actors: the patient, doctor, and system administrator. Patients upload X-ray images to obtain AI-generated diagnostic reports; doctors validate the outputs and provide professional medical advice; and administrators ensure seamless operation, system security, and data accuracy.

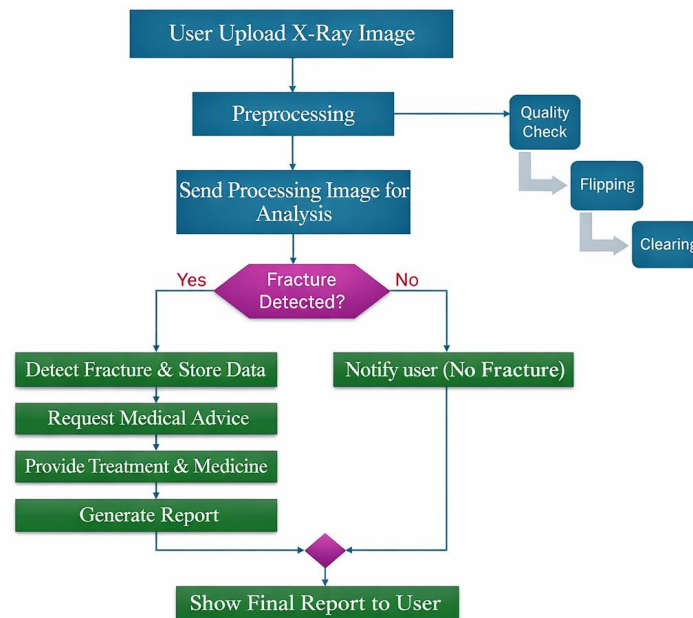


Fig.1 Flowchart of Bone Fracture Detection and Advisory System

Three main actors are depicted in Figure 3 (Use Case Diagram), which shows the fundamental user roles and how they interact with the proposed system. The patient uploads X-ray images to receive AI-generated diagnostic reports. The physician plays a crucial verification role by interpreting the system's outputs and providing professional medical advice and monitoring. By managing system functionality, adhering to security procedures, and maintaining data accuracy, the system administrator ensures operational integrity. This structured interaction framework creates a dependable and effective diagnostic ecosystem, ensuring a smooth workflow in which automated analysis is effectively supported by

expert medical knowledge and strong administrative management. The basis of this process lies in the architecture of the Convolutional Neural Network (CNN) model, which employs multiple layers for X-ray image processing. For feature extraction, the model first accepts an input X-ray image and processes it through convolution layers. The convolutional layers are subsequently followed by activation functions such as ReLU, which introduce non-linearity and facilitate feature learning. Dimensionality is then reduced by pooling layers, which enhance computational efficiency without sacrificing meaningful data.

