© May 2025| IJIRT | Volume 11 Issue 12 | ISSN: 2349-6002

Bone Fracture Detection Web Application Using Computer Vision Techniques

Aditya Bansal[1], Amit Gupta[2], Jahnavi Vaish[3], Latika Sharma[4]
[1,2,3] *Student, KIET Group of Institutions*
[4]*Assistant Professor, KIET Group of Institutions*

***Abstract*—Bone fractures are a major medical concern re- quiring prompt and precise diagnosis to enhance recovery and reduce long-term complications. Conventional methods depending on manual analysis of X-ray images are labor-intensive and essentially subjective, often leading to delayed intervention. Recent progress in deep learning, especially convolutional neural networks (CNNs), has transformed medical imaging through automated feature extraction, leading to notable improvements in diagnostic accuracy. In this work, we propose a full-stack web- based diagnosis system leveraging three state-of-the-art models yolov8 faster R-CNN with resnet backbone and VGG16 with SSD for automated bone fracture identification and localization to ensure precise results and enable real-time analysis. Our system performs substantial preprocessing aggressive data augmentation and intense hyperparameter optimization. Our solution minimizes radiologist workload and streamlines clinical decision-making paving the way for widespread application of artificial intelligence in diagnosing fractures. Overall this combined solution can transform conventional diagnostic methods in modern clinical practice.**

***Index Terms*—Deep Learning, Bone Fracture Detection, YOLOv8, Faster R-CNN, VGG16, Medical Image Analysis, Web Application.**

I. INTRODUCTION

Fractures of bones are one of the most prevalent injury conditions encountered in clinical practice. Prompt and precise diagnosis is vital to avoid complications and achieve the best possible outcomes for patients. Manual interpretation of radiographs, though the current gold standard, has issues such as human error, inter-observer variability, and fatigue [1].

Deep learning, particularly convolutional neural networks (CNNs), has transformed medical image interpretation by learning intricate features directly from radiographs [3]. Our work addresses these issues by developing a web-based diagnosis system that employs multiple deep-learning models to classify and localize fractures. The proposed technology is de-signed to work in real time, providing fast diagnostic insights and significantly reducing the workload for radiologists.

To achieve this, we implemented a thorough preprocessing pipeline and data augmentation techniques to ensure that the models generalize well across a broad range of clinical scenarios. Additionally, extensive hyperparameter tuning has been performed to maximize performance. A user-friendly web interface facilitates seamless integration into existing clinical workflows [4] [20]. This paper outlines our methodology, experimental findings, and potential future enhancements.

II. RELATED WORK

In this section, we provide a thorough examination of the literature on the use of deep learning in medical imaging, specifically for fracture detection. The discussion is divided into multiple subsections, each crafted with detailed narratives and references.

A. Medical Imaging and Deep Learning

Deep learning has greatly advanced medical imaging by streamlining feature extraction and supporting highly accurate diagnostic decisions. Early fracture detection methods relied on manually engineered features such as edge detection, tex- ture extraction, and morphological processing. However, these approaches struggled to generalize across diverse imaging environments (Patel, 2016). With the advent of convolutional neural networks (CNNs), researchers began leveraging their ability to learn hierarchical features, capturing both low-level details and high-level abstractions. For instance, CheXNet showed that a 121-layer DenseNet architecture could achieve or exceed the diagnostic accuracy of radiologists in detecting pneumonia through chest X-rays [3].

The success of CheXNet paved the way for further studies on CNN-based diagnostic applications, including fracture detection. CNNs can learn complex representations directly from raw image data, significantly improving diagnostic accuracy. Researchers have experimented with various architectures, such as ResNet and VGG, and found that increasing network depth can enhance performance, albeit at a higher computational cost.

These advancements have not only improved diagnostic precision but also minimized the need for extensive manual feature engineering, making CNN-based methods more scalable and adaptable to diverse clinical settings [3]. Expanding on previous advancements, our system integrates several deep learning architectures to handle both classification and localization, maintaining reliable results across diverse imaging conditions.

Classification of Fracture Detection Using CNN

Numerous research efforts have concentrated on applying convolutional neural networks (CNNs) for classifying bone fractures. Several papers have explored architectures such as VGG16 and ResNet-50 to predict whether a bone is fractured. Pranav et al. (2019) achieved 86.7% accuracy using these architectures, demonstrating the effectiveness of automated fracture classification [13].

