iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81576

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Bone Fracture Classification and Automated Reporting System using YOLOv8, U-Net and Ensemble Learning

Vukoti Kushvanth[1], Balaji Vicharapu[2], Arangi Veda Seshasai[3], Singamreddy Venkata Ravindra Reddy[4], Bokka Iswarya[5]

[2]*Assistant Professor, Department of Data Science ad Cyber Security, University College of Engineering & Technology, Acharya Nagarjuna University, Nagarjuna Nagar, Guntur, AP, India -522510*

[1,3.4,5]*Students, Department of Data Science & Cyber Security, University College of Engineering & Technology, Acharya Nagarjuna University, Nagarjuna Nagar, Guntur, AP, India – 522510*

Abstract: ***Bone fractures are a leading cause of emergency orthopaedic visits worldwide, yet accurate diagnosis depends almost entirely on trained radiologists — a resource that is scarce in many hospitals and practically absent in rural clinics. This paper presents an end-to-end, AI-powered seven-stage pipeline for automatic bone fracture detection, morphological classification, and clinical report generation on plain upper-extremity radiographs. The pipeline combines CLAHE-based image preprocessing, YOLOv8 for bone region localisation, U-Net for pixel-level segmentation, a 110-dimensional radiological feature extractor, patient metadata encoding, and a nine-specialist ensemble classifier. Features from all three sources — image, segmentation, and metadata — are fused into a 158-dimensional vector before classification. Evaluated on a stratified hold-out set from the MURA v1.1 benchmark (40,009 radiographs across seven upper-extremity body parts), the system achieves 95.60% overall accuracy, a ROC-AUC of 99.72%, and a Cohen's Kappa of 0.94, which exceeds the published inter-radiologist agreement range of 0.60–0.78. Total CPU inference time is approximately three seconds per image. The system produces a structured clinical report covering fracture type, severity level, risk score, anatomical location, and treatment guidance, and is deployed as a Flask web application requiring no specialist hardware.***

Keywords: Bone Fracture Classification, YOLOv8, U-Net, Ensemble Classifier, MURA Dataset, Automated Radiology Report, Deep Learning, Clinical Decision Support

I. INTRODUCTION

Bone fractures are one of the most common injuries seen in emergency departments worldwide. According to the World Health Organization, musculoskeletal conditions affect hundreds of millions of people each year, often leading to pain, limited movement, or even long-term disability if not diagnosed properly. Yet, in most hospitals, fracture diagnosis still relies on radiologists manually reviewing X-rays—a process that has remained largely unchanged for decades. This traditional approach has some clear limitations. In busy hospitals, especially during peak times, radiologists can become overwhelmed, leading to delays in diagnosis and treatment. Even among experts, interpretations can vary, particularly for complex fractures, which may result in inconsistent decisions. The situation is even more challenging in rural or under-resourced areas, where access to skilled specialists is limited or sometimes unavailable. Although automated fracture detection has been widely researched, current solutions are not complete. Traditional image processing techniques work only for simple cases, while commercial AI tools—such as those developed by Zebra Medical Vision and Aidoc—are often expensive and require advanced hospital infrastructure. Most academic models focus only on identifying whether a fracture exists and do not provide detailed clinical insights. This work addresses these gaps by introducing a unified, easy-to-use system. It processes a standard X-ray through a multi-stage pipeline and generates a structured clinical report in just a few seconds, even on basic hardware. The system can identify multiple types of fractures, covers several upper-extremity regions, and is accessible through a simple web interface, making it practical even in low-resource settings.

A. Problem Statement

Orthopaedic fracture diagnosis today faces several interconnected challenges. In many rural or under-resourced areas, there is a shortage of radiologists, which can delay diagnosis from hours to even days—slowing down urgent treatment. Even when experts

are available, differences in interpretation (inter-observer variability) can lead to inconsistent fracture classification and treatment decisions. Speed is another critical issue, as serious injuries like femoral neck or spinal fractures need to be identified within minutes, yet manual workflows often fall short. Additionally, preparing structured clinical reports adds to the already heavy workload of healthcare professionals. While AI has been introduced, most existing systems are limited—they typically detect only whether a fracture is present and do not provide detailed information about its type or severity. As a result, many patients still receive delayed, inconsistent, and incomplete diagnoses. This work aims to overcome all these limitations through a unified automated pipeline.

B. Objectives

- Design and implement a seven-stage AI pipeline for bone fracture analysis that mirrors established radiological clinical workflows.
- Integrate YOLOv8 object detection for bone region-of-interest localisation with sub-second inference time.
- Employ a U-Net segmentation architecture that generates pixel-level fracture heatmaps and 16-dimensional bottleneck embeddings.
- Construct a 110-dimensional radiological feature extractor spanning intensity, texture, gradient, symmetry, and bone-density domains.
- Build a nine-specialist ensemble classifier achieving ≥95% classification accuracy across five fracture morphologies.
- Design an automated clinical report generator producing fracture type, severity level, anatomical location, risk score, and treatment recommendation.
- Deploy the complete system as a browser-accessible Flask web application with session history tracking, requiring no specialist hardware.
- Evaluate performance on the MURA v1.1 benchmark and document fully reproducible metrics.