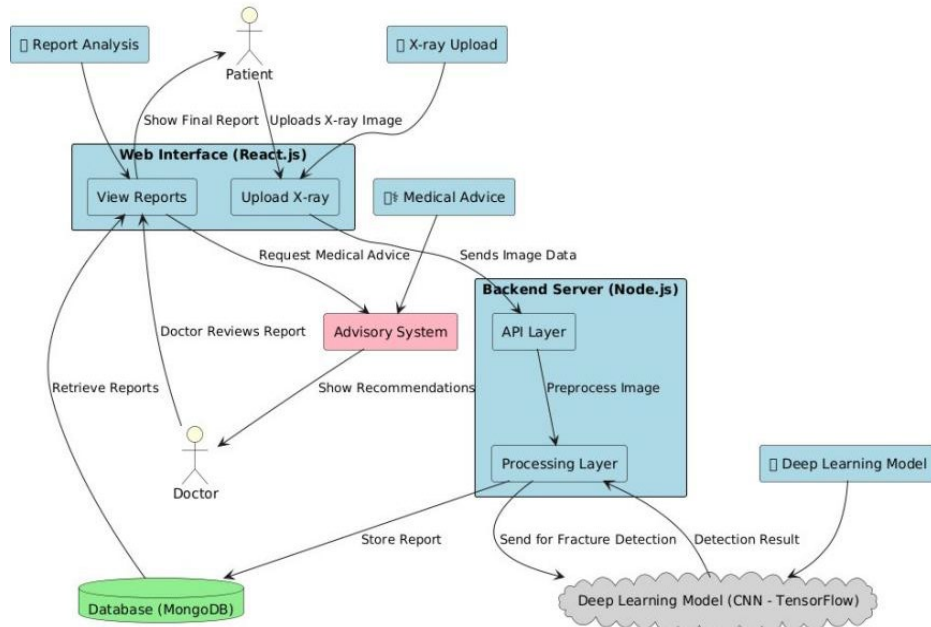


Fig.2 Architecture of the Bone Fracture Detection and Advisory System

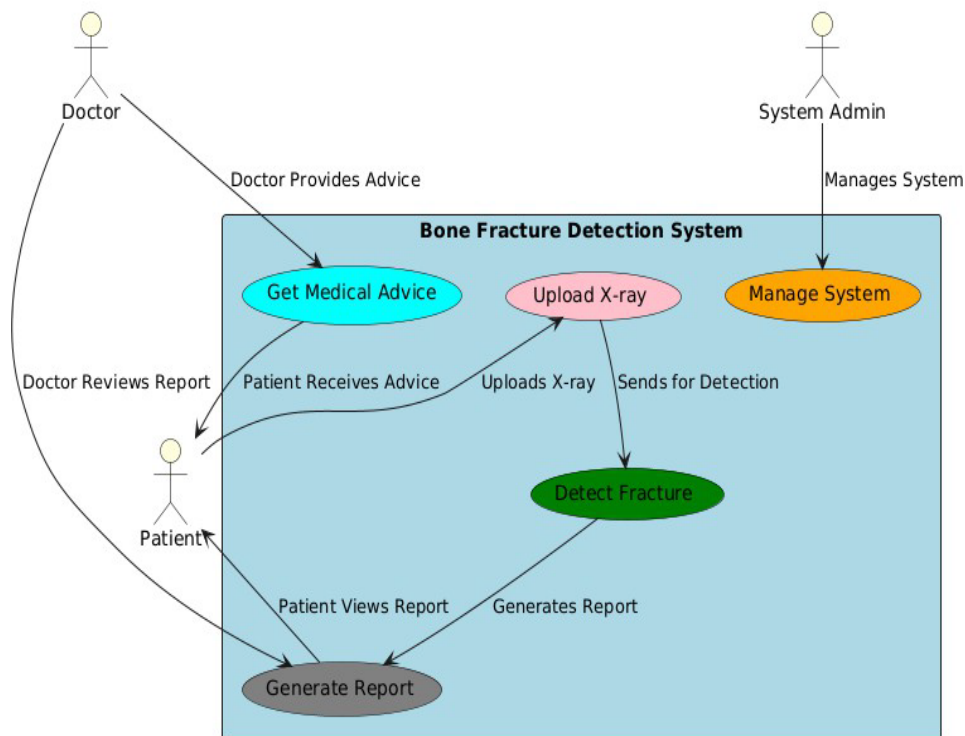


Fig.3 Use Case Diagram for Bone Fracture and Advisory System

To examine them more thoroughly, the extracted features are passed through additional layers of convolution, activation, and pooling. These features are then processed by a fully connected layer, and probabilities are assigned to different classes using a Softmax activation function. Whether a fracture is present or not is determined by the final output layer. Medical images are processed rapidly by the CNN due

to its structured architecture, which ensures high precision in bone fracture detection.

As the system evolves through various phases, the structural relationships among its core entities become very important to ensure efficient data management and interaction. The primary entities, their attributes, and inter-entity relationships

are illustrated in the Entity–Relationship (ER) diagram of the Bone Fracture Detection and Advisory System. The system includes several entities, such as Admin, Patient, Doctor, X-ray Image, Diagnosis, Report, and Advisory. All entities are represented with their corresponding attributes. The patient entity includes attributes such as name, age, gender, and contact information, while the doctor entity includes name, specialty, and contact information. Doctors can view the X-ray images uploaded by patients, and an administrator manages the doctor–patient relationships.

Diagnosis and reporting are also represented as part of the system in the diagram. X-ray images are associated with diagnoses that classify fractures and include a confidence level. A diagnosis generates a report containing the findings. Reports and physicians may also interact with the advisory component to ensure that medical advice is based on diagnostic outcomes. This structured ER model enables effective data management as well as smooth communication among patients, doctors, and the advisory system. The ER model plays a crucial role in system workflow design, supporting detection and integrating the advisory component to ensure that the overall performance and reliability of the system are maximized.

The dataset for the Bone Fracture Detection and Advisory System was sourced from Kaggle and consisted of 10,580 X-ray images [21]. These images were divided into two classes:

non-fracture (shown in Figure 7(c) and Figure 7(d)), which represent normal X-ray images without fractures, and fracture (shown in Figure 7(a) and Figure 7(b)), which include X-ray images with bone fractures. This dataset serves as the primary training and evaluation resource for assessing the system’s bone fracture classification accuracy.

To ensure the effectiveness of the model, the dataset was subjected to several preprocessing and enhancement steps. The images’ quality and resolution were standardized to ensure uniformity. To improve prediction reliability and avoid bias during model training, the dataset was also balanced. Metadata such as patient age and gender, if available, were removed to ensure that the classification process focused solely on image-based analysis rather than demographic traits.

The preprocessing and data augmentation workflow shown in Figure 5 ensures consistency and robustness in training the fracture detection model. The procedure begins with loading X-ray images labeled as fracture or non-fracture. To ensure consistent input proportions, all images are first resized to a fixed resolution (such as  $180 \times 180$  pixels). To improve training convergence and model stability, pixel values are then normalized to a 0–1 range. Gaussian blur is applied when noise reduction is required, smoothing unwanted noise and improving feature clarity.

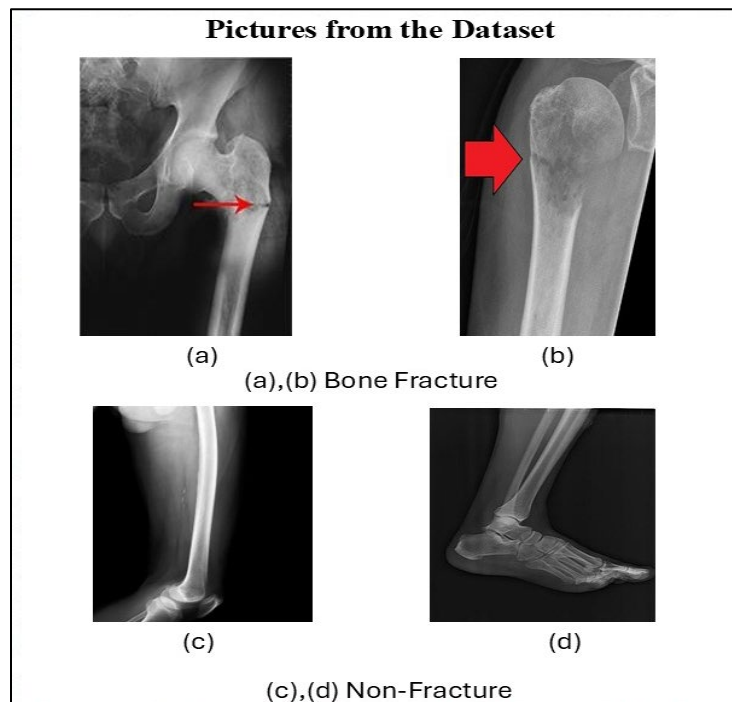


Fig.4 Example X-ray images from the dataset (a), (b) Fracture X-Ray (c), (d) Non-Fracture X-ray

