



FractureSight-HCI: A Human-Centered Interactive System for AI-Assisted X-ray Fracture Visualization

Asma Mustafa Alhadi¹, Fathi A. Hamhoum²

¹Department of Computer Engineering and Sciences, Faculty of Science, University of Zawia, Libya

Email: ssoma4482@gmail.com

²Computer Science Department, Faculty of Science, University of Zawia, Libya

Corresponding Author's Email: Fathi.Hamhoum@zu.edu.ly

¹ <https://orcid.org/0000-0002-3615-842X>

² <https://orcid.org/0009-0005-1903-0853>

Received: 10/8/2026 Revised: 15/8/2026 Accepted: 05/9/2026 Publication: 30/9/2026

Abstract

Fracture detection in X-ray images is a critical task in medical image analysis, where accurate localization of suspected fracture regions can support faster, clearer, and more focused clinical interpretation. In this work, we present FractureSight-HCI, a human-centered interactive system for AI-assisted X-ray fracture visualization and clinical decision support. The proposed system combines an image processing and deep learning pipeline with an interactive Human–Computer Interaction interface to transform model predictions into clear visual outputs that can be reviewed by physicians. The technical component of the system was developed using a YOLOv8-based object detection model trained on more than 20,000 radiographic fracture samples. The training process was conducted using Google Colab and continued for more than three hours, reflecting a practical experimental workflow rather than a simple static demonstration.

The trained model was evaluated using standard object detection metrics. The preliminary experimental results achieved a Precision of 89.5%, Recall of 79.5%, mAP50 of 89.6%, and mAP50-95 of 49.3%. These results indicate promising fracture localization capability, especially for identifying suspicious regions that require visual clinical review. However, the contribution of this work is not limited to model performance. A major focus is the design of an interpretable and physician-centered interface that displays the original X-ray image, presents the AI-assisted fracture localization result, and highlights the suspected fracture region using a bounding box. This visual presentation reduces ambiguity in the AI output and supports a more understandable decision-support workflow.

From the HCI perspective, the system was designed according to a human-in-the-loop approach, where the AI model provides visual assistance while the physician remains responsible for the final medical decision. The interface supports image selection, visual analysis, highlighted fracture visualization, and clinical review in a simplified workflow. By integrating large-scale model training, image processing, object detection, visual explanation, and human-centered interaction, FractureSight-HCI demonstrates a practical and structured approach for presenting medical AI results in a clinically meaningful and user-friendly manner. The proposed system highlights the importance of combining technical accuracy with interface clarity, interpretability, and human control in AI-assisted medical imaging applications.

Index Terms— Human–Computer Interaction, Image Processing, X-ray Fracture Detection, YOLOv8, Object Detection, Medical Image Analysis, Clinical Decision Support, Human-in-the-loop, AI-Assisted Visualization, Visual Explanation.

I. INTRODUCTION

Medical imaging plays a fundamental role in modern clinical diagnosis, particularly in the interpretation of X-ray images for bone fracture detection. X-ray imaging is widely used in emergency departments, orthopedic clinics, and radiology centers because it is fast, accessible, and effective for evaluating skeletal injuries in different anatomical regions. However, accurate fracture interpretation can still be challenging due to variations in image quality, fracture type, bone structure, patient positioning, and the visual subtlety of some fracture patterns. In busy clinical environments, missed or delayed fracture identification may affect the speed and quality of medical decision-making. Therefore, computer-assisted fracture detection has become an important research direction in medical image analysis.

Recent advances in deep learning and object detection have shown strong potential for supporting radiographic fracture analysis. Public radiographic datasets and annotated X-ray collections have enabled the development and evaluation of machine learning models for fracture detection and localization [1,2]. Deep learning-based object detection models have also been investigated for detecting fracture regions in X-ray images, demonstrating that artificial intelligence can support localization of suspected fracture areas rather than producing only image-level classification outputs [3]. More recently, YOLO-based approaches have achieved promising results in X-ray fracture detection tasks, showing the suitability of real-time object detection architectures for medical image localization [4]. These studies confirm the value of integrating image processing and AI-based detection methods into fracture analysis systems.

Despite these advances, technical model performance alone is not sufficient for practical clinical adoption. A model may achieve promising Precision, Recall, or mAP values, but its output must still be presented in a way that is clear, interpretable, and useful for the medical user. In clinical settings, physicians need to understand where the suspected abnormality is located and how the AI output should be considered during medical review. For this reason, artificial intelligence in medical imaging should not be treated only as a prediction engine, but as part of a human-centered decision-support workflow. Recent studies on explainable medical imaging AI emphasize that medical AI systems require human-centered design, transparency, and careful presentation of model outputs [11]. In addition, guidelines for human-AI interaction highlight the importance of keeping users informed, supporting meaningful control, and avoiding over-reliance on automated systems [19].

Human-Computer Interaction is especially important in AI-assisted medical imaging because the interface acts as the communication layer between the model and the physician. If the interface is unclear, overloaded, or difficult to interpret, the clinical value of the AI model may be reduced even if the model performs well technically. Human-centered explainable AI research also emphasizes that model explanations should be evaluated from the user's perspective, not only from the algorithmic perspective [13]. In healthcare, trustworthy AI systems must combine technical performance with transparency, usability, and responsible human oversight [16]. Therefore, a practical AI-assisted fracture detection system should not only detect suspected fracture regions, but also present them in a visually understandable manner while preserving the physician as the final decision-maker.

To address this need, this paper presents **FractureSight-HCI**, a human-centered interactive system for AI-assisted X-ray fracture visualization. The proposed system combines an image processing and deep learning workflow based on **YOLOv8** with an interactive HCI interface that displays the original X-ray image and highlights the suspected fracture region using a bounding box. The AI model was trained on more than **20,000 radiographic fracture samples** using **Google Colab**, with a training duration exceeding **three hours**. Preliminary experimental results achieved a **Precision of 89.5%**, **Recall of 79.5%**, **mAP50 of 89.6%**, and **mAP50-95 of 49.3%**. These results indicate that the model is capable

of localizing suspicious fracture regions and producing visual outputs that can support further clinical review.

