Vol. 06, No. 01, pp. 36 –42 (2025)
ISSN: 2708-0757

JASTT

JOURNAL OF APPLIED SCIENCE AND TECHNOLOGY TRENDS

www.jastt.org

Real-Time Bone Fracture Detection Using MobileNetV2 and Explainable AI for Clinical Integration

Tanmay Paradava, Rituraj Jain*, Harsh Patania, Tirth Patel

Department of Information Technology, Marwadi University, Rajkot, Gujarat, India, tanmayparadava1452@gmail.com, jainrituraj@yahoo.com, harshpatania2907@gmail.com, tirthpatel0313@gmail.com
**Correspondence: jainrituraj@yahoo.com*

Abstract

Bone fractures are among the most frequent injuries requiring immediate diagnosis, yet traditional X-ray analysis is time-consuming and reliant on expert interpretation. As medical AI advances, there is an increasing requirement in terms of effective and implementable diagnostic tools. The purpose of the study is to create a real-time, clinically practical system to detect a fracture combining lightweight deep learning, interpretability, and system-level integration. A convolutional neural network with MobileNetV2 architecture was trained on a stratified dataset of the elbow X-ray images, which have been divided into three classes: normal, hairline, and displaced fractures. Generalization and explainability were performed with data augmentation, two-phase fine-tuning, and Grad-CAM. This model had an accuracy of 89.26 percent, a precision of 91.52 percent, F1 score of 89.04 percent and a minimum false negative of 14 cases out of 1018 cases. The system is delivered using Docker on the AWS EC2 and available as a web interface implemented using Flask, which provides an opportunity to apply it in distant clinical facilities. The suggested pipeline merges both deep learning research and clinical practice domains because it provides a system allowing one to detect bone fractures quickly, interpretably, and scale up, and is the first of its kind to provide an entity that is accurate, can be used in real-time, and be deployable end-to-end.

Keywords: *bone fracture detection, deep learning, explainable ai, medical image analysis, mobilenetv2*
Received: May 6th, 2025 / Accepted: June 18th, 2025 / Online: June 19th, 2025

I. INTRODUCTION

One of the most common body injuries across the globe is that of bone fractures, which immediately require diagnosis and treatment to avoid complications. The existing markets for fracture detection include standard X-ray and CT scan, which often demand great expertise and are time-consuming. With advances in medical imaging technologies, deep learning algorithms are beginning to emerge as viable systems to automate and expedite the process of fracture detection. This work is concerned with the development of an automated fracture detection system for bone images based on deep learning that will adopt a MobileNetV2 architecture. The model would improve its accuracy and robustness through data augmentation, evaluation metrics, and regularization techniques. The model will be built from deep learning by training on a dataset of labeled bone X-ray images so that important identification and classification of fractures may help healthcare professionals by reducing the time and effort for diagnosis. This system describes the building, training, and evaluation processes of the deep learning model, the methods to ensure high performance and generalization across many fracture scenarios, and the envisaged clinical application of the system and its consequent effects on the health system worldwide. An intelligent deep learning model would be developed in such a way that this would automatically detect fractures on an X-ray image with very high accuracy and reliability. The objective the proposed system are:

1) *To design and train a convolutional neural network employing the MobileNetV2 architecture.*
2) *To improve the generalization power of the model with the use of advanced data augmentation techniques.*
3) *To analyze the decision-making process of the model using Class Activation Mapping (CAM).*
4) *To test real images and produce diagnosis output.*
5) *To create an end-user system that is capable of use for clinical decision support.*

In contrast to existing approaches, this study focuses not only on classification accuracy but also on the deployment of a

doi: 10.38094/jastt61236

ipAcademia www.ipacademia.org

lightweight, interpretable, and real-time system optimized for clinical settings.