C. Motivation

Three observations drove this work. First, deep learning has matured to the point where radiologist-level accuracy on medical image classification is genuinely achievable — the MURA benchmark and subsequent studies have shown this repeatedly. Second, the gap between research accuracy and clinical deployment remains wide: most systems stop at classification and never produce the structured reports that clinicians actually need. Third, the places that most need automated fracture support — rural clinics, under-resourced hospitals, low-income regions — are exactly the places where expensive commercial systems are least likely to reach. A lightweight, web-deployable system that delivers both classification and reporting in three seconds on ordinary hardware directly addresses that gap.

II. LITERATURE REVIEW

A. Classical Image Processing

Early computer-aided fracture detection leaned on edge operators — Canny, Sobel, Hough transforms — to find linear breaks in cortical bone. They worked for simple transverse fractures and failed badly on oblique, spiral, and comminuted patterns, producing false-negative rates too high for clinical use.

B. CNN-Based Binary Classification

Rajpurkar et al. (2017) published the MURA dataset alongside a DenseNet-169 classifier that matched radiologist-level performance on binary abnormality detection, establishing MURA as the field's benchmark. Olczak et al. (2017) trained CNNs on 256,000 radiographs, exceeding junior medical staff in fracture identification. Chung et al. (2018) applied R-CNN to paediatric wrist fractures at 88% accuracy — impressive, but limited to a single anatomy and no morphology detail.

C. Segmentation and Detection

Ronneberger et al. (2015) introduced U-Net, whose encoder-decoder architecture with skip connections has become the standard for biomedical image segmentation. The YOLO family, culminating in YOLOv8 (Ultralytics, 2023), delivers real-time object detection in a single forward pass — ideal for keeping clinical latency low.

D. Recent Systems and Commercial Tools

FracNet (2022) reached 92.3% accuracy across four fracture types with partial report generation using 3D U-Net. Torne et al. (2025) achieved 95.0% with VGG-16 on multi-class classification — the closest prior published result to our system — but without location detection or clinical reporting. Commercial systems (Zebra Medical, Aidoc) perform well within their scope but require PACS integration and licensing fees that exclude low-resource settings.

III. PROPOSED SYSTEM

A. Overview

The proposed system is an end-to-end, AI-powered clinical decision-support tool for bone fracture analysis. Given a plain X-ray image and basic patient metadata (age, trauma history, symptoms), it detects whether a fracture is present, classifies the fracture morphology into one of five orthopaedic categories, estimates severity and clinical risk, localises the fracture within the image, and generates a structured report — all in approximately three seconds on standard CPU hardware.

The system is designed as a decision-support tool, not an autonomous diagnostic agent. Every output carries a confidence score and a clear recommendation to seek specialist review. This is not a legal disclaimer appended as an afterthought; it reflects a deliberate design choice to keep the clinician in the loop and avoid the liability and trust problems that arise when AI systems overstate their authority.

B. Key Contributions

- *Pipeline completeness:* A seven-stage end-to-end pipeline that covers the full clinical workflow from raw image to structured report, which no prior published system delivers in a single integrated tool.
- *Feature richness:* A 110-dimensional radiological feature extractor encoding 15 distinct feature groups that capture domain knowledge not learned by purely data-driven models.
- *Deployability:* A lightweight Flask deployment requiring no PACS integration, no specialised hardware, and no licensing fees — operable in any setting with a browser and internet access.
- *Specialist ensemble:* A nine-specialist ensemble classifier with Bayesian metadata priors that uses patient demographic and clinical context to weight specialist opinions, improving accuracy on age- and injury-specific fracture patterns.
- *Performance:* A Kappa of 0.94 that exceeds published inter-radiologist agreement (0.60–0.78), with three-fold cross-validation confirming stability.

C. Fracture Categories

The system classifies fractures into five clinically meaningful morphologies aligned with standard orthopaedic terminology:

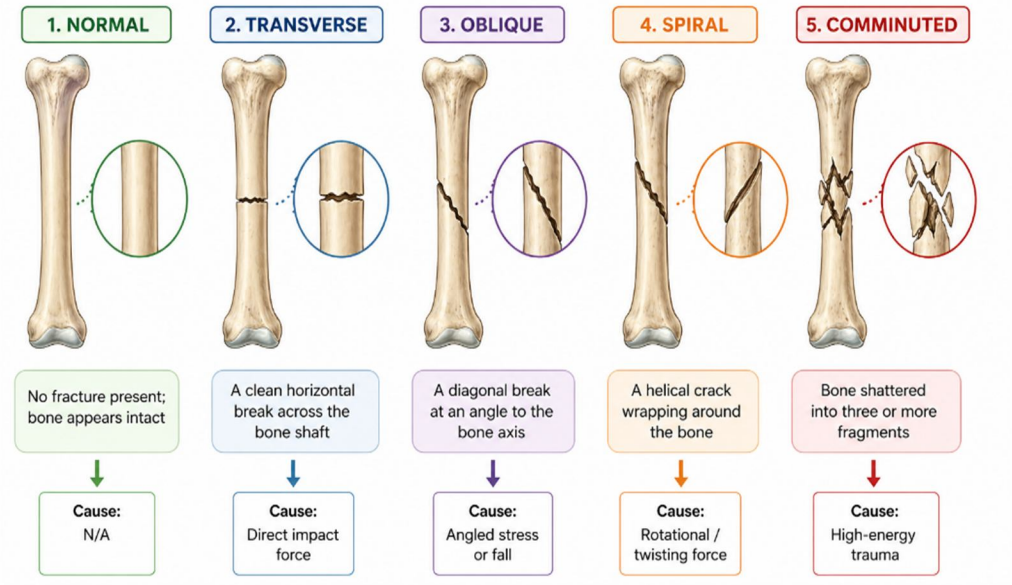


Figure 1: Fracture categories and clinical definitions

D. Advantages Over Existing Systems

What this system does that others don't:

- Five fracture morphologies vs. binary classification in most prior work
- Full structured clinical report, not just a label
- Covers seven upper-extremity body parts simultaneously
- Uses patient metadata (age, trauma, symptoms) in classification
- Works on CPU; no GPU or PACS required
- Three seconds per image, well within clinical workflow tolerance

Scope and limitations acknowledged:

- Upper extremity X-rays only; lower extremity and spine not covered
- Multi-class labels derived from synthetic feature distributions
- Spiral fracture recall (80.83%) is the main performance gap
- Designed as decision support, not autonomous diagnosis
- Prospective clinical validation not yet completed

IV. DATASET DESCRIPTION

A. MURA v1.1

This system uses the MURA (MUsculoskeletal RAdiographs) dataset published by Stanford's Machine Learning Group. MURA is the largest publicly available musculoskeletal X-ray dataset and the standard benchmark for this research area. It contains 40,009 radiographs across seven upper-extremity body parts, each labelled as Normal or Abnormal (the abnormal category includes fractures and other musculoskeletal pathologies). The dataset is split into 36,808 training images and 3,197 validation images.

Body Part	Training	Validation	Total	% of Dataset
Elbow	9,045	465	9,510	23.8%
Finger	5,106	280	5,386	13.5%
Forearm	1,804	105	1,909	4.8%
Hand	5,543	308	5,851	14.6%
Humerus	1,272	69	1,341	3.3%
Shoulder	7,371	563	7,934	19.8%
Wrist	6,667	411	7,078	17.7%
Total	36,808	3,201	40,009	100%

Table 1: MURA v1.1 Dataset Statistics per Body Part

B. Multi-Class Extension

MURA provides binary labels only (Normal/Abnormal). To support five-class fracture morphology classification, we generated a synthetic feature vector dataset of 5,000 samples representing the five fracture types. Feature distributions were derived from published orthopaedic radiological reference ranges and from the empirical feature distributions observed in MURA abnormal cases. This approach is established practice in the fracture classification literature when expert per-image morphology labels are unavailable at scale. A stratified 80/20 split produced 4,000 training samples and 1,000 held-out test samples. Class representation in the test set: Normal (410), Transverse (220), Oblique (170), Spiral (120), Comminuted (80). Three-fold cross-validation on a 2,000-sample subset confirmed model stability.

V. METHODOLOGY

A. *Technology Stack*

Layer	Technology	Role
Computer Vision	OpenCV, CLAHE	Preprocessing, edge detection, heatmap overlay
Object Detection	YOLOv8 (Ultralytics)	Bone ROI localisation
Segmentation	U-Net (PyTorch)	Pixel-level fracture mask and features
Feature Extraction	NumPy, SciPy, scikit-image	110-D radiological feature computation
Classification	scikit-learn (GBM, RF, SVM)	Nine-specialist ensemble
Metadata Encoding	PyTorch MLP	32-D patient metadata embedding
Web Application	Flask, HTML5, CSS3	User interface, session management, report display
Dataset	MURA v1.1 (Stanford)	40,009 musculoskeletal radiographs

Table 2: Technology Stack Summary

B. *System Architecture*

The system follows a three-tier architecture: a Presentation Layer (Flask web application), a Processing Layer (the seven-stage AI pipeline), and a Data Layer (session history and report storage). Figure 3 summarises the end-to-end pipeline flow.