The proposed system is designed according to a human-in-the-loop principle. The AI component provides visual assistance by identifying and highlighting the suspected fracture region, while the physician remains responsible for the final clinical decision. This design helps reduce ambiguity in AI output and supports a clearer decision-support workflow. Instead of replacing medical judgment, **FractureSight-HCI** aims to strengthen clinical attention by transforming model predictions into an interpretable visual format that can be reviewed by users.

The main contributions of this paper are summarized as follows:

- Development of an AI-assisted X-ray fracture localization workflow based on **YOLOv8**.
- Training of the object detection model on more than **20,000 radiographic fracture samples** using **Google Colab**.
- Evaluation of the trained model using standard object detection metrics, including **Precision, Recall, mAP50**, and **mAP50-95**.
- Design of a human-centered interactive interface that presents the original X-ray and highlights the suspected fracture region for physician review.
- Integration of Image Processing and Human–Computer Interaction principles into a practical human-in-the-loop clinical decision-support workflow.

The remainder of this paper is organized as follows. Section II reviews related work in AI-based fracture detection, medical image analysis, explainable AI, and human-centered clinical decision support. Section III describes the proposed methodology and system design. Section IV presents the implementation details of the AI model and interactive interface. Section V discusses the evaluation and results. Section VI presents the discussion, limitations, and future work, followed by the conclusion.

II. RELATED WORK

AI-assisted fracture detection and human-centered medical interfaces have become increasingly important research directions in medical image analysis and clinical decision-support systems. Since the proposed **FractureSight-HCI** system combines **Image Processing** and **Human–Computer Interaction**, this section reviews related work from two main perspectives: AI-based X-ray fracture detection and human-centered interaction for interpretable medical AI systems.

A. AI-Based X-ray Fracture Detection

Fracture detection in X-ray images has received significant attention due to the wide use of radiography in emergency and orthopedic diagnosis. The availability of annotated radiographic datasets has supported the development of machine learning and deep learning models for detecting and localizing bone fractures. The GRAZPEDWRI-DX dataset, for example, provides a large collection of pediatric trauma X-ray images with annotations that support fracture detection and localization research [1,2]. Although this dataset focuses on wrist trauma, it has become an important reference in the broader field of radiographic fracture detection because it provides structured X-ray data and bounding-box annotations suitable for object detection experiments.

Deep learning-based object detection has been widely investigated for fracture localization in radiographic images. Hardalaç et al. studied fracture detection in X-ray images using deep learning-based

object detection models and demonstrated that bounding-box-based detection can help identify suspected fracture regions in medical images [3]. Such methods are important because they provide visual localization rather than only classifying the entire image as normal or abnormal. This localization ability is directly related to clinical decision support, where the physician needs to know the exact region that requires attention.

Recent YOLO-based approaches have shown strong potential for real-time fracture detection. Ju and Cai investigated **YOLOv8** for X-ray fracture detection and reported promising results using an object detection framework [4]. Although many existing studies focus on specific anatomical regions, the underlying method of detecting suspicious fracture regions using bounding boxes can be extended to broader X-ray fracture visualization workflows. Review studies have also shown that deep learning is increasingly used for skeletal fracture detection and that AI-assisted systems can support radiographic interpretation when properly validated and integrated into clinical workflows [5,6].

In addition to academic research, clinical and health-technology evaluations have examined the value of AI-assisted fracture detection in urgent care. NICE reviewed artificial intelligence technologies designed to help detect fractures on X-rays, emphasizing their potential role as supportive tools in clinical settings [7]. Similarly, early value assessment studies have discussed the use of AI software to assist fracture detection in radiographic images [8]. These references support the idea that AI-assisted fracture detection is clinically relevant, but they also highlight that AI outputs must be integrated carefully into decision-support workflows rather than replacing clinical judgment.

B. Medical Image Analysis and Object Detection Metrics

Object detection systems are commonly evaluated using metrics such as **Precision**, **Recall**, **mean Average Precision at an IoU threshold of 0.5 (mAP50)**, and **mean Average Precision across multiple IoU thresholds (mAP50-95)**. These metrics are important because fracture localization requires both correct detection and accurate spatial alignment between the predicted bounding box and the true fracture region. Precision reflects how many detected fracture regions are correct, while Recall reflects how many actual fracture regions are successfully detected. mAP-based metrics provide a more complete evaluation of detection quality by considering localization accuracy and detection confidence.

Large-scale radiographic datasets such as MURA have also contributed to abnormality detection research in musculoskeletal radiographs [9]. Although MURA mainly supports abnormality classification rather than bounding-box fracture localization, it demonstrates the importance of large medical imaging datasets for training and evaluating AI models. Other studies have explored calibrated deep learning ensembles for abnormality detection in musculoskeletal radiographs, emphasizing the need for reliable and interpretable AI outputs in medical applications [10]. These studies support the technical foundation of the proposed system, where model performance is evaluated quantitatively before the output is presented through a clinical interface.

C. Human-Centered AI and Explainable Medical Imaging

While technical accuracy is important, medical AI systems must also be understandable, usable, and trustworthy for clinical users. Chen et al. argued that explainable medical imaging AI needs human-centered design because model explanations must be meaningful to the people who use them in real settings [11]. This is directly relevant to **FractureSight-HCI**, where the model output is presented as a visual bounding box to help the physician understand the suspected fracture region.

Human-centered evaluation of explainable AI applications has also been emphasized in recent sys-

tematic reviews. Kim et al. reviewed human-centered evaluation methods for explainable AI and highlighted the importance of evaluating AI systems from the user’s perspective, not only from the algorithmic perspective [12]. Rong et al. also discussed user studies for model explanations and emphasized that explanations should support human understanding, decision-making, and trust [13]. In medical imaging, this means that an AI model should not only produce a prediction, but should present the result in a way that supports clinical reasoning.

Frameworks for explainable AI design and evaluation further support the need for transparency, usability, and meaningful user control. Mohseni et al. proposed a multidisciplinary framework for designing and evaluating explainable AI systems, showing that explanation quality should be connected to user needs and task context [14]. In medical imaging specifically, the CLAIM checklist provides guidance for reporting AI studies and encourages clear documentation of data, model design, evaluation, and clinical relevance [15]. These works support the design direction of **FractureSight-HCI**, where the model output is translated into a clear visual result that can be reviewed by the physician.