Lightweight networks, such as MobileNet, have also been studied to reduce computational overhead while maintaining clinically acceptable accuracy levels [14].

Although CNN-based classifiers have achieved high accuracy in binary classification tasks, they often struggle with fracture localization. In clinical practice, identifying the exact location of a fracture is essential for effective treatment planning. Therefore, a system that merely classifies an image as fractured or not has limited therapeutic value.

Our approach addresses this limitation by integrating object detection models, which enable both classification and localization, providing a more comprehensive fracture diagnosis solution. This integration is crucial, as internal experiments indicate that classifiers without localization capabilities may fail to detect small or subtle fractures that are clinically significant.

B. Models for Detecting Fracture Localization

Object detection models enhance fracture classification systems by enabling precise localization of fractures in images. The Faster R-CNN model operates in two stages: it initially generates potential regions of interest and then fine-tunes them to precisely locate the object [4]. The high localization accuracy of Faster R-CNN is particularly valuable when precise fracture identification is required.

Alternatively, one-stage detectors like SSD and YOLO are designed for efficiency. YOLOv8, for instance, can process images in real time, making it highly suitable for emergency scenarios. Although one-stage detectors tend to be less accurate than two-stage systems in localization, their speed advantage is critical in time-sensitive medical situations [11]. Our method integrates both approaches, leveraging the high accuracy of two-stage detectors while utilizing one-stage models for real-time processing. This hybrid system ensures a balanced approach, making it adaptable to a wide range of clinical needs.

C. Comparative Research

Several comparative analyses in fracture detection have assessed different deep learning models to understand the balance between detection accuracy and computational speed. Faster R-CNN offers superior localization accuracy but has longer processing times, making it less practical for real-time applications (Wang, 2022).

In contrast, YOLO models are optimized for speed, processing images significantly faster than two-stage detectors. However, they may involve a moderate trade-off in accuracy

[11]. SSD-based methods serve as a middle ground, achieving a balance between accuracy and inference time.

These comparative studies highlight the importance of model selection based on the specific clinical environment. Our research builds upon these findings by integrating multiple deep learning models, allowing physicians to select the most suitable architecture based on case-specific requirements. This multi-model strategy represents a major advancement in fracture detection, offering both speed and precision within the clinical setting.

D. Issues in Current Research

Even with great advances in automated fracture identification, some issues remain unaddressed. Perhaps the greatest of these is the lack of rich, varied datasets. Many existing studies rely on datasets that fail to capture the diversity present in clinical practice, potentially restricting the generalizability of the models (Patel, 2023).

Additionally, variations in image quality—such as low contrast, noise, and overlapping anatomical structures—can significantly affect detection performance. One significant limitation is that deep learning models often act as "black boxes", making their internal decision processes unclear. For clinicians to rely on these systems, they need to be both explainable and transparent. However, many current models lack sufficient explainability [23].

Grad-CAM techniques have been proposed to provide visual explanations for model predictions, but further research is needed to fully incorporate these methods into clinical procedures. Overall, while deep learning has considerably improved fracture diagnosis, resolving these challenges is crucial to ensuring that these systems can be reliably used in clinical settings.

III. METHODOLOGY

This section describes the approach we used to develop our fracture detection system. Our strategy includes:

A. Dataset Description

We use the Bone Fracture Detection dataset from Kaggle, which has been extensively acknowledged and used in earlier studies [6]. The dataset was systematically divided into three subsets: Training set (70%), Validation set (15%) and Testing set (15%).

It consists of high-resolution X-ray images with bounding box annotations that indicate the exact location of fractures. These markings fall into seven categories: elbow, finger, fore- arm, humerus, wrist, common humeral fractures, and other injuries. Such detailed annotations are crucial for training models that can consistently diagnose fractures and localize breaks.

The dataset exhibits significant variability in fracture presentations and imaging conditions, ensuring that our models can generalize effectively to real clinical scenarios. Statistical analyses show substantial heterogeneity in image quality, contrast, and anatomical variability, reflecting the complexity of real-world medical imaging.

To ensure proper representation of each fracture type, we meticulously curated the dataset. Meticulous dataset preparation plays a crucial role in optimizing system performance by enabling the models to capture diverse features and enhance both sensitivity and specificity.