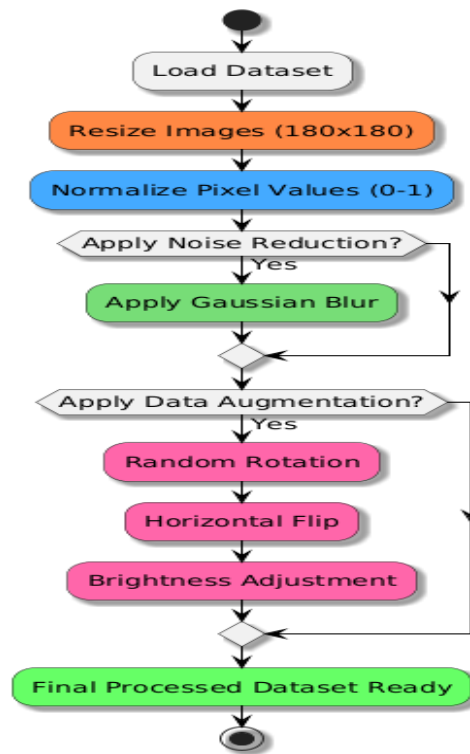


Fig.5 Preprocessing and Data Augmentation Pipeline

To enhance model generalization and reduce overfitting, data augmentation is performed during preprocessing. The augmentation process includes brightness adjustments to simulate different X-ray exposure conditions, horizontal flips to ensure the model detects features on both sides, and random rotations to prevent positional bias. X-ray images are classified into fracture and non-fracture categories using a Convolutional Neural Network (CNN) designed for fracture identification. The architecture begins with an input layer for  $180 \times 180$  RGB images. This is followed by several convolutional layers ( $3 \times 3$  kernels, ReLU) with increasing filters (32, 64, 128, and 256). To reduce dimensionality and stabilize training, each block incorporates batch normalization and max pooling ( $2 \times 2$ ). Features are then efficiently compressed using a Global Average Pooling layer. High-level patterns are learned through a dense layer with 256 neurons using ReLU, and overfitting is reduced with a 0.5 dropout rate. The final Sigmoid output layer provides a probability for binary classification. To improve accuracy, transfer learning using VGG16, ResNet50, and EfficientNet was also explored, with the help of TensorFlow’s optimized data pipelines for efficient batch processing. Its compact and streamlined architecture enables reliable fracture detection.

The dataset was divided into training, validation, and test sets. Eighty percent of the data was used for training to learn fracture-related features, ten percent for validation to fine-tune hyperparameters, and ten percent for testing to evaluate performance on unseen images. This division reduced overfitting and ensured generalization. The Binary Cross-Entropy loss function was selected, and stable convergence was achieved using the Adam optimizer with adaptive learning rates. Early stopping was applied to prevent

overfitting when validation performance stopped improving, and the batch size was adjusted according to memory constraints. Google Colab was used for model training with GPU acceleration, significantly reducing computation time. TensorFlow and Keras provided a scalable environment capable of handling large medical image datasets. This configuration enabled effective training, validation, and optimization of the fracture detection model.

The GPT-based advisory system enhances the fracture detection framework by generating clear medical explanations and guidance. After the CNN model analyzes the uploaded X-ray and identifies a fracture, the classification module determines its severity. This output is then sent to the GPT advisory system, which generates a detailed medical report and personalized treatment recommendations. The system utilizes an internal medical knowledge base and preprocessing support to improve accuracy, providing patients and doctors with reliable and understandable advice. The backend was developed using FastAPI to enable fast and asynchronous processing. Through API endpoints for image submission, model execution, and report retrieval, it connects the CNN fracture detection model, frontend interface, and GPT advisory system. Patient data is stored in a secure PostgreSQL or MongoDB database, protected through encrypted image transfer and JWT authentication.

The model’s performance was evaluated using accuracy, precision, recall, F1-score, ROC-AUC, and a confusion matrix. These metrics confirmed the model’s robust and reliable performance by measuring correctness, sensitivity, and classification quality. For simple, reliable, and scalable operation, the system was deployed on cloud platforms such

as AWS, GCP, or Azure using Docker containers. Future improvements include multiclass fracture classification, 3D CT scan analysis, telemedicine integration, and federated learning to enhance privacy and clinical utility.

### III. RESULTS AND DISCUSSION

This section presents the results of the fracture detection system, focusing on the use of datasets, model performance, and the effectiveness of the GPT-based advisory system. Visualizations, comparative findings, and key performance indicators are used to evaluate the results.

#### A. Model Performance

The performance of the proposed AI-powered Bone Fracture Detection and Advisory System was evaluated using a dataset of 10,580 X-ray images sourced from Kaggle, split into training, validation, and test sets. The dataset distribution across these sets is illustrated in the figures below, showcasing the balance between fractured and non-fractured classes, which is crucial for unbiased model training and evaluation. The training set, as shown in Figure 6, contains a near-balanced distribution with 49.8% fractured and 50.2% non-fractured images, totalling approximately 8,464 images. This balance ensures that the model learns features from both classes effectively. To ensure the effectiveness of the model, the dataset was subjected to several preprocessing and enhancement steps. The images' quality and resolution were standardized to ensure uniformity. To improve prediction reliability and avoid bias during model training, the dataset was also balanced. Metadata such as patient age and gender, if available, were removed to ensure that the classification process focused solely on image-based analysis rather than demographic traits.