D. Human-in-the-Loop Clinical Decision Support

Human-in-the-loop design is an important principle for clinical AI systems because medical decisions require professional judgment, responsibility, and contextual understanding. Trustworthy AI guidelines in healthcare emphasize that AI systems should be safe, reliable, transparent, and deployable in ways that preserve human oversight [16]. In this context, AI should support the physician rather than replace the physician.

Explainable AI studies also highlight the risks of black-box systems in sensitive domains such as healthcare. Barredo Arrieta et al. discussed the concepts, opportunities, and challenges of explainable AI, emphasizing the need for responsible AI systems that users can understand and trust [17]. Adadi and Berrada also reviewed explainable AI approaches and described the importance of opening the black box when automated systems influence decision-making [18]. These ideas are particularly relevant in fracture detection because the physician must be able to inspect the suspected region visually before making a final interpretation.

Guidelines for Human-AI Interaction proposed by Amershi et al. provide practical principles for designing AI interfaces, including making system status visible, supporting user control, and helping users understand system behavior [19]. In addition, Holzinger et al. discussed causability and explainability in medical AI, emphasizing the importance of explanations that are useful to human experts [20]. These works support the HCI design of **FractureSight-HCI**, where the interface presents the original radiograph, highlights the suspected fracture region, and keeps the final decision under physician control.

E. Research Gap

Based on the reviewed literature, several gaps can be identified. First, many studies focus mainly on model performance and detection accuracy, while less attention is given to how the detection result is presented to the physician. Second, existing fracture detection systems often report quantitative metrics but do not always provide a clear human-centered interface that supports clinical interpretation. Third, explainable AI research emphasizes the importance of transparency and user-centered evaluation, yet many medical AI applications still present outputs in ways that may be difficult for users to interpret. Finally, there remains a need for systems that integrate AI-based fracture localization with a simple, visual, and physician-centered HCI workflow.

To address these gaps, **FractureSight-HCI** integrates **YOLOv8-based fracture localization** with a human-centered interface that presents the suspected fracture region as a visual bounding box. The sys-

tem is designed to support clinical attention, reduce ambiguity in AI output, and preserve the physician as the final decision-maker. This combination of image processing, object detection, visual explanation, and human-in-the-loop interaction distinguishes the proposed system from approaches that focus only on algorithmic performance.

Table 1: Summary of related work and relevance to the proposed FractureSight-HCI system.

Research Area	Representative References	Relevance to This Work
Radiographic fracture datasets	GRAZPEDWRI-DX dataset [1,2]	Supports the use of annotated X-ray data and bounding-box annotations for fracture detection and localization research.
Deep learning-based fracture detection	Hardalaç et al. [3], Ju and Cai [4]	Provides the technical foundation for using object detection models such as YOLO-based architectures to localize suspected fracture regions.
Skeletal fracture detection reviews	Su et al. [5], Husarek et al. [6]	Shows the growing role of deep learning in skeletal fracture detection and highlights the need for reliable AI-assisted radiographic interpretation.
Clinical AI fracture detection assessment	NICE [7], Farmer et al. [8]	Supports the clinical relevance of AI-assisted fracture detection while emphasizing its role as a supportive tool rather than a replacement for physicians.
Musculoskeletal radiograph analysis	MURA [9], He et al. [10]	Demonstrates the importance of large-scale radiographic data and reliable model outputs in medical image analysis.
Human-centered explainable AI	Chen et al. [11], Kim et al. [12], Rong et al. [13]	Supports the need for medical AI outputs that are understandable, user-centered, and useful for clinical decision-making.
Explainable AI frameworks and reporting	Mohseni et al. [14], CLAIM checklist [15]	Supports transparent design, reporting, and evaluation of AI systems in medical imaging.
Trustworthy and human-controlled AI	FUTURE-AI [16], Amer-shi et al. [19], Holzinger et al. [20]	Supports the human-in-the-loop design where AI provides visual assistance while the physician remains the final decision-maker.

III. METHODOLOGY AND SYSTEM DESIGN

This section describes the methodology and system design of the proposed **FractureSight-HCI** system. The system was designed to combine AI-based fracture localization with a human-centered clinical interface. The methodology consists of five main stages: radiographic image collection and preparation, model training using **YOLOv8**, fracture localization, visual result presentation, and physician-centered review through a human-in-the-loop workflow.

F. System Overview

The proposed **FractureSight-HCI** system follows a structured workflow that starts with X-ray image

preparation and ends with an interpretable visual output displayed to the user. The input consists of radiographic X-ray images containing bone structures from different fracture-related cases. The images are prepared and used to train an object detection model capable of localizing suspected fracture regions. After training, the model produces bounding-box predictions that identify the possible fracture area. These outputs are then presented through an interactive HCI interface that displays the original X-ray image and highlights the suspected fracture region in a clear visual format.

The system is not designed to replace clinical judgment. Instead, it acts as an AI-assisted visual decision-support tool. The AI model provides a localized visual suggestion, while the physician remains responsible for reviewing the image and making the final clinical decision. This design follows a human-in-the-loop approach, where automated detection and human expertise work together.

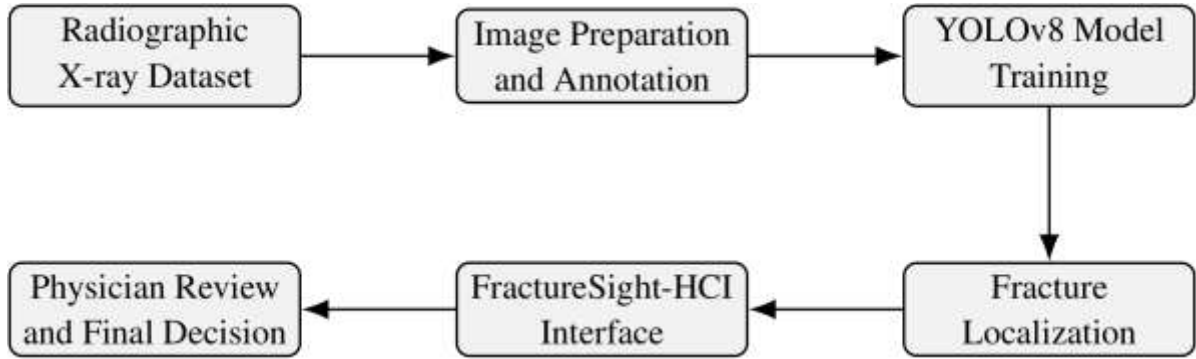


Figure 1: Overall architecture of the proposed FractureSight-HCI system.