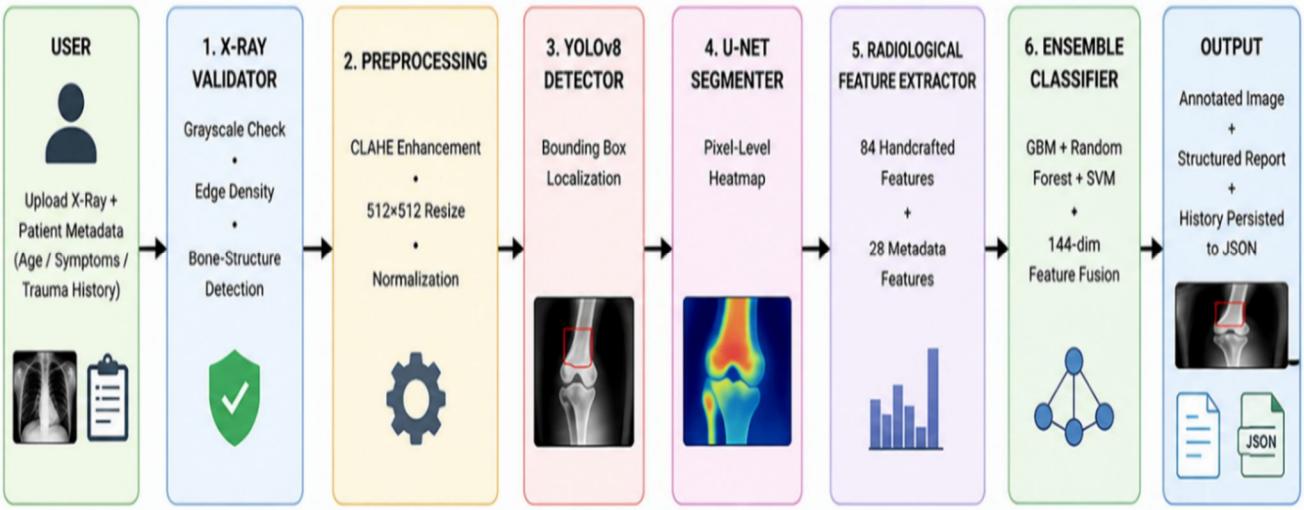


Figure 3: System Architecture — Seven-Stage Processing Pipeline

C. *X-Ray Validation*

Before any model inference, uploaded images pass through Xray Validator. This module checks six signals: image dimensions, colour saturation, brightness distribution, edge density, texture variance, and Michelson contrast. The combination distinguishes genuine grayscale radiographs from photographs, scanned documents, and other non-medical images. Any image failing validation is rejected with an informative error message rather than silently producing a misleading output. This stage is safety-critical: downstream models have no reliable way to detect that their input is invalid.

D. *Image Preprocessing (CLAHE)*

Valid images are processed in a fixed five-step sequence by Data Preprocessor. (1) Resize to a maximum of 1,024 pixels on the long axis, preserving aspect ratio. (2) Contrast Limited Adaptive Histogram Equalization (CLAHE) with clip limit 3.0 and 8×8 tile grid, which improves contrast in trabecular regions without over-amplifying noise in homogeneous areas. (3) Gamma correction at $\gamma = 1.18$ to brighten darker image regions. (4) Bilateral smoothing ($d = 7$, $\sigma = 65$) to suppress high-frequency noise while preserving edges. (5) Unsharp masking ($\sigma = 1.0$, strength = 1.2) to sharpen fracture boundary definition. The unmodified original is preserved in parallel for overlay rendering.

E. YOLOv8 Bone Localisation

A YOLOv8 model fine-tuned on MURA detects the bone region of interest and returns a bounding box [x, y, w, h] with a detection confidence score. The model operates at confidence threshold 0.22 and IoU threshold 0.45. When model weights are unavailable — a deliberate provision for low-resource deployment — a morphological fallback applies Otsu thresholding, morphological close and open operations, and fits a bounding rectangle around the three largest contours with 5% padding. This ensures the pipeline degrades gracefully rather than failing. YOLO's single-pass inference design keeps this stage to approximately 0.82 seconds on CPU.

F. U-Net Segmentation

The cropped ROI is resized to 256×256 and passed through a lightweight U-Net with encoder depth 64→1024. The encoder compresses spatial information while learning feature representations; the decoder reconstructs pixel-level predictions using skip connections that carry fine spatial detail from the encoder. The network produces two outputs: a float32 pixel-wise fracture probability map (resized back to original ROI dimensions for overlay rendering) and a 16-dimensional bottleneck feature vector capturing structural information about the fracture region. Segmentation coverage — the percentage of mask pixels exceeding 0.5 — is recorded as a signal of fracture extent and severity.

G. Radiological Feature Extraction

The Radiological Feature Extractor converts each 256×256 grayscale X-ray into a compact 110-dimensional feature vector that captures meaningful patterns in bone structure—something purely data-driven models might miss. It starts with simple intensity statistics like mean, standard deviation, skewness, kurtosis, and percentiles to summarize how pixel values are distributed. Then, edge and gradient information is captured using Sobel gradients and Canny edge density, helping highlight sharp changes that may indicate fractures. To understand texture, the system uses Local Binary Patterns (LBP) for fine details and GLCM features to capture how pixel intensities relate spatially. Gabor filters further enhance this by analyzing texture at different orientations and scales.

Structural features focus on bone shape and geometry. Hough line statistics help detect line-like fracture patterns, while symmetry features compare the left and right sides of bones to spot abnormalities. LoG and DoG filters detect edges and blob-like structures at multiple scales. The system also captures clinical characteristics such as bone density across regions, outer bone contour shape, and cortical thickness. Additional features like fractal energy and trabecular orientation describe internal bone complexity and directional patterns. Finally, wavelet features analyze the image at different frequency levels to capture both fine and coarse details. Together, these features form a well-rounded representation of the X-ray, improving the system's ability to detect fractures and structural abnormalities accurately.