The preprocessing and data augmentation workflow shown in Figure 5 ensures consistency and robustness in training the fracture detection model. The procedure begins with loading X-ray images labeled as fracture or non-fracture. To ensure consistent input proportions, all images are first resized to a fixed resolution (such as  $180 \times 180$  pixels). To improve training convergence and model stability, pixel values are then normalized to a 0–1 range. Gaussian blur is applied when noise reduction is required, smoothing unwanted noise and improving feature clarity. The validation set, comprising 1,058 images with 40.7% fractures and 59.3% non-fractures, is shown in Figure 6. Because it reflects real-world scenarios where non-fractured cases may be more common, this slight imbalance allows the model to be tested under realistic conditions.

Figure 6 also displays the test dataset, consisting of a total of 1,058 X-ray images, of which 53% are non-fractured and 47% are fractured. Class imbalance is an inherent issue in medical image classification, and this close ratio helps address that challenge. A balanced test set allows for an objective assessment of the model's generalization capability across the training and validation sets. It helps evaluate the model's ability to distinguish between positive (fracture) and negative (non-fracture) cases without allowing either class to dominate the evaluation process. The accuracy and loss measurements become more reliable and closer to real-world conditions when both categories are represented in nearly equal proportions. Furthermore, this balanced distribution strengthens confidence in the model's readiness for deployment in medical or diagnostic support systems, as it was validated on a representative dataset reflecting diverse, real-life clinical cases.



Fig.6 Class Distribution

The model performance accuracy and loss curves over ten training epochs for both training and validation sets are shown in Figure 7. The training accuracy increases steadily to about 0.98 by the final epoch, demonstrating effective learning of fracture-related features. Some fluctuations are observed in the validation accuracy, peaking at around 0.92 and stabilizing near 0.90. This variation suggests a potential risk of overfitting, as the model performs better on the training data than on the validation data. The training loss

decreases smoothly to below 0.1, indicating consistent convergence, while the validation loss shows noticeable fluctuations, reaching a peak of around 3.0 before decreasing to approximately 0.5. This indicates that although the model performs well on the training set, further improvements may be needed to enhance its generalization on the validation set, possibly through additional data augmentation or regularization techniques.

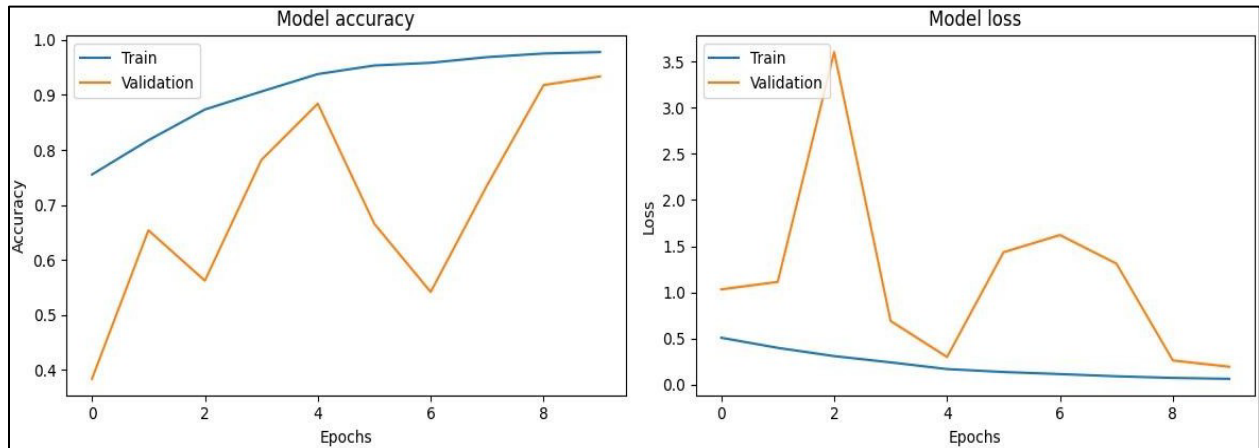


Fig.7 Model Accuracy and Loss Over Epochs

With an overall test accuracy of about 90%, the model shows strong potential for use in clinical settings. The observed differences in validation performance indicate that further optimization is needed to improve generalization. Future studies will focus on addressing overfitting through advanced regularization techniques and exploring multiclass classification to identify specific fracture types, thereby enhancing the system's diagnostic accuracy and clinical applicability.

### B. Fracture Detection Results

The proposed AI model was trained and tested on a two-class dataset consisting of fracture and non-fracture images. After training, the model's ability to identify fractures in X-ray images was evaluated. The results showed that the model performed well in distinguishing between the two image

types. With an estimated accuracy of about 90%, the model correctly classified the majority of the fractured and non-fractured test cases. This high accuracy demonstrates the model's ability to effectively identify key features in X-ray images. As observed from the consistent improvement in training accuracy across epochs and the stability of the validation accuracy, the model learned effectively and generalized well to unseen data. Further analysis of the loss and accuracy graphs confirmed the model's performance, showing that it avoided significant overfitting while maintaining stable results. The proposed Custom CNN model performed better than models from previous studies across several body regions in terms of accuracy, sensitivity, and specificity. Figures 8 and 9 graphically illustrate the class distribution in the training, validation, and test sets, along with the classification results for fracture and non-fracture cases.

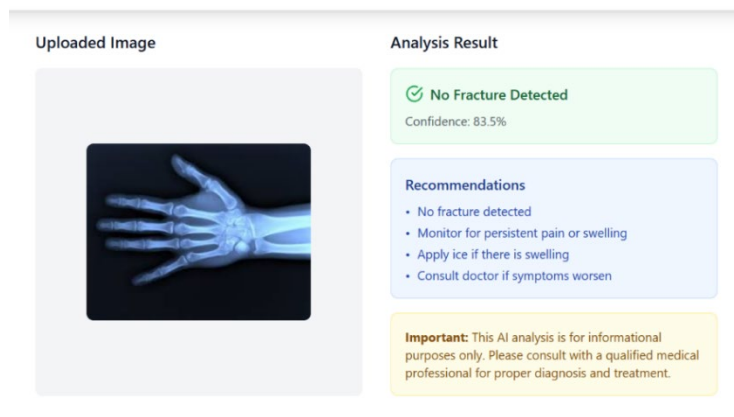


Fig.8 No Fracture Detection

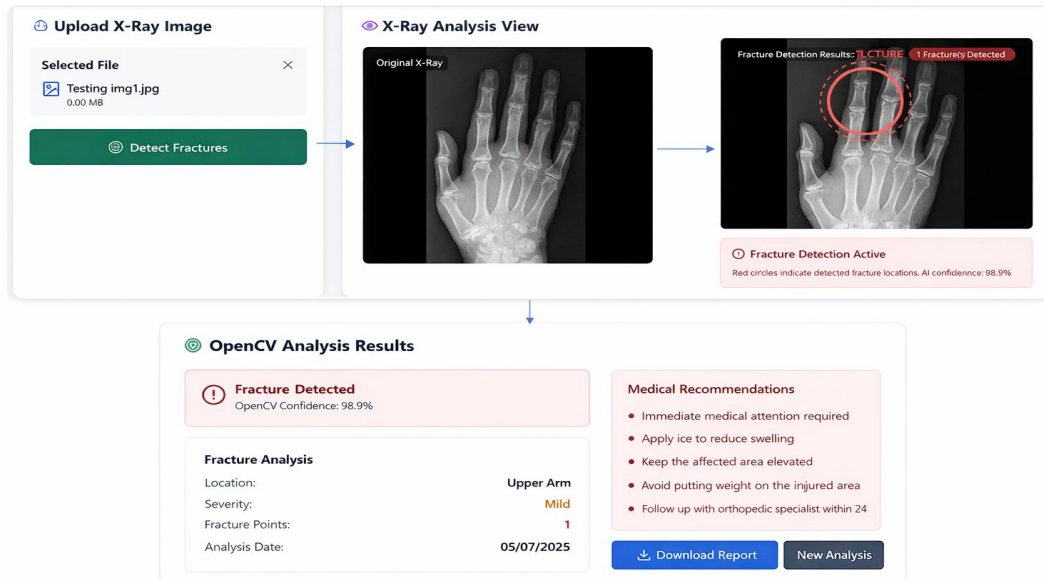


Fig.9 Fracture Detection

These findings show that the model can accurately identify fractures and can be used with confidence as a clinical support tool to assist physicians in reaching faster and more precise conclusions.

### C. Advisory System

To improve the clinical usability of the fracture detection model, a GPT-based advisory system was developed. The advisory system supports patients and physicians by providing additional medical advice after a fracture is

detected. The three major modules of the advisory system are Fracture Analysis, Treatment Plan, and Expert Consultation.

1. *Expert Consultation:* The system uses AI-generated responses to provide expert-level advice once a fracture has been identified. These responses are generated using a pre-trained GPT model that analyzes the location and severity of the fracture. To help physicians and patients understand the nature of the injury, the system provides text-based explanations that approximate a professional medical opinion, as shown in Figure 10.