G. Dataset Preparation and Image Processing

The model was trained using more than **20,000 radiographic fracture samples**. The dataset preparation stage involved organizing X-ray images and their corresponding fracture annotations. Since the proposed system is based on object detection, the annotation process focuses on identifying the suspected fracture region using bounding boxes. Each bounding box defines the spatial location of a fracture-related area within the X-ray image.

Before training, the images were prepared to be compatible with the **YOLOv8** training pipeline. This preparation included image organization, annotation formatting, dataset splitting, and verification of image-label consistency. The dataset was divided into training and validation subsets to allow the model to learn fracture localization patterns and to evaluate its performance on unseen data.

From an image processing perspective, the X-ray images represent grayscale radiographic data where fracture regions may appear as subtle discontinuities, irregular edges, or abnormal bone structures. Therefore, the role of the object detection model is to learn visual patterns associated with suspected fracture regions and produce localized predictions that can be interpreted visually.

H. YOLOv8-Based Fracture Localization

YOLOv8 was selected as the object detection model because it provides a strong balance between detection accuracy and processing speed. Object detection is suitable for this task because fracture analysis requires not only determining whether a fracture may exist, but also identifying where the suspicious region is located. The model receives an X-ray image as input and outputs bounding boxes, confidence values, and class predictions for suspected fracture regions.

The training process was conducted using **Google Colab** and continued for more than **three hours**. The model was trained on more than **20,000 samples**, which reflects a practical experimental setup rather than a simple interface demonstration. After training, the model was evaluated using standard object detection metrics, including **Precision**, **Recall**, **mAP50**, and **mAP50-95**.

The Intersection over Union (IoU) is used to measure the overlap between the predicted bounding box and the ground-truth bounding box. It is defined as follows:

$$IoU = \frac{Area(B_p \cap B_g)}{Area(B_p \cup B_g)} \quad (1)$$

where B_p represents the predicted bounding box and B_g represents the ground-truth bounding box. Precision measures the proportion of correct detections among all predicted detections:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall measures the proportion of actual fracture regions successfully detected by the model:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

where TP represents true positives, FP represents false positives, and FN represents false negatives.

The F1-score provides a balanced measure between Precision and Recall:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Mean Average Precision (mAP) summarizes detection performance across classes or IoU thresholds:

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (5)$$

where AP_i represents the average precision for class or threshold i , and N represents the number of evaluated classes or thresholds.

I. Human-Centered Interface Design

The HCI component of **FractureSight-HCI** was designed to make the AI output understandable and clinically useful. Instead of presenting only numerical results, the interface displays the original X-ray image and then presents the AI-assisted result with a visible bounding box around the suspected fracture region. This visual explanation helps users quickly understand the location of the AI-detected region.

The interface supports a simplified interaction workflow. The user selects or provides an X-ray image, views the original radiograph, and then reviews the AI-assisted visualization. The system presents the suspected fracture region in a clear visual form while avoiding unnecessary interface complexity. This design supports usability, interpretability, and focused clinical review.

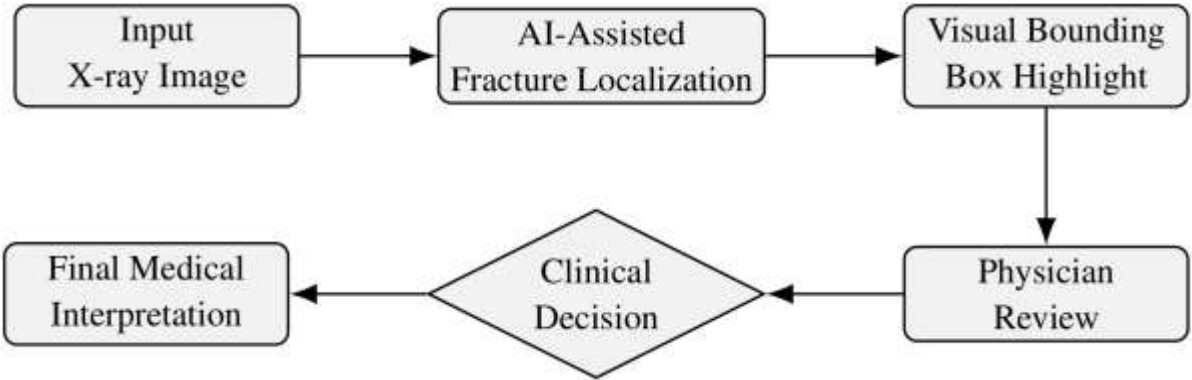


Figure 2: Human-in-the-loop workflow of the proposed FractureSight-HCI system.

J. Human-in-the-Loop Workflow

The proposed system follows a human-in-the-loop workflow. In this workflow, the AI model performs fracture localization and provides a visual suggestion, but the physician remains responsible for the final medical interpretation. This is important because medical decisions require clinical context, professional judgment, and patient-specific information that cannot be replaced by an automated model.

The human-in-the-loop design reduces the risk of over-reliance on AI by presenting the model output as decision support rather than a final diagnosis. The bounding box acts as a visual guide that directs attention to a suspicious region, while the final decision remains under human control.

K. System Components

The proposed system consists of three main components. The first component is the AI model training pipeline, which includes dataset preparation, annotation, **YOLOv8** training, and performance evaluation. The second component is the fracture localization module, which processes X-ray images and generates bounding-box predictions. The third component is the HCI interface, which presents the original image and AI-assisted visualization in a clear and user-centered manner.

Together, these components create a complete workflow that connects image processing, object detection, and human-centered interaction. This integration allows the system to transform technical AI predictions into visual decision-support outputs that are easier for medical users to interpret.

Table 2: Main components of the proposed FractureSight-HCI system.