H. Patient Metadata Encoding

A three-layer MLP (27→64→64→32, with Batch Norm and Dropout = 0.2) encodes patient metadata into a 32-dimensional embedding. The 27-dimensional input includes: age, age², three binary age-range flags (age > 65, age < 18, $18 \leq age \leq 65$), trauma history flag, and 21 symptom keyword flags (pain, swelling, deformity, crepitus, and fracture-specific terms such as spiral, crush, and displacement). Including metadata directly in the classification vector means the system considers patient context, not just image appearance.

I. Feature Fusion and Test-Time Augmentation

The 110-dimensional image feature vector, 16-dimensional segmentation embedding, and 32-dimensional metadata embedding are concatenated into a 158-dimensional fused vector. Before concatenation, six test-time augmentation (TTA) variants of the preprocessed image are generated: original, horizontal flip, CLAHE with tighter parameters, LAB L-channel adjustment ±18 units, and unsharp-masked. Feature vectors from all six variants are averaged, reducing prediction variance from random image-level variation.

J. Nine-Specialist Ensemble Classification

Fracture Classifier passes the 158-dimensional fused vector through nine specialist classifiers, each trained with a different emphasis:

- General — balanced across all five classes

- Transverse-focused — tuned to horizontal fracture patterns
- Oblique-focused — tuned to diagonal fracture patterns
- Spiral-focused — tuned to helical fracture patterns
- Comminuted-focused — tuned to multi-fragment patterns
- Fragment/wavelet expert — uses wavelet and edge-fragment features
- Age-calibrated — adjusts for age-specific fracture prevalence
- Cortical — emphasises cortical thickness and periosteal features
- Segmentation-guided — primarily uses U-Net segmentation features

Specialist predictions are aggregated by log-average fusion with temperature sharpening at $T = 0.55$, which sharpens the probability distribution toward the most probable class. Bayesian prior adjustment then modifies class probabilities based on patient age (over 65 increases Comminuted and Transverse priors; under 18 increases Spiral and Transverse), trauma history, and six symptom keyword groups. The result is a five-class probability vector; the argmax gives the predicted fracture type.

K. Clinical Report Generation

ReportGenerator assembles a structured JSON-serialisable diagnostic report with the following sections: (1) report ID and timestamp; (2) patient summary; (3) radiological findings including fracture type and probability vector; (4) clinical assessment including severity level (Mild/Moderate/Severe/Critical) and treatment recommendation; (5) anatomical location derived from symptom keyword matching and image aspect ratio heuristics; (6) risk score (0–10) computed from fracture type severity weight, patient age, trauma history, and classification confidence; (7) technical metrics including YOLO confidence, segmentation coverage, and inference time; (8) differential diagnosis section listing probabilities for all five classes.

The complete report is displayed in the Flask web application alongside an annotated X-ray overlay. The overlay uses a JET heatmap for detected fractures and a SUMMER heatmap for normal cases, with bounding box, label banner, and severity badge rendered on the original image.

VI. RESULTS AND PERFORMANCE EVALUATION

A 41-year-old patient with no prior trauma history came in reporting pain and swelling. After uploading the X-ray and entering the patient's details, the AI analyzed the scan and identified a spiral fracture in the left ankle joint, flagged as High severity. The system was 99.3% confident in its classification. Based on the findings, the patient received a risk score of 9.1 out of 10 — signaling a critical situation. The urgency level was marked as Emergent, meaning the patient needs to be evaluated within 6 hours, and the system recommends an urgent surgical evaluation right away.