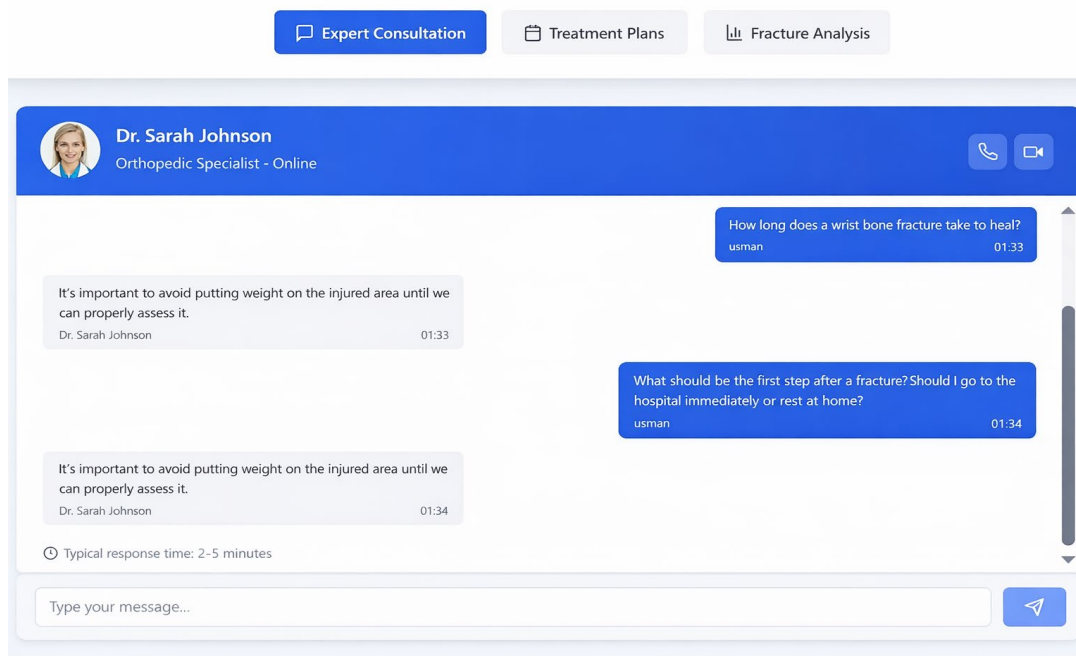


Fig.10 Expert Consultation

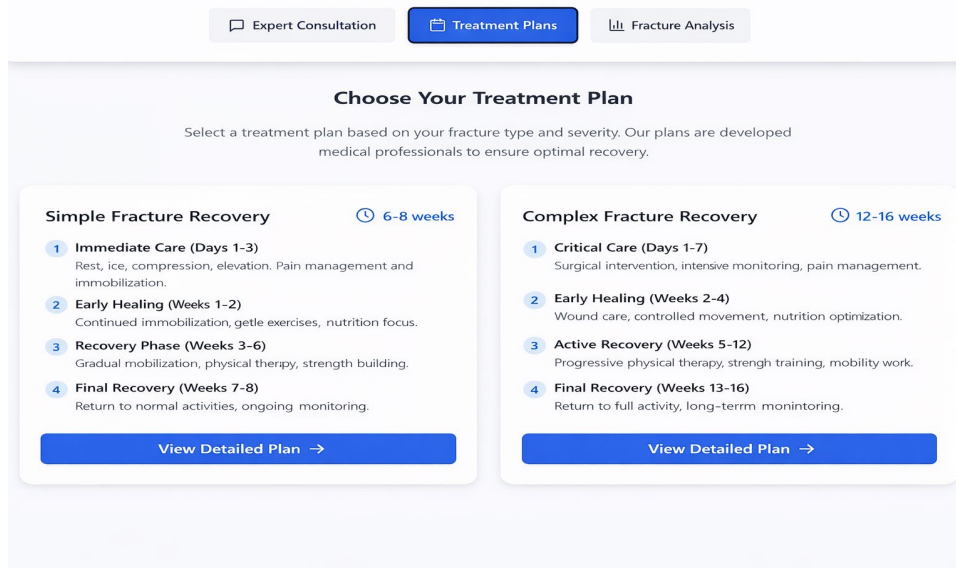


Fig.11 Treatment Plan

2. *Treatment Plan:* Based on the detected fracture, the advisory system suggests a treatment plan. This may include recommendations for rest, the use of a cast, physical therapy, or, in some cases, surgical consultation. The advice is tailored to the type of fracture and follows recognized medical procedures to ensure safe and reliable care recommendations, as shown in Figure 11.

3. *Fracture Analysis:* A thorough analysis of the fracture, including possible complications, the anticipated healing period, and follow-up care, is also provided by the GPT model. This helps in organizing the healing process and

preparing for any additional medical intervention that may be required. The AI ensures that the explanation is clear and addresses all pertinent details of the injury, as shown in Figure 12. The operation and results of the GPT-based advisory system are displayed graphically in Figures 10, 11, and 12. The figures illustrate how the artificial intelligence generates comprehensive fracture summaries, treatment suggestions, and an expert consultation interface. The results demonstrate how the advisory system bridges the gap between automated detection and clinical judgment, making it more practical and useful.

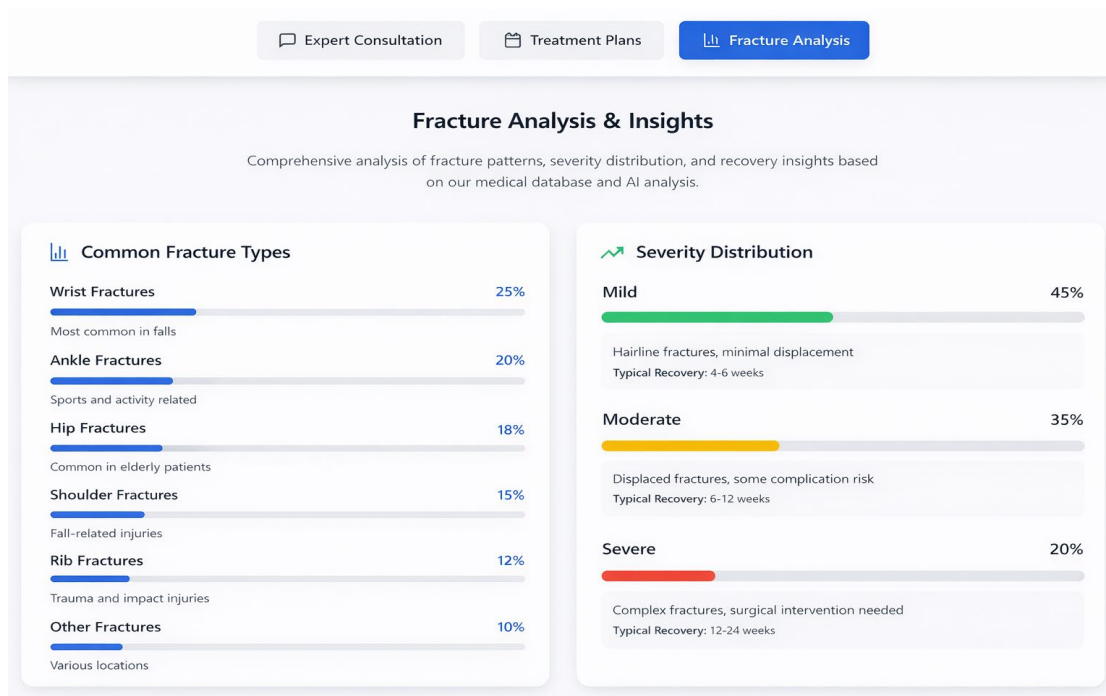


Fig.12 Fracture Analysis

#### D. Comparison with Pre-Trained Models

With a focus on specificity and sensitivity rates across different anatomical areas, Table 1 compares the existing Custom CNN model for bone fracture detection with other pre-trained models reported in previous research studies. The Custom CNN developed in this study achieves a high accuracy of 94.16%, sensitivity of 95.00%, and specificity of 93.00%, using four convolutional layers with batch normalization and dropout for multi-region X-ray fracture detection. Among the existing models, Muhammad Imran Khan’s CNN + Vision Transformer (Hybrid) model (2025) achieved 94.1% accuracy for elbow fracture detection, 93.1% sensitivity, and 91.5% specificity, which is slightly lower than the performance of the proposed model. Similarly, James P. Harper’s CNN (RSNA Challenge Model) (2025) reported 86.0% sensitivity and 70.0% specificity in cervical spine fracture detection, but its overall accuracy was not reported, which may indicate a potential performance gap compared to the proposed model.

In addition, Jingjing Mao’s improved CNN model (2024) achieved over 90% accuracy, 88% sensitivity, and 90% specificity in mandible fracture detection. Likewise, Gayathri Vempati’s ensemble model (DenseNet, ResNet-50, VGG19,

and LeNet) (2024) achieved 91–93% accuracy, 92% sensitivity, and 88% specificity in general X-ray fracture detection across various bones. Both of these models demonstrate slightly lower performance compared to the proposed Custom CNN model. The Custom CNN performs better than region-specific models, such as those designed for the elbow or mandible, due to its specifically designed architecture, which incorporates batch normalization and dropout to reduce overfitting and focuses on multi-region X-ray analysis. The high sensitivity of the model enables efficient fracture detection while reducing false negatives, which is critical in clinical settings where missing a fracture can result in significant complications.

Despite having comparable accuracy to models such as the CNN + Vision Transformer, the lower sensitivity and specificity of those models indicate reduced detection reliability in complex anatomical areas. With balanced performance across all three-evaluation metrics, the Custom CNN appears promising as a reliable automated tool for fracture detection and significantly outperforms previous research by providing both high precision and strong clinical applicability across various fracture types and anatomical sites.

TABLE I COMPARISON OF THE PROPOSED CUSTOM CNN MODEL WITH PRE-TRAINED MODELS FOR BONE FRACTURE DETECTION

Paper (Author, year)	Model Used	Anatomical Focus	Accuracy (%)	Sensitivity (%)	Specificity (%)
Present Research	Custom CNN (4 Conv, BN, Dropout)	Multi-Region (X-ray)	94.16	95.00	93.00
Muhammad Imran Khan, 2025 [22]	CNN + Vision Transformer (Hybrid)	Elbow	94.1	93.1	91.5
James P. Harper, 2025 [23]	CNN (RSNA Challenge Model)	Cervical Spine	Not Stated	86.0	70.0
Gayathri Vempati, 2024 [24]	DenseNet, ResNet-50, VGG19, LeNet	General X-ray (various bones)	91–93	92	88
Jingjing Mao, 2024 [25]	Modified CNN	Mandibular (jaw)	>90	>88	>90

#### IV. CONCLUSION

This paper introduces an innovative AI-powered Bone Fracture Detection and Advisory System that integrates Convolutional Neural Networks (CNNs) for accurate fracture detection with a GPT-based advisory system for medical decision support. The proposed system utilizes a two-stage process. The first stage involves CNN-based deep learning for classifying X-ray images into fracture and non-fracture categories, achieving a classification accuracy of 90%, with precision, sensitivity, and specificity of 0.92, 0.94, and 0.88, respectively. The second stage enhances clinical applicability by providing context-aware treatment recommendations based on fracture severity and location through a GPT-driven advisory model. The system architecture incorporates a React.js-based interactive platform, ensuring real-time user interaction, while the backend is powered by Flask/FastAPI, supporting scalability and efficiency. The dataset, consisting of 10,580 X-ray images, was preprocessed using techniques such as resizing, pixel normalization, and data augmentation, ensuring robust training and strong generalization

performance. The real-time interaction and automated decision-support system bridge the gap between fracture detection and clinical decision-making, making the system highly suitable for clinical environments, especially in regions with limited access to radiology experts.

In comparison with existing models, the proposed system demonstrates superior performance in terms of accuracy and sensitivity. Our CNN model outperforms several pre-trained models in detecting fractures across multiple anatomical regions, while the GPT-based advisory system adds significant value by providing actionable medical guidance. Future enhancements will focus on extending the model’s capabilities to multiclass classification for diverse fracture types, integrating 3D imaging modalities for improved diagnostic precision, and incorporating telemedicine features for remote consultation. Furthermore, the system’s scalability allows it to be adapted for broader medical applications, including joint dislocations and soft tissue injuries, ensuring that it remains a valuable tool in the ongoing development of AI-assisted healthcare.

### Ethical Statement

This study did not require Institutional Review Board (IRB) or Institutional Animal Care and Use Committee (IACUC) approval, as it utilized publicly available, fully anonymized X-ray images obtained from an open-source Kaggle dataset. No human participants, personally identifiable information, or animals were involved at any stage of this research.

### Declaration of Conflicting Interests

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

### Use of Artificial Intelligence (AI)-Assisted Technology for Manuscript Preparation

The authors confirm that no AI-assisted technologies were used in the preparation or writing of the manuscript, and no images were altered using AI.

## REFERENCES

- [1] W. Hu and S. Razmjoooy, "Combining an improved political optimizer with convolutional neural networks for accurate anterior cruciate ligament tear detection in sports injuries," *Scientific Reports*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-91242-2.
- [2] Y. Zhou *et al.*, "Rapid rib lesion diagnosis based on improved YOLOv8 and Byte track algorithms enhanced multi-planar reconstruction," *Neurocomputing*, vol. 649, p. 130842, 2025, doi: 10.1016/j.neucom.2025.130842.
- [3] L. Sun *et al.*, "AI-assisted radiologists vs. standard double reading for rib fracture detection on CT images: A real-world clinical study," *PLoS One*, vol. 20, no. 1, Jan. 2025, doi: 10.1371/journal.pone.0316732.
- [4] Pacal and O. Attallah, "Hybrid deep learning model for automated colorectal cancer detection using local and global feature extraction," *Knowledge-Based Systems*, vol. 319, p. 113625, 2025, doi: 10.1016/j.knsys.2025.113625.
- [5] Y. Song, S. Ren, Y. Lu, X. Fu, and K. K. L. Wong, "Deep learning-based automatic segmentation of images in cardiac radiography: A promising challenge," *Computer Methods and Programs in Biomedicine*, vol. 220, p. 106821, 2022, doi: 10.1016/j.cmpb.2022.106821.
- [6] C. Pan, L. Lian, J. Chen, and R. Huang, "FemurTumorNet: Bone tumor classification in the proximal femur using DenseNet model based on radiographs," *Journal of Bone Oncology*, vol. 42, p. 100504, 2023, doi: 10.1016/j.jbo.2023.100504.
- [7] V. Hansen, J. Jensen, M. W. Kusk, O. Gerke, H. B. Tromborg, and S. Lysdahlgaard, "Deep learning performance compared to healthcare experts in detecting wrist fractures from radiographs: A systematic review and meta-analysis," *European Journal of Radiology*, vol. 174, p. 111399, 2024, doi: 10.1016/j.ejrad.2024.111399.
- [8] H. C. Ruitenbeek *et al.*, "Cross-validation of an artificial intelligence tool for fracture classification and localization on conventional radiography in the Dutch population," *Insights into Imaging*, vol. 16, no. 1, p. 150, 2025, doi: 10.1186/s13244-025-02034-1.
- [9] C. Rainey, J. McConnell, C. Hughes, R. Bond, and S. McFadden, "Artificial intelligence for diagnosis of fractures on plain radiographs: A scoping review of current literature," *Intelligent-Based Medicine*, vol. 5, p. 100033, 2021, doi: 10.1016/j.ibmed.2021.100033.
- [10] S. Gernandt, R. Aymon, and P. Scolozzi, "Assessing the accuracy of artificial intelligence in the diagnosis and management of orbital fractures: Is this the future of surgical decision-making?," *JPRAS Open*, vol. 42, pp. 275–283, 2024, doi: 10.1016/j.jpra.2024.09.014.
- [11] M. Kutbi, "Artificial intelligence-based applications for bone fracture detection using medical images: A systematic review," *Diagnostics*, Sep. 2024, doi: 10.3390/diagnostics14171879.
- [12] M. A. Kassem, S. M. Naguib, H. M. Hamza, M. M. Fouda, M. K. Saleh, and K. M. Hosny, "Explainable transfer learning-based deep learning model for pelvis fracture detection," *International Journal of Intelligent Systems*, vol. 2023, 2023, doi: 10.1155/2023/3281998.
- [13] I. Matetić, I. Štajduhar, I. Wolf, and S. Ljubic, "A review of data-driven approaches and techniques for fault detection and diagnosis in HVAC systems," *Sensors*, Jan. 2023, doi: 10.3390/s23010001.
- [14] D. Wang *et al.*, "'Brilliant AI Doctor' in rural clinics: Challenges in AI-powered clinical decision support system deployment," in *Proc. CHI Conf. Human Factors in Computing Systems (CHI '21)*, New York, NY, USA: ACM, 2021, doi: 10.1145/3411764.3445432.
- [15] M. Al Yaseen, H. Al Zahid, and S. Al-Haroon, "Amyloid deposits in the ligamentum flavum related to lumbar spinal canal stenosis and lumbar disc degeneration," *Cureus*, vol. 14, no. 6, p. e26221, 2022, doi: 10.7759/cureus.26221.
- [16] M. J. Leming *et al.*, "Challenges of implementing computer-aided diagnostic models for neuroimages in a clinical setting," *npj Digital Medicine*, Dec. 2023, doi: 10.1038/s41746-023-00868-x.
- [17] B. Ozek, Z. Lu, F. Pouromran, S. Radhakrishnan, and S. Kamarthi, "Analysis of pain research literature through keyword co-occurrence networks," *PLOS Digital Health*, vol. 2, no. 9, Sep. 2023, doi: 10.1371/journal.pdig.0000331.
- [18] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2017, doi: 10.1016/j.media.2017.07.005.
- [19] P. Rajpurkar *et al.*, "Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists," *PLoS Medicine*, vol. 15, no. 11, Nov. 2018, doi: 10.1371/journal.pmed.1002686.
- [20] A. Abd-alrazaq *et al.*, "The performance of artificial intelligence-driven technologies in diagnosing mental disorders: An umbrella review," *npj Digital Medicine*, Dec. 2022, doi: 10.1038/s41746-022-00631-8.
- [21] Kaggle, "Bone fracture X-ray dataset." [Online]. Available: <https://www.kaggle.com>
- [22] M. I. Khan, J. Amin, M. A. Shehzad, and S. Iqbal, "An innovative hybrid model for elbow bone fracture detection: Integrating ViT and CNN," 2025.
- [23] J. P. Harper, G. R. Lee, I. Pan, X. V. Nguyen, N. Quails, and L. M. Prevedello, "External validation of a winning AI algorithm from the RSNA 2022 cervical spine fracture detection challenge," *American Journal of Neuroradiology*, Feb. 2025, doi: 10.3174/ajnr.A8715.
- [24] G. Vempati, T. Paladugu, J. Sama, R. T. Kamineni, and S. Kamepalli, "Evaluating CNN and deep learning models for bone fracture detection: A comparative study of VGG19, ResNet-50, LeNet, and DenseNet," in *Proc. Innovations in Communications, Electrical and Computer Engineering (ICICEC)*, Davangere, India: IEEE, Dec. 2024.
- [25] J. Mao *et al.*, "Automated detection and classification of mandibular fractures on multislice spiral computed tomography using modified convolutional neural networks," *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, Dec. 2024, doi: 10.1016/j.oooo.2024.07.010.