II. Literature Review

Automated bone fracture detection using medical imaging has garnered significant interest, with machine learning (ML) and deep learning (DL) methods increasingly employed to support clinical diagnostics. Early classical ML approaches, such as the study by [1], showed encouraging results on X-ray images using linear discriminant analysis (LDA). This technique integrated multiple image processing methods like Canny and Sobel edge detection, along with Hough line feature extraction, and was evaluated across several classifiers with hyperparameter tuning and cross-validation. However, the study lacked insights into clinical interpretability and the use of diverse datasets.

With the advancement of DL, newer methods have offered enhanced capabilities, though results vary depending on dataset composition and imaging modalities. In [2] authors utilized a combination of YOLACT++ segmentation and YOLOv4 detection on an augmented arm bone X-ray dataset, outperforming some existing models. Still, the small dataset used limits generalizability, underscoring the need for validation on larger, multicentric datasets. Authors in [3] conducted a systematic review highlighting DL's potential in pediatric fracture detection, particularly for elbow injuries. Their analysis emphasized the predominance of single-institution studies, pointing to a need for broader and more diverse data to enhance model robustness. In the domain of CT imaging, authors of [4] investigated 3D CNNs for pelvic fracture detection but faced challenges in accurately modeling complex bone structures. Alternatively, Weikert et al. demonstrated the potential of DL for rib fracture detection in whole-body trauma scans, achieving notable specificity and sensitivity, though still with limitations in clinical coverage.

For skull fractures, authors of [5] compared YOLOv3 and attention-based U-Net segmentation models. While both techniques demonstrated effective localization, they did not yet align with ideal benchmarks expected for clinical deployment. More recent efforts have aimed at improving detection sensitivity and reducing false positives. In [6] authors introduced a multi-stage 3D DL algorithm for rib fractures that showed enhanced detection with refinement, though it still requires broader validation. In X-ray imaging, author applied YOLO for rapid fracture localization [7], while authors used a fusion of feature and image pyramid networks to improve detection accuracy in thighbone fractures [8]. In [9] authors addressed the detection of rare acetabular fractures using a DCNN, though their work was constrained by limited data availability.

Several other researchers [10-12] employed CNN-based models such as AlexNet, DenseNet, and VGG19 across larger datasets, achieving superior performance compared to traditional ML methods. These works demonstrate the evolving accuracy and efficiency of DL models in medical image classification. In [13], researchers provided a comprehensive review on hip fracture detection, highlighting the advantage of AI-human collaboration in enhancing diagnostic reliability. In [14], researchers explored integrated learning techniques on the MURA dataset with promising classification outcomes, though real-world deployment remains to be explored further. Moreover, researchers in [15] focused on vertebral fracture detection using DL, reporting performance comparable to clinical experts, while Hsieh et al. extended DL's application to predict fracture risks related to bone mineral density in osteoporosis.

Even though other researches already addressed diverse deep learning models in fracture detection, the present research proposal contributes to the research field by a combination of a light-weight model (MobileNetV2) with a explainable AI framework (Grad-CAM) trained according to a two-phase fine-tuning approach and implemented as a real-time web interface. Our method, unlike the traditional ones, is a full-end-to-end pipeline, inclusive of data preprocessing, deployment on the cloud to Flask, Docker and AWS EC2. The suggested system focuses not only on the accuracy of classification but also the usability, transparency, and scalability of such an idea to the real world clinical environment with a focus on resource limited conditions.

Overall, AI-driven bone fracture detection shows substantial promise but challenges remain in dataset diversity, clinical interpretability, minimizing false positives, and validating models across multiple institutions and imaging modalities to ensure real-world applicability. Table I summarizes the key research gaps identified in the domain of automated bone fracture detection and outlines the specific contributions of the proposed framework in addressing each of these challenges. By leveraging a lightweight MobileNetV2 backbone, extensive data augmentation, transfer learning, and explainability via Grad-CAM, the system not only enhances detection performance but also ensures real-time usability and clinical integration through a web-based interface and full-stack deployment strategy.

TABLE I. Summary of Research Gaps and Corresponding Solutions Proposed in This Study

Research Gap Identified	Impact on Prior Studies	How This Research Addresses It
Lack of lightweight, deployable models for real-time detection	Limited deployment in clinical or low-resource settings	Uses MobileNetV2, a lightweight and fast model suitable for low-resource environments
Limited dataset diversity and poor generalization	Overfitting and low robustness in varied clinical scenarios	Applies data augmentation and transfer learning to improve generalization across classes
High false negative rates in clinical diagnostic tools	Missed fracture cases, risking delayed diagnosis	Achieves very low false negatives (14/1018), enhancing diagnostic reliability
Absence of user-friendly interfaces for clinicians	Reduced clinical trust in model decisions	Provides a responsive web-based interface with image upload, preview, and real-time output
Lack of explainability in predictions	Models remained inaccessible to clinicians	Integrates Grad-CAM for visual feedback on model attention, enhancing interpretability
Lack of end-to-end deployment for clinical adoption	Gaps in practical application and scalability	Implements complete pipeline with Flask, Docker, and AWS EC2 for scalable, remote clinical use

III. METHODOLOGY

The proposed system for automated bone fracture detection integrates deep learning with a lightweight MobileNetV2 model and a user-centric web interface to provide real-time diagnostic support from X-ray images. This section outlines the end-to-end methodology, including dataset preparation, preprocessing, model training, hyperparameter tuning, system architecture, and deployment.

A. Dataset Description

The dataset used in this study was sourced from a publicly available repository which contains X-ray images of multiple anatomical regions, including the elbow, hand, and shoulder. This was further discerned into three viable clinical groups of normal, hairline and displaced fractures. The method of stratified sampling was utilized with the following division of the data partition into the training (70%), validation (15%), and testing (15%) sets, where each of them run through the defined number of classes in a balanced distribution. To improve model generalization and to avoid overfitting, a complex augmentation pipeline was used, including rotations (±20°), zooming (up to 20%), horizontal and vertical shifts (±20%), brightness variation (±20%) and horizontal/vertical flipping. These distortions simulate randomness, which is typically present in healthcare radiography and enhance the generalizations of the model in practice.

B. Preprocessing Pipeline

All input images underwent standardized preprocessing prior to model training. Images were resized to 224×224 pixels to comply with MobileNetV2's input specifications. Pixel values were normalized to a [0, 1] range for improved training convergence. Channels were formatted in RGB to align with the pre-trained network expectations. Moreover, stratified sampling ensured class-balanced distributions across training, validation, and test sets, promoting unbiased learning and evaluation.

C. System Architecture

As depicted in Fig. 1, the system comprises four primary modules: the user interface, preprocessing pipeline, deep learning inference engine, and result display unit. The client-side interface, built using HTML, CSS, and JavaScript, allows users to upload X-ray images through a drag-and-drop interface or manual selection. These images are processed via a Flask backend, which executes preprocessing, invokes the trained MobileNetV2 model for inference, and returns the prediction (fracture or no fracture) to the frontend dynamically without reloading the page. This modular design ensures efficient operation, low-latency feedback, and intuitive clinical usability.

D. Model Description and Training Strategy

In this study, we used MobileNetV2 as the backbone architecture, a well-known lightweight convolutional neural network which is efficient and highly accurate at the same time in low resource environments. In order to have the same input dimensions of ImageNet pretrained models which are (224 × 224 × 3), the model was initialized with this dimension, as summarized in Table II. This was set to 0.75 to decrease model complexity while maintaining a little drop in the feature extraction skills. It was removed the top layers of classification (include_top = False) in order to receive a custom head classification for the binary fracture detection task. ImageNet Pretrained weights were used to leverage the transfer learning to allow extraction of robust and generalized image representations.

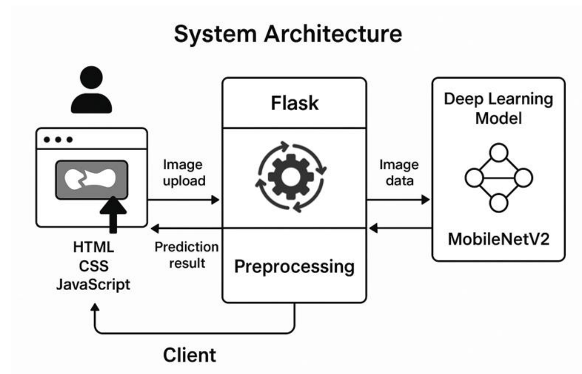


Fig. 1. End-to-End System Architecture for Real-Time Bone Fracture Detection Using MobileNetV2

Two separate phases of model training were optimized for performance and generalization. First, the base MobileNetV2 layers were frozen and only the custom classification layers were trained which enabled the domain adaptation to be effective. Then in the second stage the last 30 layers of the backbone were unfrozen to fine tune the network to learn task specific features. The Adam optimizer was used for training with the learning rate of 0.0001 and the batch size of 64. Dropout layers with dropout rates of 0.7 and 0.5 were applied on various stages to avoid overfitting while the use of L2 regularization with a weight decay coefficient (λ = 0.01) was also employed. Manual search over learning rate schedules, batch size, dropout rates and regularization values were performed based on surrogate performance on the validation set. This systematic fine tuning approach ensured that the balance between the computational efficiency and accuracy of classification productions was optimal.

TABLE II. MOBILENETV2 BASE MODEL CONFIGURATION

Parameter	Value	Description
Input Shape	(224, 224, 3)	Standard ImageNet size
Alpha	0.75	Width multiplier (reduced from 1.0)
Include Top	False	Exclude final classification layers
Weights	ImageNet	Pre-trained weights
Trainable Layers	Last 30	Fine-tuning approach

E. Hyperparameter Optimization

In order to have optimal model performance while being efficient, a systematic hyperparameter tuning has been implemented. Then the learning rate was manually explored in the range [0.1, 0.01, 0.001, 0.0001] and 0.0001 was found to provide the best convergence and stability. Batch size experiments were studied with values of 32, 64, 128 and 64 was chosen as the best value that maintains a balance between model generalization and computational requirement. At these two important stages in the custom classification head, dropout regularization was applied at rates of 0.3, 0.5 and .0. A (validation) performance of 0.7 in the first dropout layer and 0.5

in the subsequent layer had better results. We also tested L2 regularizations with λ values of 0.1, 0.01 and 0.001, we found that a trade-off between mitigating overfitting and training loss stability is with $\lambda = 0.01$. Also, the fine tuning of the width multiplier (α), with MobileNetV2 was done for $\alpha \in [0.35, 0.5, 0.75, 1.0]$, where $\alpha = 0.75$ was found to theoretically provide the best balance between model expressiveness and classification accuracy. During model optimization, the search space of the hyperparameters explored is outlined in Table III, learning rate, batch size, dropout rates L2 regularization and alpha scaling. A specification is given for test ranges, optimal values chosen and tuning methods for each parameter. Through this systematic tuning, we ended up with a well-balanced configuration which simultaneously achieved maximum possible classification performance and kept the computational expense at bay.

TABLE III. HYPERPARAMETER SEARCH SPACE

Parameter	Values Tested	Optimal Value	Tuning Method
Learning Rate	[0.1, 0.01, 0.001, 0.0001]	0.0001	Manual Search
Batch Size	[32, 64, 128]	64	Resource Constraints
Dropout Rates	[0.3, 0.5, 0.7]	0.7 (first), 0.5 (second)	Validation Performance
L2 Regularization	[0.1, 0.01, 0.001]	0.01	Weight Analysis
Alpha Value	[0.35, 0.5, 0.75, 1.0]	0.75	Model Size/Accuracy Tradeoff

F. Training Management and Callback Optimization

To guarantee the proper convergence and the possibility of overfitting, a well-planned training strategy was followed with three important callbacks. The EarlyStopping was set up with a patience of 5 epochs which terminated the training once the validation performance had stagnated to avoid having to spend unnecessary computations. The ReduceLROnPlateau was applied to automatically scale the learning rate with a factor of 0.2 in case no improvement was noted out of the last 3 epochs in the training, thus leaving the model to adjust its weights better. ModelCheckpoint was applied in order to track validation accuracy and store the model weights at the point when the best performance was achieved. The optimizer used to train the model was the Adam type and the model could train up to 20 epochs although with the EarlyStopping feature, it usually converged within 14th to 17th epochs.

G. Implementation and Deployment

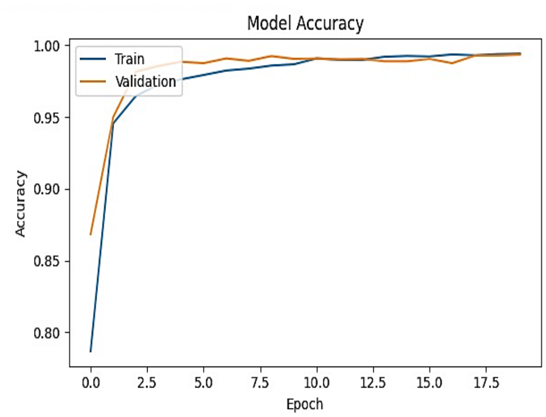
Upon completion of training, the final model was encapsulated in a production-ready system using Flask for backend integration. The user-facing web interface not only supports real-time image upload and diagnosis but also incorporates Grad-CAM visualizations for interpretability and transparency in predictions. For seamless deployment, the entire system was containerized using Docker and hosted on a single AWS EC2 instance (t2.medium, 2 vCPUs, 4 GB RAM). This lightweight setup eliminates the need for GPU acceleration, owing to the efficiency of the MobileNetV2 architecture. The deployment ensures cross-platform accessibility, scalability, and remote usability—making the system well-suited for use in hospital settings, rural clinics, and telemedicine environments.

IV. RESULT AND DISCUSSION

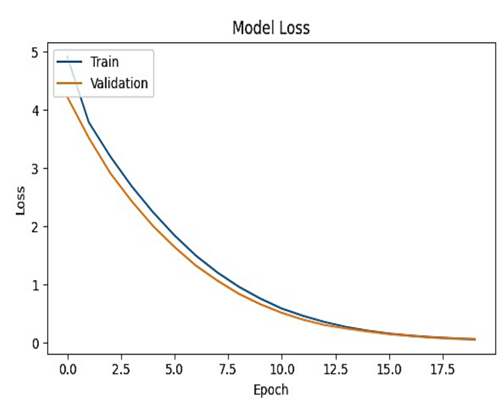
The proposed bone fracture detection system, based on a fine-tuned MobileNetV2 model, demonstrated robust classification performance, high usability, and readiness for clinical deployment. Evaluation was conducted using standard metrics, confusion matrix analysis, training dynamics, and web interface testing.

A. Training and Validation Trends

Fig. 2 illustrates the training and validation curves across 20 epochs. The model exhibited rapid learning during the early epochs, with accuracy stabilizing around epoch 10. Loss curves show consistent decline in both training and validation sets, indicating effective convergence without overfitting. The close alignment between the curves confirms that the transfer learning strategy enabled the model to generalize well on unseen X-ray data. Such consistency was obtained by the considered two-phase fine-tuning scheme: the original freezing of layers (on the first phase) was followed by the subsequent unfreezing the same final 30 layers at later stages. Regularization methods like dropout, L2 penalty and learning rate were also used as additional support. The stability of the performance proved that data augmentation and transfer learning provided good generalization on unseen data, which is very significant in medical imaging.



(a)



(b)

Fig. 2. Training and Validation curve over 20 epochs for (a) Accuracy (b) loss

B. Classification Performance

The model achieved an accuracy of 89.26%, an F1-score of 89.04%, precision of 91.52%, and recall of 89.26%, demonstrating strong reliability in binary classification of elbow X-rays. As shown in Fig. 3, the confusion matrix confirms high sensitivity and specificity, with only 2 false positives and 14 false negatives out of 2034 test images. This low false negative rate—approximately 0.69%—is especially important in clinical practice, where missed fracture cases can result in delayed or incorrect treatment. Analysis of misclassifications showed that most errors were associated with subtle hairline fractures, a known challenge even for experienced radiologists, further highlighting the value of model-assisted diagnostics.

Table IV compares the proposed model with other notable methods from the literature. Notably, our model achieved superior accuracy while being significantly more lightweight and suitable for deployment on low-resource or real-time platforms. This comparison underscores the practical value of our method in terms of the trade-off between performance and computational efficiency—an often overlooked aspect in earlier studies.

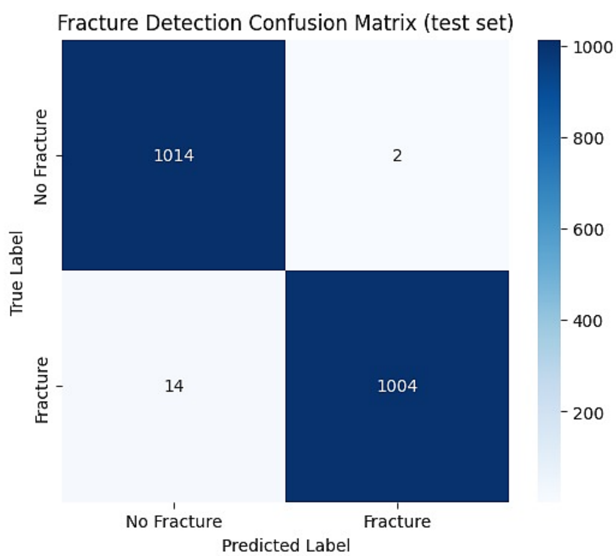


Fig. 3. Confusion matrix for the MobileNetV2-based fracture detection system.

TABLE IV. ACCURACY COMPARISON WITH SELECTED EXISTING FRACTURE DETECTION MODELS

Method - Study	Accuracy (%)
[1]	88.67
[2]	81.91
[4]	69.5
[8]	87.80
[9]	82.8
[15]	86
Proposed Method	89.26

The given method shows the higher accuracy of classification compared to a number of earlier works, and a substantial decrease of the computational overhead. Compared to systems like CNN-based 3D architectures [4] or ensembles of feature networks [8] whose training and obtaining outputs is resource-intensive, the current model is based on an efficient MobileNetV2 that supports real-time applications. The combination of the two-phase fine-tuning technique with sophisticated data augmentation techniques increase the robustness learned with respect to both domain-specific fracture features and by generalizing. Moreover, Grad-CAM integration provides a solution to a frequently missed deficiency in the previous research an inability to explain AI-based diagnostic outcomes. Such a union between model performance, interpretability and clinically deployable architecture (thanks to Flask and Docker) give rise to a harmonious framework that is not only precise but also usable and expandable advantages that have seldom been documented together in prior fracture detection research.

C. Web Interface and Real-Time Feedback

The system includes a browser-based interface, shown in Fig. 4, designed for intuitive interaction. Users can upload X-ray images via drag-and-drop or manual selection. The interface supports JPG, PNG, and JPEG formats and performs reliably across modern browsers including Chrome and Edge. Once uploaded, images are processed by a Flask backend, and predictions are displayed dynamically with no page reload. Fig. 5(a) shows output when a normal case is detected, and Fig. 5(b) shows detection of a fracture, both with real-time visual feedback. This interface supports clinician-friendly deployment and improves trust via interpretability.

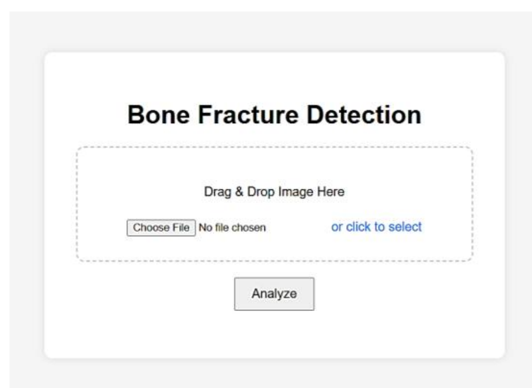
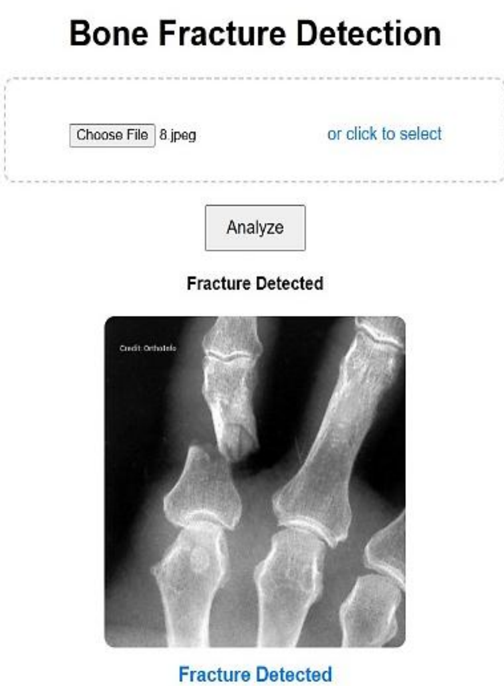


Fig. 4. User-friendly browser-based interface for uploading bone X-ray images

D. Deployment and Real-World Readiness

The model's lightweight architecture and low memory requirements enable deployment on edge devices or mobile platforms. The entire system is containerized using Docker and deployed on an AWS EC2 instance for remote access and scalability. These characteristics make the solution suitable for resource-limited settings, emergency care, and telemedicine applications. So, the proposed system integrates robust model performance with real-time usability and explainability, presenting a practical and scalable solution for automated bone fracture detection in clinical settings. Clinically, the proposed system offers significant advantages: a high level of diagnostic safety due to the low false negative rate, rapid inference for use in emergency care, and a user-friendly interface that requires minimal training. From a research standpoint, it addresses key gaps in previous literature by combining accuracy, explainability, and full-stack deploy ability in a single framework. Overall, the study demonstrates that AI-based diagnostic tools can be both powerful and practical when designed with real-world integration in mind.

(a)



(b)

Fig. 5. Real-time output showing prediction result (a) No Fracture (b) Fracture Detected

V. CONCLUSION

This study presents an end-to-end, lightweight, and explainable AI framework for automated bone fracture detection using elbow X-ray images. By leveraging the MobileNetV2 architecture, the model achieves a balance between computational efficiency and classification accuracy, making it well-suited for real-time diagnostic applications. The system was trained using a two-phase fine-tuning strategy with data augmentation and regularization to enhance generalization and reduce overfitting. It demonstrated strong performance, achieving 89.26% accuracy, 89.04% F1-score, and minimal false negatives—critical in clinical contexts where missed fractures can have serious implications. The inclusion of a browser-based interface with Grad-CAM visualizations enhances model transparency and supports clinician trust. The application is containerized and deployed on AWS, ensuring scalability for both urban and remote healthcare settings. Overall, the proposed solution bridges the gap between advanced deep learning research and practical clinical deployment, contributing to more efficient, accessible, and interpretable medical diagnostic workflows. Overall, the work distinguishes itself by not only demonstrating strong diagnostic performance but also addressing practical deployment needs, interpretability concerns, and clinical usability through a complete, lightweight, and real-time AI-driven framework—advancing beyond prior fracture detection models.