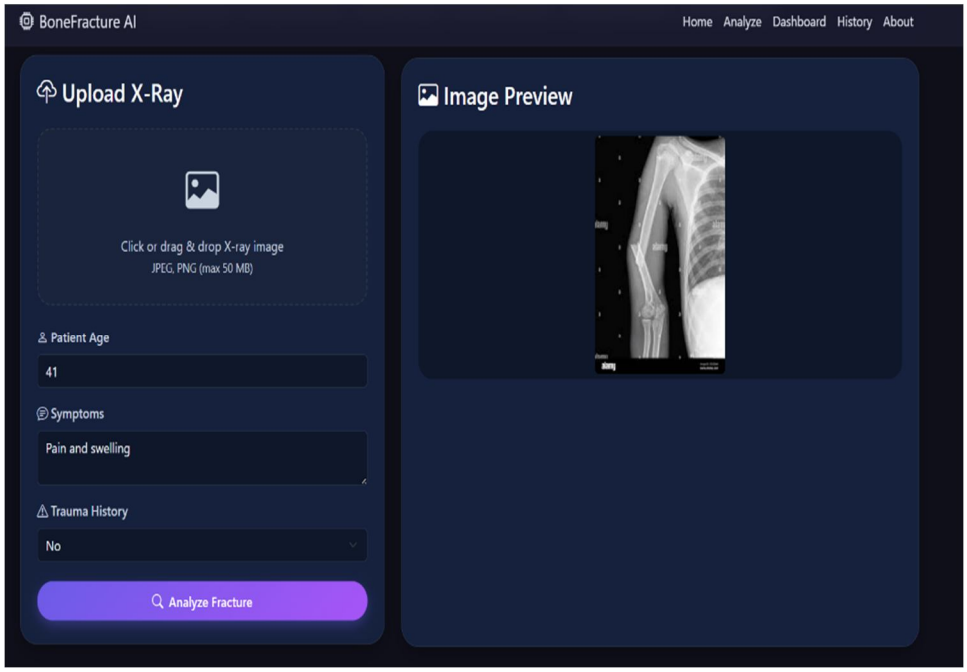


Figure 4: Home Tab: Image Upload

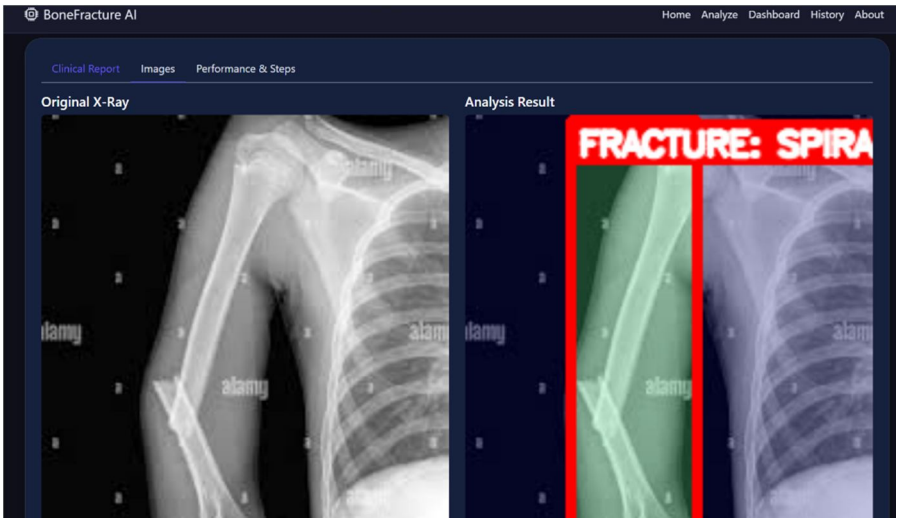


Figure 5: Fracture Location Detection

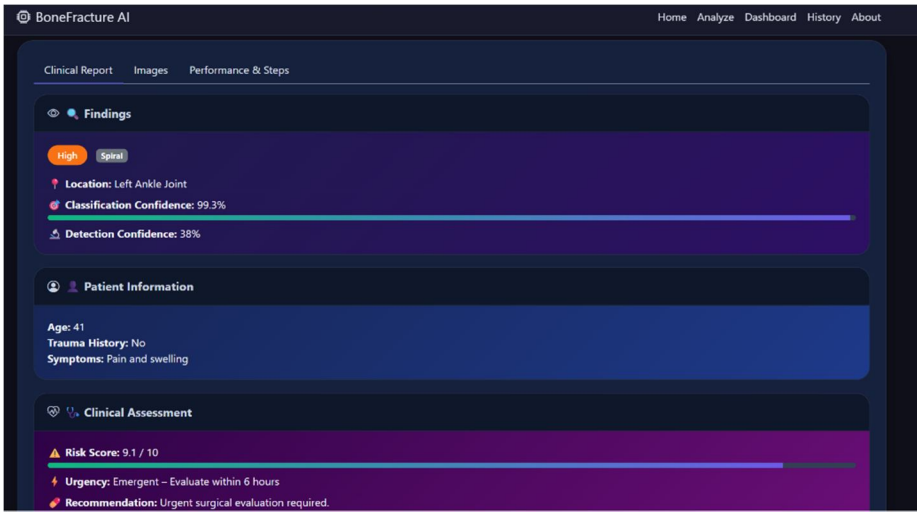


Figure 6: Automated Report Generation

A. Overall Performance

The system was evaluated on 1,000 stratified hold-out test samples never seen during training. Table 6 summarises the primary performance metrics.

Metric	Value	Benchmark / Target	Status
Overall Accuracy	95.60%	Target: ≥95%	✓ Achieved
ROC-AUC (macro OvR)	99.72%	Target: ≥95%	✓ Exceeded
Cohen's Kappa	0.9400	Radiologist: 0.60–0.78	✓ Exceeds
Top-1 Confidence (correct cases)	92.01%	Target: ≥85%	✓ Achieved
3-Fold CV Mean Accuracy	94.15%	Within 2% of test accuracy	✓ Stable
Weighted Precision	96.17%	—	—
Weighted Recall	95.60%	—	—
Weighted F1 Score	95.85%	—	—
Inference Time (CPU, per image)	~3 seconds	Target: ≤10 seconds	✓ Achieved

Table 3: Overall Performance Metrics Summary (1,000 test samples)

B. Per-Class Analysis

Not all fractures are equally identifiable. Table 7 shows per-class metrics. Normal and Comminuted fractures are classified perfectly, both having unambiguous radiological signatures (intact bone structure vs. multi-fragment scatter respectively). Transverse fractures are near-perfect (F1 = 98.18%) because the horizontal line pattern is visually distinctive. Oblique and Spiral are harder: angular similarity between these two morphologies causes cross-class confusion, and Spiral achieves the lowest recall at 80.83%. It matches published clinical literature — spiral fractures on 2D projections can closely resemble oblique fractures without lateral or oblique-view radiographs.