Component	Description
Dataset Preparation	Organizes more than 20,000 radiographic fracture samples and prepares image-label pairs for object detection training.
Image Annotation	Defines suspected fracture regions using bounding boxes to support supervised fracture localization.
YOLOv8 Training	Trains an object detection model using Google Colab for more than three hours .
Fracture Localization	Generates bounding-box predictions and confidence values for suspected fracture regions in X-ray images.
HCI Interface	Displays the original radiograph and AI-assisted fracture visualization in a clear and user-centered format.
Human-in-the-Loop Review	Keeps the physician responsible for the final clinical decision while using AI output as visual decision support.

IV. IMPLEMENTATION

This section presents the implementation details of the proposed **FractureSight-HCI** system. The implementation was designed to demonstrate a practical integration between AI-based fracture localization and a human-centered interactive interface. The system consists of two main implementation parts: the **YOLOv8-based training and fracture localization pipeline**, and the **FractureSight-HCI interface** for visual presentation and physician-centered review.

L. Development Environment

The technical development was carried out using Python-based tools and **Google Colab**. Google Colab was used for training the **YOLOv8** object detection model because it provides a practical cloud-based environment for deep learning experiments. The training process continued for more than **three hours** and was performed on more than **20,000 radiographic fracture samples**. This training stage represents the core image processing and AI component of the proposed system.

The interactive interface was implemented using **Streamlit**, which allows rapid development of web-based Python applications. Streamlit was selected because it provides a simple and effective way to display medical images, interface elements, messages, and AI-assisted visualization results. The interface was designed to be clear, direct, and suitable for demonstrating the human-centered workflow of the proposed system.

Table 3: Development environment and software tools used in FractureSight-HCI.

Component	Implementation Details
Programming Language	Python was used for model training, image handling, and interface implementation.
Training Platform	Google Colab was used to train the YOLOv8 model in a cloud-based environment.
AI Model	YOLOv8 object detection model was used for suspected fracture localization.
Dataset Size	The model was trained on more than 20,000 radiographic fracture samples .
Training Duration	The training process continued for more than three hours .
Interface Framework	Streamlit was used to implement the interactive web-based interface.
Output Type	Original X-ray image, AI-assisted result, bounding-box highlight, and visual decision-support output.

M. YOLOv8 Model Implementation

The **YOLOv8** model was implemented as an object detection model for localizing suspected fracture regions in X-ray images. The dataset was organized into image and label folders according to the YOLO training structure. Each image was associated with an annotation file containing bounding-box coordinates for the suspected fracture region. This structure allowed the model to learn the spatial relationship between fracture patterns and their locations in the radiographic image.

During training, the model learned to identify visual fracture-related features and produce bounding-box predictions. The output of the model includes the predicted location of the suspected fracture region and the confidence score associated with the detection. The trained model was evaluated using standard detection metrics, including **Precision**, **Recall**, **mAP50**, and **mAP50-95**.

The implemented training pipeline can be summarized as follows:

1. Organize radiographic X-ray images and fracture annotation files.
2. Prepare the dataset according to the YOLO object detection format.
3. Split the dataset into training and validation subsets.
4. Train the YOLOv8 model using Google Colab.
5. Evaluate the trained model using standard object detection metrics.
6. Export and use the trained model for AI-assisted fracture visualization.

N. Interface Implementation

The **FractureSight-HCI** interface was implemented to present the AI output in a form that is understandable to the user. The interface displays the system title, a short explanation of the workflow, and an image selection area. After the user selects an X-ray image, the original radiograph is displayed first. Then, the AI-assisted result is shown with a bounding box highlighting the suspected fracture region.

The purpose of this design is to allow the user to compare the original image with the AI-assisted visualization. Instead of presenting only numerical predictions, the system provides a visual explanation that directs attention to the suspected region. This makes the output easier to interpret and supports the physician-centered review process.



Figure 3: Welcome screen of the proposed FractureSight-HCI interface.



Figure 4: Image selection interface showing the available X-ray input options in the navigation panel.

O. Visual Output and Bounding Box Highlighting

The most important visual element in the system is the bounding box that highlights the suspected fracture region. The bounding box acts as an interpretable visual cue that links the AI model prediction to a specific area in the X-ray image. This is important in medical image analysis because the user needs to know not only whether a fracture may exist, but also where the suspicious region is located.

The interface uses a clear visual presentation style to avoid unnecessary complexity. The original image and the detected result are displayed in a structured layout, allowing the user to inspect the input image and the highlighted suspected region. This design supports visual comparison, interpretability, and focused review.



Figure 5: AI-assisted visualization of a representative X-ray image showing the suspected fracture region highlighted for physician review.

P. Human-in-the-Loop Implementation

The system was implemented according to a human-in-the-loop approach. In the implemented workflow, the AI model provides an assisted localization output, but the physician remains responsible for interpreting the result. The interface explicitly supports this principle by presenting the AI result as visual decision support rather than a final diagnosis.

This implementation choice is important because medical decisions require clinical expertise and context. The system therefore avoids presenting the AI output as an absolute conclusion. Instead, it highlights the suspected region and leaves the final decision to the physician. This supports safer and more responsible use of AI in medical imaging.

Q. Implemented Workflow

The implemented workflow of the proposed system begins when the user opens the FractureSight-HCI interface. The user selects an X-ray image for analysis, and the system displays the original image. After that, the AI-assisted result is shown with a bounding box around the suspected fracture region. The user can visually compare the original image and the highlighted result before clinical review.

The workflow is summarized as follows:

1. The user opens the FractureSight-HCI interface.
2. The user selects or provides an X-ray image.
3. The system displays the original radiographic image.
4. The trained YOLOv8 model provides fracture localization output.
5. The suspected fracture region is highlighted using a bounding box.
6. The physician reviews the visual output.
7. The final medical interpretation remains under physician control.

R. Implementation Summary

Table 4 summarizes the main implementation components of the proposed FractureSight-HCI system.

Table 4: Implementation summary of the proposed FractureSight-HCI system.

Implementation Part	Description
AI Training Pipeline	YOLOv8 model trained using more than 20,000 radiographic fracture samples on Google Colab .
Training Duration	The training process continued for more than three hours .
Fracture Localization	The trained model produces bounding-box predictions for suspected fracture regions.
Interface Framework	Streamlit was used to build a simple and interactive web-based interface.
Visual Explanation	The suspected fracture region is highlighted using a bounding box to improve interpretability.
User Interaction	The user can select an X-ray image, view the original image, and review the AI-assisted result.
Human-in-the-Loop Design	The AI output is presented as decision support, while the physician remains responsible for the final clinical decision.

V. EVALUATION AND RESULTS

This section presents the evaluation and results of the proposed **FractureSight-HCI** system. The evaluation was conducted from two complementary perspectives. The first perspective focuses on the technical performance of the **YOLOv8-based fracture localization model** using standard object detection metrics. The second perspective focuses on the **Human-Computer Interaction** aspect by analyzing the clarity, interpretability, and usefulness of the interface in presenting AI-assisted fracture visualization.

S. Experimental Setup

The **YOLOv8-based fracture localization model** was trained using more than **20,000 radiographic fracture samples**. The training process was performed using **Google Colab** and continued for more than **three hours**. The dataset was organized according to the YOLO object detection format, where each X-ray image was associated with a label file containing bounding-box annotations for suspected fracture regions.

The purpose of the experiment was to evaluate the ability of the trained model to localize suspected fracture regions in X-ray images. Since the proposed system is designed as a visual decision-support system, the main focus was not only classification accuracy, but also the ability to produce localized bounding-box outputs that can be presented clearly through the HCI interface.

T. Quantitative Model Evaluation

The trained model was evaluated using standard object detection metrics, including **Precision**, **Recall**, **mAP50**, and **mAP50-95**. These metrics are commonly used to evaluate detection performance because they consider both detection correctness and localization quality.

The experimental results showed that the trained model achieved a **Precision of 89.5%**, **Recall of 79.5%**, **mAP50 of 89.6%**, and **mAP50-95 of 49.3%**. The Precision value indicates that most predicted fracture regions were correct, while the Recall value shows the ability of the model to detect actual fracture regions. The mAP50 value indicates strong detection performance at an IoU threshold of 0.5, while the mAP50-95 value provides a stricter evaluation across multiple IoU thresholds.

These results demonstrate that the model has promising capability in localizing suspected fracture regions. Although **mAP50-95** is lower than **mAP50**, this is expected in medical object detection tasks because stricter IoU thresholds require highly precise bounding-box alignment. In the context of the proposed system, the bounding box is used as a visual guide for physician review rather than as an automatic final diagnosis.

Table 5: Preliminary performance results of the YOLOv8-based fracture localization model.

Metric	Value	Interpretation
Precision	89.5%	High correctness of predicted fracture regions.
Recall	79.5%	Good ability to detect actual suspected fracture regions.
mAP50	89.6%	Strong localization performance at IoU threshold 0.5.
mAP50-95	49.3%	Stricter localization evaluation across multiple IoU thresholds.

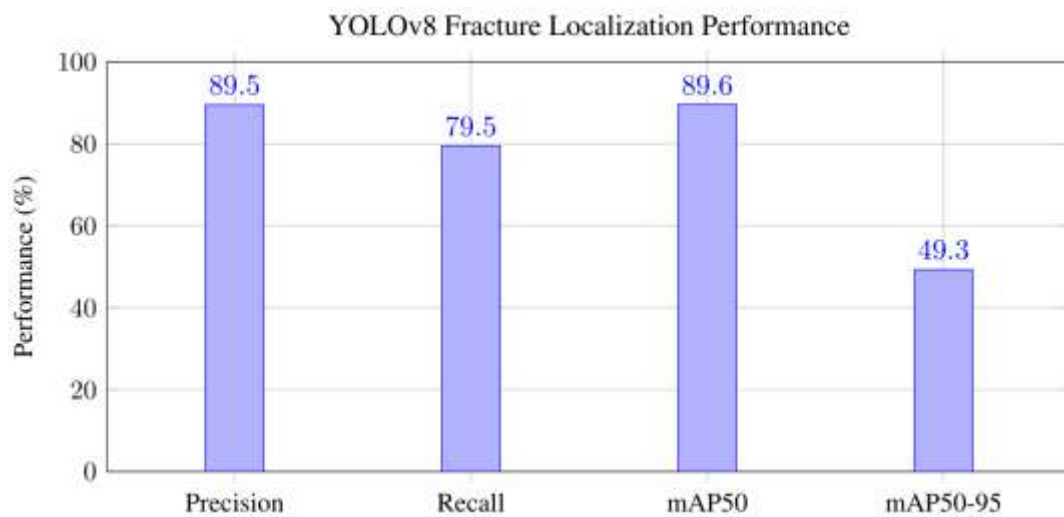


Figure 6: Bar chart showing the preliminary performance metrics of the YOLOv8-based fracture localization model.

U. Visual Output Evaluation

The output of the trained model was integrated into the **FractureSight-HCI** interface to provide visual decision support. The interface displays the original X-ray image and then presents the AI-assisted

visualization with a bounding box around the suspected fracture region. This output format is important because it allows the user to understand the location of the suspected abnormality instead of receiving only a numerical or textual prediction.

The visual result shows that the system can transform the model prediction into an interpretable interface output. The bounding box provides a direct visual cue that supports focused review and helps reduce ambiguity in the AI result. This is especially important in medical image analysis, where the location of the finding is as important as the prediction itself.

V. HCI-Based Evaluation

From the **Human–Computer Interaction** perspective, the interface was evaluated according to clarity, simplicity, visual feedback, interpretability, and physician control. The interface was designed to reduce unnecessary complexity by presenting a clear workflow: image selection, original X-ray display, AI-assisted visualization, and physician review.

The system supports visual feedback by showing the suspected fracture region using a bounding box. This helps the user understand the AI output quickly and supports a more focused review process. The interface also supports physician control by presenting the output as decision support rather than a final medical conclusion. This human-in-the-loop design helps maintain responsible use of AI in medical imaging.

Table 6: HCI evaluation criteria for the FractureSight-HCI interface.

Evaluation Criterion	Description
Visual Clarity	The interface clearly displays the original X-ray image and the AI-assisted fracture visualization.
Interpretability	The bounding box provides an understandable visual explanation of the suspected fracture region.
Ease of Use	The workflow is simple and allows the user to select an image and review the result without unnecessary complexity.
Clinical Usefulness	The highlighted region can help direct attention to the suspicious area during medical review.
Human Control	The AI output is presented as decision support, while the physician remains responsible for the final clinical decision.
Reduced Ambiguity	Visual highlighting reduces uncertainty by linking the AI prediction to a specific region in the X-ray image.

W. Preliminary Qualitative Feedback

The proposed interface was reviewed from a practical usability perspective by presenting the system workflow to medical and general users. The purpose of this review was to observe whether the interface was clear, understandable, and useful for presenting AI-assisted fracture localization. The reviewers focused on the clarity of the X-ray display, the visibility of the highlighted suspected fracture region, and the usefulness of the interface as a visual decision-support tool.

The feedback indicated that the bounding-box visualization was clear and easy to understand. Med-

ical reviewers emphasized that the system should be used as a supportive tool and not as a replacement for professional clinical judgment. General users also found the interface simple and readable because the workflow presents the original image and the AI-assisted result in a direct visual format.

This feedback supports the HCI direction of the proposed system. It suggests that the interface can present AI model outputs in a way that is understandable to users while maintaining the physician as the final decision-maker.

Table 7: Summary of preliminary qualitative feedback on the FractureSight-HCI interface.

Aspect	Feedback Summary
Visual Clarity	The highlighted suspected fracture region was clear and easy to notice.
Ease of Understanding	The interface workflow was simple and understandable.
Clinical Support	Medical reviewers indicated that the visual highlight could help direct attention to the suspected region.
Physician Role	Reviewers emphasized that the final medical decision should remain under physician control.
User Experience	General users found the interface readable and easy to follow.

X. Summary of Results

The evaluation results show that **FractureSight-HCI** successfully integrates AI-based fracture localization with a human-centered interface. The **YOLOv8** model achieved promising detection performance, particularly in terms of **Precision** and **mAP50**. The interface presented the model output in a clear visual form, allowing the suspected fracture region to be reviewed by the user.

Overall, the results demonstrate that combining object detection, image processing, and HCI principles can produce a practical AI-assisted medical imaging system. The system provides both technical detection capability and interpretable visual presentation, making it suitable as a clinical decision-support interface while preserving the physician as the final decision-maker.

VI. DISCUSSION, LIMITATIONS, AND FUTURE WORK

The results of this work show that **FractureSight-HCI** can provide a practical connection between AI-based fracture localization and human-centered medical image review. The **YOLOv8-based model** achieved promising preliminary performance, particularly in **Precision** and **mAP50**, indicating that the model can localize suspected fracture regions in X-ray images and generate visual outputs suitable for review. However, the value of the proposed system is not limited to detection performance. A key strength of the system is its ability to transform model predictions into a clear visual interface that supports interpretation and physician-centered decision-making.

From an image processing perspective, the system demonstrates how object detection can be used to identify suspicious fracture regions and represent them through bounding boxes. This is important because fracture analysis is a spatial task: the user needs to know where the suspected fracture is located, not only whether an abnormality may exist. Therefore, the use of localized visual output makes the

model result more useful for medical review.

From a **Human–Computer Interaction** perspective, the interface plays an essential role in making the AI output understandable. The system presents the original X-ray image and the AI-assisted highlighted result in a simple workflow. This reduces visual ambiguity and helps users focus on the suspected region. The human-in-the-loop design also supports responsible AI use by keeping the physician responsible for the final clinical decision. This is important in medical imaging because AI systems should support clinical attention rather than replace medical judgment.

In addition to technical evaluation, the interface was evaluated through qualitative interviews with medical and general users. The interview feedback was highly positive, especially regarding the clarity of the bounding-box visualization, the simplicity of the interaction workflow, and the usefulness of highlighting the suspected fracture region. Medical reviewers indicated that the system could help direct attention to suspicious regions in X-ray images, while general users found the interface understandable and easy to follow. This interview-based feedback supports the HCI value of the proposed system and shows that the interface can present AI-assisted outputs in a clear and user-centered manner.

Table 8: Summary of interview-based qualitative evaluation.

Evaluation Aspect	Interview Feedback Summary
Visual Clarity	Reviewers reported that the bounding-box highlight was clear and easy to identify.
Ease of Use	Users found the interface workflow simple, direct, and easy to follow.
Clinical Usefulness	Medical reviewers indicated that the system could help direct attention to suspicious fracture regions.
Interpretability	The visual output made the AI prediction easier to understand compared with numerical results alone.
Physician Control	Medical reviewers emphasized that the system should support the physician without replacing clinical judgment.
Overall Feedback	The overall interview feedback was highly positive and supported the usability direction of the proposed interface.

Despite the promising results, several limitations remain. First, although the model was trained on more than **20,000 radiographic fracture samples**, wider evaluation on additional fracture types, anatomical regions, and imaging conditions would strengthen the generalizability of the system. Second, while qualitative interviews were conducted and produced positive feedback, a larger questionnaire-based usability study with more participants would provide stronger statistical evidence. Third, the system currently focuses on visual fracture localization and does not include full clinical reporting, patient history, or integration with hospital information systems.

Future work will focus on expanding both the technical and HCI evaluation of the system. On the technical side, the dataset can be extended to include more anatomical regions, fracture types, and imaging variations. Additional experiments can also compare **YOLOv8** with other object detection models to evaluate accuracy, speed, and localization robustness. On the HCI side, future evaluation can include structured questionnaires, standardized usability scales, and larger semi-structured interviews with medical professionals. These methods can provide deeper insight into usability, trust, interpretability, perceived usefulness, and clinical workflow integration.

Future versions of **FractureSight-HCI** may also include confidence visualization, explanation notes, report generation, and integration with clinical review systems. The interface can be improved by allowing users to compare original and detected images more interactively, inspect model confidence, and provide physician feedback to improve later versions of the system. Overall, this work demonstrates that combining image processing, deep learning, visual explanation, and human-centered design can support the development of practical AI-assisted medical imaging systems.

VII. CONCLUSION

This paper presented **FractureSight-HCI**, a human-centered interactive system for AI-assisted X-ray fracture visualization and clinical decision support. The proposed system integrates image processing, deep learning-based object detection, and **Human-Computer Interaction** principles to transform AI model predictions into clear and interpretable visual outputs. A **YOLOv8-based model** was trained on more than **20,000 radiographic fracture samples** using **Google Colab**, with a training duration exceeding **three hours**. The model achieved promising preliminary results, including a **Precision of 89.5%**, **Recall of 79.5%**, **mAP50 of 89.6%**, and **mAP50-95 of 49.3%**.

The main contribution of this work is not limited to the technical performance of the model, but also includes the design of a physician-centered interface that presents the original X-ray image and highlights the suspected fracture region using a bounding box. This visual explanation helps reduce ambiguity, supports focused medical review, and improves the interpretability of AI-assisted outputs. The system follows a human-in-the-loop approach, where the AI model provides visual decision support while the physician remains responsible for the final clinical decision.

The interface was also evaluated through qualitative interviews with medical and general users, and the feedback was highly positive regarding visual clarity, ease of use, interpretability, and the usefulness of highlighting the suspected fracture region. These results demonstrate that combining AI-based fracture localization with human-centered interface design can provide a practical and understandable decision-support workflow for medical imaging applications.

Overall, **FractureSight-HCI** demonstrates the importance of integrating technical accuracy with usability, transparency, and human control in AI-assisted healthcare systems. Future work will focus on expanding the dataset, testing the system on more fracture types and anatomical regions, comparing additional object detection models, and conducting larger questionnaire-based usability evaluations with medical professionals.

References

- [1] E. Nagy, M. Janisch, F. Hrzić, E. Sorantin, and S. Tschauner, "A pediatric wrist trauma X-ray dataset (GRAZPEDWRI-DX) for machine learning," *Scientific Data*, vol. 9, Art. no. 222, 2022, doi: 10.1038/s41597-022-01328-z.
- [2] E. Nagy, M. Janisch, F. Hrzić, E. Sorantin, and S. Tschauner, "GRAZPEDWRI-DX," *figshare*, 2022, doi: 10.6084/m9.figshare.14825193.
- [3] F. Hardalaç et al., "Fracture Detection in Wrist X-ray Images Using Deep Learning-Based Object Detection Models," *Sensors*, vol. 22, no. 3, Art. no. 1285, 2022, doi: 10.3390/s22031285.
- [4] R.-Y. Ju and W. Cai, "Fracture Detection in Pediatric Wrist Trauma X-ray Images Using YOLOv8 Algorithm," *Scientific Reports*, vol. 13, Art. no. 20077, 2023, doi: 10.1038/s41598-023-47460-7.
- [5] Z. Su, X. Liu, and Z. Sun, "Skeletal Fracture Detection with Deep Learning: A Comprehensive

- Review,” *Diagnostics*, vol. 13, no. 20, Art. no. 3245, 2023, doi: 10.3390/diagnostics13203245.
- [6] J. Husarek et al., “Computer-assisted artificial intelligence fracture detection: a systematic review and meta-analysis of diagnostic test accuracy studies,” *Scientific Reports*, vol. 14, Art. no. 22475, 2024, doi: 10.1038/s41598-024-73058-8.
- [7] National Institute for Health and Care Excellence, “Artificial intelligence technologies to help detect fractures on X-rays in urgent care: early value assessment,” *NICE HealthTech Guidance*, 2025.
- [8] C. Farmer et al., “Artificial intelligence software to help detect fractures on X-rays in urgent care: an early value assessment,” *NIHR Health Technology Assessment / NCBI Bookshelf*, 2026.
- [9] P. Rajpurkar et al., “MURA: Large Dataset for Abnormality Detection in Musculoskeletal Radiographs,” *arXiv preprint arXiv:1712.06957*, 2017.
- [10] M. He, A. P. Sharp, and R. T. L. Chung, “A calibrated deep learning ensemble for abnormality detection in musculoskeletal radiographs,” *Scientific Reports*, vol. 11, Art. no. 9097, 2021, doi: 10.1038/s41598-021-88578-w.
- [11] H. Chen, C. Gomez, C.-M. Huang, and M. Unberath, “Explainable medical imaging AI needs human-centered design: Guidelines and evidence from a systematic review,” *npj Digital Medicine*, vol. 5, Art. no. 156, 2022, doi: 10.1038/s41746-022-00699-2.
- [12] J. Kim, H. Maathuis, and D. Sent, “Human-centered evaluation of explainable AI applications: A systematic review,” *Frontiers in Artificial Intelligence*, vol. 7, Art. no. 1456486, 2024, doi: 10.3389/frai.2024.1456486.
- [13] Y. Rong et al., “Towards Human-Centered Explainable AI: A Survey of User Studies for Model Explanations,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 4, pp. 2104–2122, 2024, doi: 10.1109/TPAMI.2023.3331846.
- [14] S. Mohseni, N. Zarei, and E. D. Ragan, “A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems,” *ACM Transactions on Interactive Intelligent Systems*, vol. 11, no. 3–4, Art. no. 24, 2021, doi: 10.1145/3387166.
- [15] J. Mongan, L. Moy, and C. E. Kahn Jr., “Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers,” *Radiology: Artificial Intelligence*, vol. 2, no. 2, Art. no. e200029, 2020, doi: 10.1148/ryai.2020200029.
- [16] K. Lekadir et al., “FUTURE-AI: International consensus guideline for trustworthy and deployable artificial intelligence in healthcare,” *BMJ*, vol. 388, Art. no. e081554, 2025, doi: 10.1136/bmj-2024-081554.
- [17] A. Barredo Arrieta et al., “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI,” *Information Fusion*, vol. 58, pp. 82–115, 2020, doi: 10.1016/j.inffus.2019.12.012.
- [18] A. Adadi and M. Berrada, “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence,” *IEEE Access*, vol. 6, pp. 52138–52160, 2018, doi: 10.1109/ACCESS.2018.2870052.
- [19] S. Amershi et al., “Guidelines for Human-AI Interaction,” in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–13, doi: 10.1145/3290605.3300233.
- [20] A. Holzinger, G. Langs, H. Denk, K. Zatloukal, and H. Müller, “Causability and Explainability of Artificial Intelligence in Medicine,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 9, no. 4, Art. no. e1312, 2019, doi: 10.1002/widm.1312.