REFERENCES

[1] M. E. Sahin, “Image processing and machine learning-based bone fracture detection and classification using X-ray images,” *International Journal of Imaging Systems and Technology*, vol. 33, no. 3, pp. 853–865, Jan. 2023, doi: 10.1002/ima.22849.

[2] H. P. Nguyen, T. P. Hoang, and H. H. Nguyen, *A deep learning based fracture detection in arm bone X-ray images*. 2021. doi: 10.1109/mapr53640.2021.9585292.

[3] S. C. Shelmerdine, R. D. White, H. Liu, O. J. Arthurs, and N. J. Sebire, “Artificial intelligence for radiological paediatric fracture assessment: a systematic review,” *Insights Into Imaging*, vol. 13, no. 1, Jun. 2022, doi: 10.1186/s13244-022-01234-3.

[4] N. Yamamoto et al., An automated fracture detection from pelvic CT images with 3-D convolutional neural networks. 2020, pp. 1–6. doi: 10.1109/ccs49175.2020.9231453.

[5] W. Shan *et al.*, “Automated identification of skull fractures with Deep Learning: A comparison between object detection and segmentation approach,” *Frontiers in Neurology*, vol. 12, Oct. 2021, doi: 10.3389/fneur.2021.687931.

[6] A. Niiya *et al.*, “Development of an artificial intelligence-assisted computed tomography diagnosis technology for rib fracture and evaluation of its clinical usefulness,” *Scientific Reports*, vol. 12, no. 1, May 2022, doi: 10.1038/s41598-022-12453-5.

[7] B. Kalyani, “Human fracture detection using machine learning,” *International Journal for Research in Applied Science and Engineering Technology*, vol. 12, no. 3, pp. 3344–3350, Mar. 2024, doi: 10.22214/ijraset.2024.59613.

[8] B. Guan and G. Zhang, *Thighbone fracture detection based on fused deep learning method*. 2021, pp. 8066–8071. doi: 10.23919/ccc52363.2021.9550335.

[9] F. Erne *et al.*, “Deep learning in the detection of rare fractures – Development of a ‘Deep Learning Convolutional Network’ model for detecting acetabular fractures,” *Zeitschrift Für Orthopädie Und Unfallchirurgie*, vol. 161, no. 01, pp. 42–50, Jul. 2021, doi: 10.1055/a-1511-8595.

[10] L. Chawhan, M. M. Naik, R. Madhura, V. Ramya, K. J. Bhanushree, and H. R. Poorvitha, *Bone Fracture Detection and Classification using*

Machine Learning. 2024, pp. 1–8. doi: 10.1109/nmitcon62075.2024.10699036.

[11] S. Chauhan, *Bone Fracture Detection with CNN: A Deep Learning Approach*. 2024, pp. 1253–1258. doi: 10.1109/icosec61587.2024.10722699.

[12] B. Pujitha, K. R. Sravya, N. Krishnasai, and C. Aparna, *Detection of Bone Fracture Using Deep Learning*, vol. 5. 2024, pp. 703–708. doi: 10.1109/esic60604.2024.10481535.

[13] Y. Cha, J.-T. Kim, C.-H. Park, J.-W. Kim, S. Y. Lee, and J.-I. Yoo, “Artificial intelligence and machine learning on diagnosis and classification of hip fracture: systematic review,” *Journal of Orthopaedic Surgery and Research*, vol. 17, no. 1, Dec. 2022, doi: 10.1186/s13018-022-03408-7.

[14] F. Yang and B. Ding, *Computer Aided Fracture Diagnosis Based on Integrated Learning*. 2020, pp. 523–527. doi: 10.1109/iciscae51034.2020.9236917.

[15] K. Murata *et al.*, “Artificial intelligence for the detection of vertebral fractures on plain spinal radiography,” *Scientific Reports*, vol. 10, no. 1, Nov. 2020, doi: 10.1038/s41598-020-76866-w.