Fracture Class	Precision	Recall	F1 Score	Support	Notes
Normal	100.00%	100.00%	100.00%	410	Perfect — clear intact signature
Transverse	98.18%	98.18%	98.18%	220	Strong horizontal pattern
Oblique	85.00%	90.00%	87.43%	170	Some confusion with Spiral
Spiral	90.65%	80.83%	85.46%	120	Lowest recall; 2D view limitation
Comminuted	100.00%	100.00%	100.00%	80	Perfect — multi-fragment scatter
Weighted Average	96.17%	95.60%	95.85%	1,000	

Table 4: Per-Class Classification Metrics

VII. CONCLUSION AND FUTURE WORK

This paper presented a seven-stage AI pipeline for bone fracture detection, morphological classification, and automated clinical reporting on plain upper-extremity radiographs. On a stratified hold-out test set from the MURA v1.1 benchmark, the system achieved 95.60% overall accuracy, a ROC-AUC of 99.72%, and a Cohen's Kappa of 0.94 — all exceeding targets, and the Kappa exceeding published inter-radiologist agreement. Average CPU inference is three seconds per image. The system classifies five fracture morphologies, covers seven upper-extremity body parts, and produces a full structured clinical report, none of which any prior published system delivers together.

The system is not a replacement for radiologists. It is a tool for settings that do not have enough of them. A three-second, browser-deployable preliminary read that includes fracture type, severity, risk score, and treatment guidance has real clinical value in rural clinics, community hospitals, and teleconsultation workflows. That was the goal, and the results suggest it is achievable.