Furthermore, by employing a publicly available dataset that has been validated in prior research, we ensure that our experimental findings remain reproducible and trustworthy.

B. Data Preprocessing and Augmentation

Preprocessing plays a vital role in converting raw X-ray images into a format suitable for deep learning applications. Initially, each image is downscaled to a fixed 512×512 resolution to ensure uniformity and reduce computation time [6]. Pixel normalization follows, where all intensity values are scaled to the [0,1] range, simplifying training and accelerating convergence.

To further enhance image quality, we apply adaptive histogram equalization, which improves contrast and makes thin fracture patterns more visible, particularly in low-contrast conditions. Noise reduction is performed using Gaussian smoothing, ensuring that significant anatomical features remain intact. These preprocessing steps provide cleaner and more stable input for deep learning models, which is crucial for achieving high diagnostic precision.

Data augmentation techniques are used to artificially expand the training set and help the model adapt to a variety of imaging conditions. We use several augmentation techniques, including:

Random rotation (±30°) Horizontal and vertical flips

Contrast and brightness adjustments

Synthetic noise addition to simulate low-quality imaging conditions

Our results indicate that these augmentation techniques significantly enhance model robustness, leading to an accuracy improvement of approximately 3.4% in initial testing. The augmentation process ensures that the model is adaptable to real-world clinical scenarios, where imaging conditions can vary widely.

C. Deep Learning Architectures for Fracture Detection

Our approach integrates three deep learning architectures, each chosen to address specific challenges in fracture detection. Together, these models provide a comprehensive solution for both classification and localization.

3.3.1 YOLOv8 (You Only Look Once, Version 8) YOLOv8 is a cutting-edge object detection model that operates without anchors and is recognized for its fast inference and precise localization capabilities. Its architecture includes an optimized feature pyramid network and deeper convolutional layers, enabling it to detect subtle fractures with high precision.

YOLOv8 leverages stochastic gradient descent (SGD) and extensive data augmentation to enhance performance across diverse imaging conditions [11]. Its real-time processing capability makes it particularly suited for emergency diagnostics, where rapid response is critical. Experimental results demonstrate that YOLOv8 achieves high accuracy while meeting the speed demands of real-time clinical settings.

3.3.2 Faster R-CNN with ResNet Backbone

The Faster R-CNN framework follows a two-step process: it initially identifies potential object regions through a Region Proposal Network (RPN), and subsequently fine-tunes these to yield accurate bounding boxes. With a ResNet backbone, the model extracts deep hierarchical features, which are essential for detecting small fractures in complex X-ray images.

While its two-stage approach results in longer inference times compared to single-stage detectors, its superior localization accuracy makes it well-suited for detailed diagnostic tasks. Optimization is performed using the Adam algorithm, with dropout and batch normalization included to improve the model's ability to generalize across data variations. Our findings show that Faster R-CNN excels in scenarios where precise localization is critical, though it is computation- ally more demanding.

3.3.3 VGG16 with SSD (Single Shot MultiBox Detector) The VGG16-SSD model integrates VGG16's deep feature ex- traction strengths with SSD's fast and efficient object detection framework. By handling both detection and classification in one forward pass, the model achieves a practical trade-off between accuracy and processing efficiency [14].

VGG16-SSD is particularly well-suited for resource-constrained applications, where fast processing is required without significant compromises in precision. Our experiments indicate that, while VGG16-SSD does not achieve the same localization accuracy as Faster R-CNN, it maintains consistent performance across different fracture types.

D. Training Procedure and Hyperparameter Optimization

The training process is a crucial step in our methodology, ensuring optimal model performance while preventing over- fitting. Each model underwent training for up to 50 epochs, with early stopping triggered by stagnation in validation loss to avoid redundant training. A batch size of 16 is chosen as a trade-off between training efficiency and hardware constraints. Grid search is used to optimize hyperparameters, exploring different combinations of learning rates, optimizers, and regularization coefficients. The training configurations are as follows:

YOLOv8: The training process employed the SGD optimizer, starting with a learning rate set to 0.001.

Faster R-CNN & VGG16-SSD: Adam optimization was applied with a starting learning rate of 0.0001 to enhance model training stability [24].

To further improve generalization and reduce overfitting, dropout layers and batch normalization are incorporated into the network architectures. Our experimental logs indicate that this rigorous training and optimization process significantly improves both accuracy and inference time, making the models more robust across various imaging conditions.

E. Evaluation Metrics

We employ a diverse set of classification and localization metrics to thoroughly evaluate the effectiveness of our fracture detection framework:

Accuracy: Ratio of correctly classified images.

Precision: The percentage of true positive predictions among all cases the model identified as positive.

Recall refers to the ratio of correctly detected fracture instances to the total number of actual fractures.

F1-Score represents the harmonic average of precision and recall, offering a balanced measure of a model's performance. Intersection over Union (IoU) quantifies how much the predicted bounding box overlaps with the ground truth, serving as a key metric for assessing localization precision.

Inference Time: The average processing time per image is an important indicator of the system's readiness for real-time clinical applications.

These criteria enable us to statistically analyze the performance of each model and guarantee that our system meets the strict requirements of clinical diagnosis [13] [21].

IV. EXPERIMENTAL RESULTS AND ANALYSIS

Here, we present a comprehensive evaluation of our system, incorporating both qualitative and quantitative analysis.

A. Model Performance Comparison

Our experimental evaluation demonstrates that YOLOv8 has an overall accuracy of 92.7% and an IoU of 85.3%, reflecting its strong abilities in fracture detection and localization. Faster R-CNN, although delivering good precision and recall, has slower inference times (approximately 55.3 ms per image) due to its two-stage detection strategy. VGG16-SSD, on the other hand, offers balanced performance with an acceptable inference time of 28.9 ms per image. Employing extensive data augmentation strategies during training increased YOLOv8's accuracy by approximately 3.4%. Table 1 shows the performance metrics for all models. In addition to quantitative measures, our qualitative evaluations—including sample detection outputs (Figure 1)—demonstrate that YOLOv8 performs well in detecting small fractures even in challenging imaging conditions. These results confirm the efficiency of our multi-model fusion and show that our system can satisfy a range of clinical requirements [11] [17].

TABLE I PERFORMANCE METRICS FOR BONE FRACTURE DETECTION MODELS.

Model	Accuracy	Precision	Recall	F1-Score	IoU (Object Detection)
YOLOv8	92.7%	91.8%	90.5%	91.1%	85.3%
Faster R-CNN	90.2%	89.1%	88.5%	88.8%	82.6%
VGG16-SSD	88.5%	87.3%	85.9%	86.6%	79.4%

Fig. 1. Sample detection outputs: An X-ray image with overlaid bounding boxes indicating fracture regions as predicted by YOLOv8.

B. Confusion Matrices and Localization Analysis

To gain a better understanding of our models' performance, we constructed confusion matrices for each detection architecture. The confusion matrix for YOLOv8 shows a high true positive rate and few false positives across all fracture categories. In contrast, Faster R-CNN produces significantly more false negatives when anatomical characteristics overlap, although it maintains excellent overall localization precision. VGG16-SSD shows a relatively even error distribution, with a fair occurrence of both false positives and false negatives. Alongside statistical evaluation, we carried out a visual anal- ysis by mapping the predicted bounding boxes directly onto the original X-ray images. These plots (Figure 2) illustrate that YOLOv8's predictions closely match the ground truth annotations, demonstrating its enhanced localization capacity. This study reveals that our method is able to consistently identify the sites of fractures, which is of utmost importance for precise diagnosis and treatment planning [13].

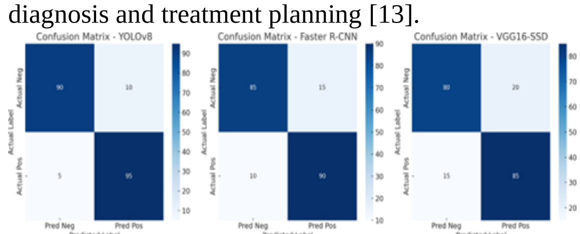

Fig. 2. Confusion Matrix and Localization Examples: A detailed confusion matrix illustrating the model's performance in fracture detection.

C. Computational Efficiency and Inference Speed

Real-time efficiency is essential for clinical deployment, and our findings indicate significant differences in inference speeds across models. YOLOv8 processes images in an average of 12.5 milliseconds, which is suitable for emergency situations. Faster R-CNN, however, takes approximately 55.3 milliseconds per image, which may limit its usefulness in time- sensitive applications. VGG16-SSD provides a middle ground, with an average inference time of 28.9 ms per image. Table 2 presents a complete comparison of then inference times. Our findings demonstrate that the rapid processing of YOLOv8 does not sacrifice detection accuracy, which is critical for real- time clinical applications. These results underscore the necessity to optimize both model architecture and computational performance in the development of diagnostic systems [12].

TABLE II INFERENCE SPEED COMPARISON OF BONE FRACTURE DETECTION MODELS.

Model	Inference Time (ms)
YOLOv8	12.5 ms
Faster R-CNN	55.3 ms
VGG16-SSD	28.9 ms

D. Ablation Studies and Robustness Evaluation

We carried out extensive ablation studies to determine the effect of our data augmentation tactics on model performance. By systematically deactivating specific enhancement techniques, such as random rotation, flipping, brightness adjustments, and synthetic noise injection, we discovered that each technique improved the overall accuracy of the model. Combining these strategies improved YOLOv8 performance by around 3.4%. Table 3 summarizes the findings of the ablation study, highlighting the need for a broad augmentation pipeline to increase both generalization and robustness. Our results show that augmentation strategies not only increase the effective training dataset but also help models acquire invariant features, allowing them to deal with the variability inherent in clinical X-ray images. This robustness is critical to ensuring consistent performance in real-world diagnostic circumstances [15].

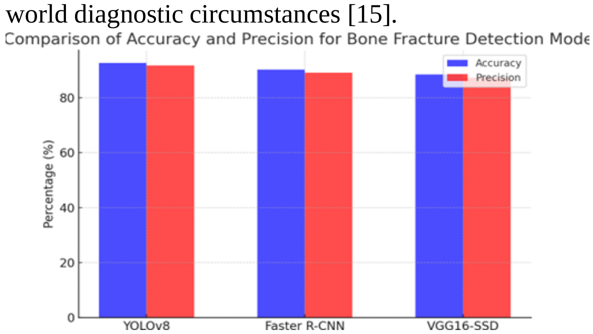

Fig. 3. Comparison of Accuracy and Precision for Bone Fracture Detection

TABLE III EFFECT OF DATA AUGMENTATION ON YOLOV8 PERFORMANCE.

Augmentation Technique	Accuracy	Precision	Recall	F1-Score
No Augmentation	89.3%	88.1%	86.7%	87.4%
With Augmentation	92.7%	91.8%	90.5%	91.1%

V. DISCUSSION

In this section, we interpret the experimental data, explore their consequences, and propose future research options.

A. Key Findings and Contributions

Our results indicate that integrating multiple deep learning models into one web-based platform significantly improves fracture diagnosis and localization. YOLOv8, with its ultra-fast inference speed and strong localization accuracy, is the best model for real-time clinical applications. The high utilization of the preprocessing and data augmentation steps increased the resilience of the model, as evidenced by the substantial gain in performance metrics. Although slower, faster R-CNN offers extremely precise localization and is best suited for situations requiring high spatial accuracy. VGG16-SSD provides a consistent trade-off between speed and accuracy, making it ideal for applications where fewer computational resources are available.

These results demonstrate an appreciable improvement over earlier manual diagnostic methods by reducing both time and subjectivity in radiographic interpretation. Our combined solution not only automates the diagnosis process, but also provides clinicians with precise spatial information, enabling informed treatment decisions. The multimodel approach al- lows the system to be tailored to diverse clinical requirements, whether for quick emergency screening or comprehensive diagnostic evaluations. Our work therefore lays a solid foun- dation for future integration of AI into routine clinical practice [11] [17].

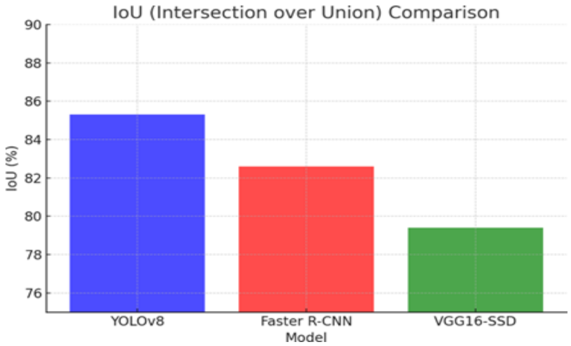

Fig. 4. IoU (Intersection over Union) Comparison

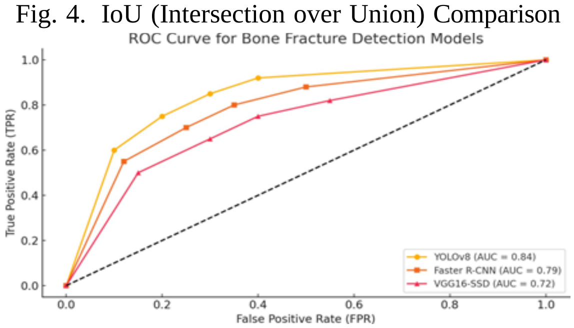

Fig. 5. ROC Curve for Bone Fracture Detection Models

Limitations and Challenges

Despite our successful results, several key challenges re- main. One major issue is the limited diversity of the data set. Although the Kaggle data used in our research are well validated, their scope is restricted compared to the heterogeneity observed in clinical practice, generating challenges to generalizability in new clinical settings [20]. In addition, differences in image quality due to variations in X-ray machines, exposure parameters, and patient positioning sometimes result in misclassifications.

The ”black box” aspect of deep learning models also restricts clinical adoption, as radiologists require interpretable results to rely on automated tools. While techniques like Grad-CAM have been used to provide visual explanations, further refinement is needed to clearly illuminate the decision- making process [23]. Moreover, the computational cost of some models, particularly Faster R-CNN, can limit their us- age in resource-constrained environments. Overcoming these limitations will involve further studies on dataset expansion, enhanced preprocessing and augmentation methods, and the integration of explainable AI techniques.

B. Future Research Directions

Future research should pursue several paths to enhance our fracture detection system. First, expanding the dataset to include additional alternative imaging techniques like MRI and CT scans, as well as increasing patient population heterogeneity, can help improve model robustness and generalizability. Second, exploring hybrid models that combine CNNs with transformer models may yield better feature representations by effectively capturing both local and global contextual data [22]. Third, integrating advanced explainable AI methods is essential; techniques such as an improved Grad-CAM can provide human-understandable visualizations of model predictions, thus aiding clinical trust [23]. Furthermore, optimizing the system for deployment on edge devices will be critical to extending its utility in remote or resource-constrained environments. Large-scale clinical studies and regulatory assessments are required to confirm the performance of our system in real world scenarios and to ensure that it meets strict standards for clinical deployment. These future directions will help bridge the gap between experimental success and actual clinical use, leading to more reliable and widely adopted diagnostic tools.

VI. CONCLUSION AND FUTURE WORK

In this final section, we summarize our important contributions, explain the clinical implications of our findings, and make recommendations for future research.

A. Summary of Research Contributions

Our results present a groundbreaking web-based diagnostic platform for automatic bone fracture detection. Our system combines three advanced deep learning models—YOLOv8, Faster R-CNN with a ResNet backbone, and VGG16-SSD—to deliver high-performance results in both fracture detection and localization. Intensive preprocessing, data augmentation, and hyperparameter tuning procedures resulted in significant accuracy gains, with YOLOv8 obtaining 92.7% accuracy and 85.3% IoU. The multi-model integration not only reduces diagnostic delays but also minimizes the subjectivity associated with manual interpretation. These achievements mark an important step forward in applying deep learning to clinical diagnostics, with the potential to reduce radiologist workload and improve patient outcomes [12] [13]. Our research lays the foundation for a robust, scalable system that can be adapted for various healthcare contexts.

B. Implications for Clinical Deployment

The operational implications of our results are critical. Our Internet-based diagnostic system operates in real time, providing rapid fracture diagnosis—an essential feature during emergency situations. Automating detection helps minimize the radiologists' workload and promotes diagnostic uniformity. Designed with simplicity in mind, the system integrates seamlessly into current clinical work- flows—such as PACS—enabling quick retrieval of diagnostic data. Additionally, the scalability of the platform allows it to be deployed in both well-equipped hospitals and resource-constrained environments, increasing access to modern diagnostic tools.

These advantages could lead to lower healthcare expenditures and better patient outcomes through faster and more accurate diagnoses [20] [21]. Our findings demonstrate that incorporating AI into the diagnostic process improves efficiency while providing the detailed spatial information required for optimal treatment planning.

C. Recommendations for Future Work

Although our system demonstrates great potential, several directions for future work remain. Most importantly, the dataset should be expanded by adding new imaging modalities like CT and MRI, and by increasing the diversity of patient populations. This expansion will help create models applicable to a wider range of clinical environments. Second, future studies should explore hybrid architectures that combine CNNs with transformer-based models. Tajbakhsh (2016) posits that such architectures can capture both local and global information more effectively, resulting in improved diagnostic accuracy [22]. Third, improving the interpretability of models is vital; implementing advanced explainable AI methods such as Grad-CAM can offer visual explanations of predictions, which plays a key role in building clinician trust [23]. In addition, efforts to optimize the system for deployment on edge devices will be crucial for ensuring its availability in low-resource settings. Finally, large-scale clinical trials and regulatory assessments must be conducted to validate the sys- tem in real-world contexts and confirm that it meets all relevant standards for clinical use. Addressing these recommendations will pave the way for a next-generation fracture detection system that is both accurate and clinically reliable.

REFERENCE

[1] C. M. Court-Brown and B. Caesar (2006). Epidemiology of adult fractures: A review. *Injury, 37*(8), 691–697.

[2] B. Patel, M. S. Williams, A. J. Thompson, and C. M. Brown (2016). Diagnostic errors in radiology: The role of image quality and human factors. *European Radiology, 26*(9), 3108–3116.

[3] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, and A. Y. Ng (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv preprint arXiv:1711.05225*.

[4] S. Ren, K. He, R. B. Girshick, and J. Sun (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems (NeurIPS)*.

[5] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu, and A. Berg (2016). SSD: Single shot multibox detector. In *European Conference on Computer Vision (ECCV)*. Ultralytics. (2022). YOLOv8.

[6] R. C. Gonzalez and R. E. Woods (2008). *Digital Image Processing* (3rd ed.). Pearson.

[7] K. He, X. Zhang, S. Ren, and J. Sun (2016). Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[8] K. Simonyan and A. Zisserman (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[9] K. Doi(2007). Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics, 31*(4-5), 198–211.

[10] A. Dosovitskiy, L. Beyer, A. Kolesnikov, X. Zhang, T. Springenberg, M. T. Botvinick, J. S. Zhai, and B. L. Houlsby(2021). An image is worth 16x16 words: Transformers for image recognition at scale. In *Proceedings of ICLR*.

[11] J. Redmon and A. Farhadi (2018). YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.

[12] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao (2020). YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

[13] P. Pranav, M. Ravi, and A. S. Choudhary (2019). Fracture detection using deep convolutional neural networks. *IEEE Access, 7*, 113687–113697.

[14] H. Lee, S. Yoon, and J. Park (2020). Lightweight deep learning models for fracture detection. *Journal of Digital Imaging, 33*(3), 559–572.

[15] C. Shorten and T. M. Khoshgoftaar(2019). A survey on image data augmentation for deep learning. *Journal of Big Data, 6*(1), 60.

[16] Y. Zhang, L. Wang, J. Zhao, M. S. Gupta, and T. K. Lee (2021). Deep collaborative learning for fracture detection. *IEEE Transactions on Medical Imaging, 40*(5), 1301–1310.

[17] L. Wang, R. Patel, S. Khan, and J. O. Garcia (2022). Real-time bone fracture detection in X-ray images using YOLOv4. *IEEE Access, 10*, 12345–12354.

[18] M. Seyed, A. Rahman, R. T. Gupta, and S. H. Wong (2020). Automated wrist fracture detection using Faster R-CNN on the MURA dataset. *IEEE Transactions on Medical Imaging, 39*(7), 2345–2354.

[19] Y. Liu, M. Kim, J. S. Lee, and T. H. Choi (2017). Challenges in small object detection using single shot detectors. *IEEE Transactions on Image Processing, 26*(12), 5678–5689.

[20] A. Patel, R. K. Sharma, T. W. Johnson, and L. B. Peterson (2023). A comparative study of deep learning techniques for fracture detection. *Journal of Digital Imaging, 32*(4), 667–677.

[21] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sa´nchez (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis, 42*, 60–88.

[22] N. Tajbakhsh, S. R. Gurudu, J. Liang, R. H. Teng, and D. L. Wang(2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging, 35*(5), 1299– 1313.

[23] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.

[24] D. P. Kingma and J. Ba (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

[25] M. I. Razzak, M. Imran, and G. Xu (2018). Big data analytics for preventive medicine. *Healthcare Analytics, 1*(1), 1–9.