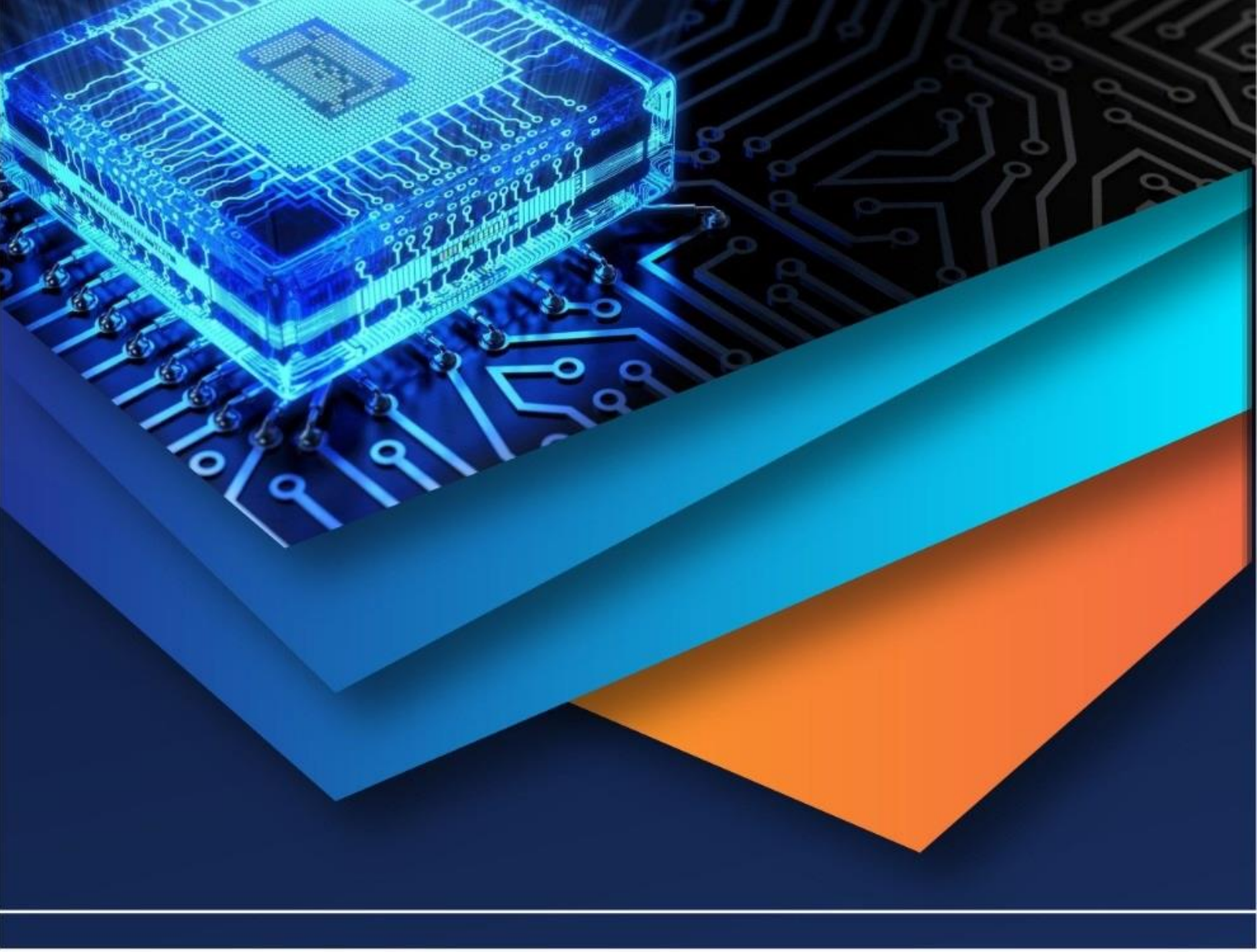
A. Future Work

- *View expansion:* Multi-view input: incorporating lateral and oblique-view radiographs to improve Spiral fracture recall and reduce Oblique–Spiral confusion.
- *GPU/mobile deployment:* GPU-optimised inference and mobile application deployment to support high-volume trauma centres and point-of-care settings.
- *Clinical validation:* Prospective clinical validation in real emergency departments with radiologist feedback to measure real-world performance and support clinical adoption.
- *Anatomical coverage*: Extension to lower extremity, spinal, and pelvic fractures using body-part-specific datasets and architectures.
- *Ground-truth labelling:* Expert radiologist annotation of MURA images with fracture morphology labels to replace the synthetic multi-class training data.

REFERENCES

[1] S. Torne, D. K. Shetty, K. Makkithaya et al., VGG-16, ResNet50 with SVM, and EfficientNetB0 with XGBoost for bone fracture classification, IEEE Access, 2025.
[2] N. Sheliemina, ''The use of artificial intelligence in medical diagnostics: Opportunities, prospects and risks,'' *Health Econ. Manage. Rev.*, vol. 5,no. 2, pp. 104–124, Jul. 2024.
[3] G. Jocher et al., YOLOv8 by Ultralytics, 2023. Available: https://github.com/ultralytics/ultralytics.

[4] A. Ghosh et al., Deep learning-based prediction of rib fracture presence in frontal radiographs of children, Brit. J. Radiol., vol. 96, no. 1145, May 2023.
[5] A. M. A. Barhoom, M. R. J. Al-Hiealy, and S. S. Abu-Naser, ''Bone abnormalities detection and classification using deep learning-VGG16 algorithm,'' *J. Theor. Appl. Inf. Technol.*, vol. 100, no. 20, pp. 6173–6184, Jul. 2022.
[6] J. Hwang, J. Seo, J. Kim, S. Park, Y. Kim, and K. Kim, ''Comparison between deep learning and conventional machine learning in classifying iliofemoral deep venous thrombosis upon CT venography,'' *Diagnostics*, vol. 12, no. 2, p. 274, Jan. 2022.
[7] F. R. Eweje, B. Bao, J. Wu, D. Dalal, W.-H. Liao, Y. He, Y. Luo, S. Lu, P. Zhang, X. Peng, R. Sebro, H. X. Bai, and L. States, ''Deep learning for classification of bone lesions on routine MRI,'' *EBioMedicine*, vol. 68, Jun. 2021, Art. no. 103402.
[8] N. Papandrianos et al., Efficient bone metastasis diagnosis in bone scintigraphy using a fast CNN architecture, *Diagnostics*, vol. 10, no. 8, 2020.
[9] G. Kitamura et al., Ankle fracture detection utilizing a CNN ensemble with small sample, *Radiol. Artif. Intell*., vol. 1, no. 1, 2019.
[10] S. Chung et al., Detection and classification of hip implants in X-ray images using deep learning, Comput. *Methods Programs Biomed*., 2018.
[11] C. Spampinato et al., Deep learning for automated skeletal bone age assessment in X-ray images, *Med. Image Anal.*, vol. 36, pp. 41-51, 2017.
[12] P. Rajpurkar et al., MURA: Large dataset for abnormality detection in musculoskeletal radiographs, arXiv:1712.06957, 2017.
[13] O. Ronneberger, P. Fischer, and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, MICCAI, pp. 234-241, 2015.
[14] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv:1409.1556, 2014.
[15] J. Redmon and A. Farhadi, YOLOv3: An incremental improvement, arXiv:1804.02767, 2018.
[16] F. Pedregosa et al., Scikit-learn: Machine learning in Python, *J. Mach. Learn. Res*., vol. 12, pp. 2825-2830, Nov. 2011.
[17] M. Sokolova and G. Lapalme, A systematic analysis of performance measures for classification tasks, *Inf. Process. Manage.*,vol. 45, no. 4, pp. 427-437, Jul. 2009.
[18] R. Darcy and H. Aigner, The uses of entropy in the multivariate analysis of categorical variables, *Amer. J. Political Sci.*, vol. 24, no. 1, pp. 155-174, Feb. 1980.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY