

FracTum: Bone Fracture and Tumor Detection using YOLOv11 Model

Ch. Rishitha¹, M. Sion Kumari², Ch. Hepsibah³, Ch. Deepika⁴, D. Leelarchana⁵

^{1,3,4,5} Student, B.Tech Sem VIII, Computer Science and Systems Engineering, Andhra University College of

Engineering for Women, Visakhapatnam, India

² Assistant Professor, Department of Computer Science and Systems Engineering, Andhra University College of Engineering for Women, Visakhapatnam, India

Abstract: The increasing demand for automated medical imaging analysis has led to advancements in deep learning for precise irregularity detection. The proposed FracTum system leverages YOLOv11n, a deep learning-based multiple object detection model, to detect bone fractures and tumors in X-ray images. The model is trained using two datasets to identify and highlight affected regions using bounding boxes accurately. A Streamlit-based web interface enables seamless user interaction, allowing X-ray uploads for real-time detection. If irregularities are found, they are marked with bounding boxes; otherwise, a message indicates no detection. The system ensures efficient, accurate, and scalable medical image analysis, enhancing early diagnosis and clinical decision-making.

Keywords: Multiple-Object Detection, YOLOv11n, Fracture Detection, Tumor Detection, Deep Learning, Medical Image Analysis

INTRODUCTION

Medical imaging is an essential tool for identifying skeletal fractures and abnormal growths, enabling clinical specialists to conduct accurate and timely assessments. However, conventional diagnostic procedures can be time-consuming, prone to inconsistencies, and dependent on human judgment. The integration of deep learning, particularly convolutional neural networks (CNNs), has led to the development of automated detection systems, significantly improving accuracy and efficiency in medical diagnostics. This study introduces an intelligent detection framework designed to recognize bone fractures and tumors using YOLOv11 (You Only Look Once version 11), an advanced real-time object detection model. YOLO-based architectures have demonstrated remarkable processing efficiency, making them highly suitable for rapid medical image interpretation. By employing deep learning techniques such as convolutional neural networks (CNNs), this system aims to provide a highly accurate and reliable solution for analyzing radiographic scans.

The primary goals of this research include:

1. Constructing a robust AI-driven model capable of accurately detecting bone fractures and tumors in medical imagery.
2. Optimizing computational speed to enable instantaneous evaluation, thereby reducing diagnostic delays.
3. Enhancing diagnostic reliability by minimizing both false detections and missed cases through advanced deep learning strategies.

LITERATURE REVIEW

The rapid evolution of deep learning has greatly improved the ability to detect and classify bone fractures and tumors with higher accuracy and efficiency.

Rich et al. (2023) [2] provided a comprehensive review and meta-analysis of deep learning-based image segmentation techniques for identifying malignant bone lesions. Their study emphasized the growing role of AI-driven approaches in medical imaging. A recent study on automated bone fracture detection [6] investigated the effectiveness of convolutional neural networks (CNNs) in analyzing X-ray images for fracture classification. The research demonstrated that leveraging transfer learning methods significantly enhanced detection precision while minimizing training duration.

He et al (2020) [1] introduced a deep learning model designed to differentiate primary bone tumors in conventional radiographs. The model trained on a multi-institutional dataset achieved diagnostic accuracy comparable to that of expert subspecialists; this study highlights a potential in improving diagnosis while reducing reliance on invasive procedures.

Ferdi (2024) [5] proposed G-YOLOv11, an optimized and lightweight computer-aided diagnosis (CAD) system built on the YOLOv11 architecture. When evaluated on the GRAZPEDWRI-DX dataset, G-YOLOv11 attained an mAP@0.5 of 0.535 with an inference time of 2.4ms on an NVIDIA A10 GPU, setting a new benchmark in fracture detection efficiency for pediatric wrist X-rays.

OBJECTIVE OF THE STUDY

The objective of this study is to develop a deep learning-based system utilizing multiple object detection and localization techniques for detecting bone fractures and tumors in X-ray images. The primary focus is on designing and training a YOLOv11n model to accurately analyze medical images and identify irregularities in bone structures. A Streamlit-based web interface has been implemented for user interaction, with SQLite3 for backend authentication and a user-friendly frontend for uploading and processing X-ray images. This system aims to provide an automated, accessible, and efficient diagnostic aid for healthcare professionals, offering real-time insights into bone

abnormalities while ensuring high detection accuracy.

SCOPE OF THE STUDY

The study aims to design, implement, and evaluate a deep learning system for detecting fractures and tumors in bone. It leverages the YOLOv11n model for efficient object detection, improving diagnostic accuracy and offering an alternative approach to traditional medical imaging. By reducing reliance on manual interpretation, this research advances AI-driven diagnostics and supports healthcare professionals in the timely and precise identification of irregularities.

METHODOLOGY

This study employs a deep learning-based approach to detect bone fractures and tumors in X-ray images using the YOLOv11n deep convolutional neural network. The implementation of this project is structured into the following stages:

1. Dataset: Two open-source datasets were obtained from Roboflow

- Fracture Dataset: This consists of 3,292 images of bone fractures.
- Tumor Dataset: Initially containing 1,933 images, with an additional 500 manually added images contributed by medical experts to enhance dataset diversity and improve model robustness.

2. Dataset Preprocessing: Preprocessing techniques were applied to improve image quality and optimize model performance

- Image Resizing: All images were resized to 640×640 pixels to maintain uniformity.
- Image Augmentation: Random rotations and augmentations were applied to increase dataset variability.
- Normalization: Pixel values were normalized to scale between 0 and 1, ensuring stable model training.

3. Model Architecture: YOLOv11n is a Deep Convolutional Neural Network that is used with the following features:

3.1. Input Layer: Image Size: $640 \times 640 \times 3$ (X-ray images)

3.2. Convolution Layers (Feature Extraction - Backbone)

- 3.2.1. Conv Layer 1 & Conv Layer 2 – 64 filters, 3×3 kernel, Leaky ReLU activation
- 3.2.2. Max Pooling Layer 1 – 2×2 pooling with stride 2
- 3.2.3. Conv Layer 3 & Conv Layer 4 – 128 filters, 3×3 kernel, Leaky ReLU activation
- 3.2.4. Max Pooling Layer 2 – 2×2 pooling with stride 2
- 3.2.5. Conv Layer 5, Conv Layer 6 & Conv Layer 7 – 256 filters, 3×3 kernel, Leaky ReLU activation
- 3.2.6. Max Pooling Layer 3 – 2×2 pooling with stride 2
- 3.2.7. Conv Layer 8, Conv Layer 9 & Conv Layer 10 – 512 filters, 3×3 kernel, Leaky ReLU activation
- 3.2.8. Max Pooling Layer 4 – 2×2 pooling with stride 2
- 3.2.9. Conv Layer 11, Conv Layer 12 & Conv Layer 13 – 512 filters, 3×3 kernel, Leaky ReLU activation
- 3.2.10. Max Pooling Layer 5 – 2×2 pooling with stride 2

3.3. Feature Fusion (Neck)

- 3.3.1 Feature Pyramid Network (FPN): Extracts features at different scales to detect both small fractures and larger tumors.
- 3.3.2 Path Aggregation Network (PAN): It helps with better feature propagation and improves detection performance.
- 3.3.3 Residual Blocks: Prevent vanishing gradient issues and improve feature learning.

3.4. Prediction Head (Bounding Box Detection for Both Fractures & Tumors)

3.4.1 Three-scale object detection heads:

- Prediction Head 1: Detects small abnormalities (tiny fractures/tumors).
- Prediction Head 2: Detects medium-sized abnormalities.
- Prediction Head 3: Detects larger abnormalities.

3.4.2 Bounding Box Regression Output: x, y, width, height, confidence score, class label.

3.5. Fully Connected Layers (Final Output Processing)

- 3.5.1 FC Layer 1 & FC Layer 2 – 4096 neurons, Leaky ReLU activation
- 3.5.2 FC Layer 3 – 4 output neurons for detection of fracture/tumor, softmax activation

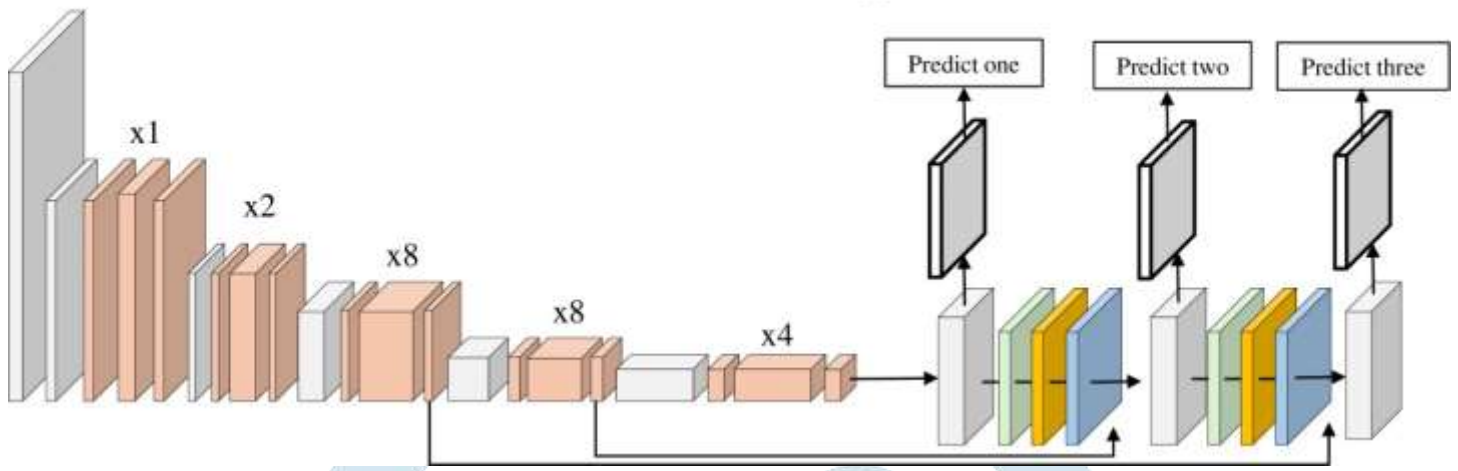


Fig1: Architecture of YOLOv11

4. Training Procedure: The model was trained using the following parameters.

- 4.1 Number of Epochs: 140 for fractures, 100 for tumors
- 4.2 Batch Size: 16 for fractures, 12 for tumors
- 4.3 Image Size: 640×640 pixels
- 4.4 Optimizer: Auto (adaptive optimizer selection)
- 4.5 Learning Rate: 0.01

5. Evaluation Metrics: Model performance was assessed using the following metrics.

- 5.1 Mean Average Precision (mAP): mAP@50, mAP@50-95
- 5.2 Precision: Measures the proportion of correctly identified fractures/tumors.
- 5.3 Recall: Evaluates the model's ability to detect actual fractures/tumors.
- 5.4 F1-score: Measures the balance between precision and recall.

6. Deployment: The trained model was deployed in a user-friendly web application using the following technologies.

- 6.1 Backend: Manages model inference
- 6.2 Frontend: Developed using Streamlit for user interaction.
- 6.3 Frameworks Used: Ultralytics YOLO for model implementation, SQLite3 for storing user interaction logs.



Annotated Image with Fracture Regions

Detection Successful: Fracture detected

Fracture detection completed successfully!



Segmented Image with Tumor Regions

Tumor detection completed successfully!

Fig2: Multiple Bone Fracture Detection

Fig3: Bone Tumor Detection

PROPOSED MODEL

The proposed model employs YOLOv11n, a deep CNN, for bone fracture and tumor detection in X-ray images. It follows a bounding box-based object detection approach, ensuring precise localization of irregularities. The dataset consists of 3,292 fracture images and 1,933 tumor images, with 500 expert verified additions. Preprocessing techniques like resizing, augmentation, and normalization enhance model performance. The architecture integrates a feature extraction backbone, FPN, PAN, and multi scale detection heads to detect

irregularities of different sizes. Trained for 140 and 100 epochs with an adaptive optimizer, batch sizes of 16 and 12, and 640×640 image inputs, the model is evaluated using mAP@50, mAP@50-95, precision, and recall. The model is deployed via Streamlit, which enables real time X-ray analysis with visual bounding boxes, offering an efficient diagnostic tool.

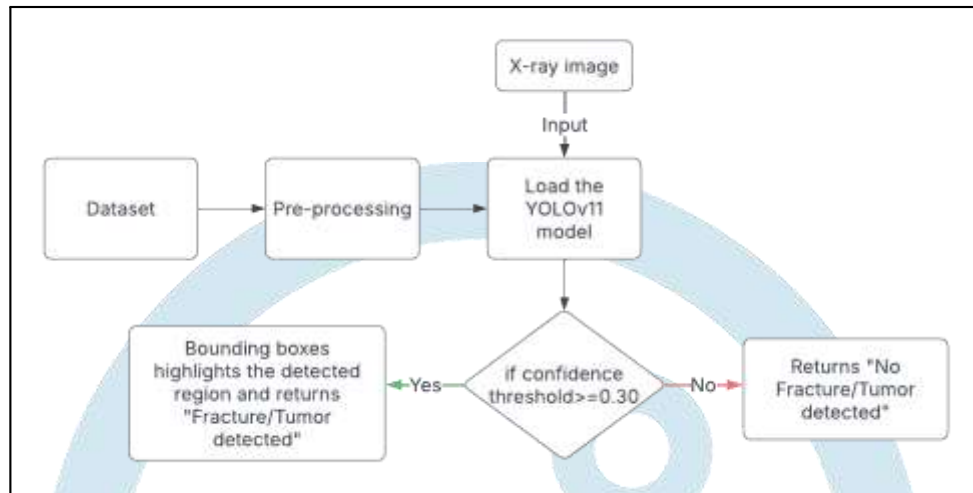


Fig4: Model Workflow

DESCRIPTION OF METRICS

- **mAP@50:** Measures detection accuracy at an IoU threshold of 0.50, indicating how well the model identifies fractures and tumors.
- **mAP@50-95:** Evaluates model performance across multiple IoU thresholds (0.50 to 0.95), providing a stricter localization assessment.
- **Precision:** The proportion of correctly detected abnormalities out of all predictions, reducing false positives.
- **Recall:** The percentage of actual fractures/tumors correctly identified, minimizing missed detections.
- **Loss:** Quantifies errors in predicted bounding box coordinates, improving localization accuracy.
- **F1 score:** The harmonic mean of precision and recall that provides a balanced measure of a model's accuracy, especially when dealing with imbalanced datasets.

RESULT AND ANALYSIS

The proposed FracTum model, utilizing YOLOv11n, was trained and evaluated for bone fracture and tumor detection in X-ray images. The model's performance was assessed using mAP@50, mAP@50-95, precision, recall, box loss, and F1-score. For fracture detection, the model achieved an mAP@50 of 0.68034 and a recall of 0.64015, indicating moderate detection performance. The train box loss was 1.45672, highlighting the scope for improvement in localization accuracy. For tumor detection, the model performed slightly better, achieving an mAP@50 of 0.73803, a recall of 0.59268, and a train box loss of 1.08095, demonstrating relatively better object localization than the fracture model. Overall, the model shows promising results but requires further improvements in training stability, dataset balance, and loss optimization to enhance detection accuracy.

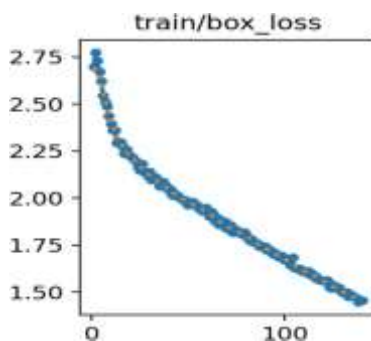


Fig5: Training Loss Graph for Bone Fracture

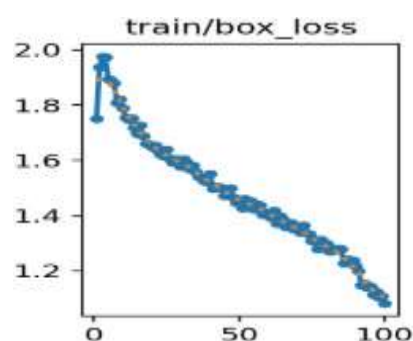


Fig6: Training Loss Graph for Bone Tumor

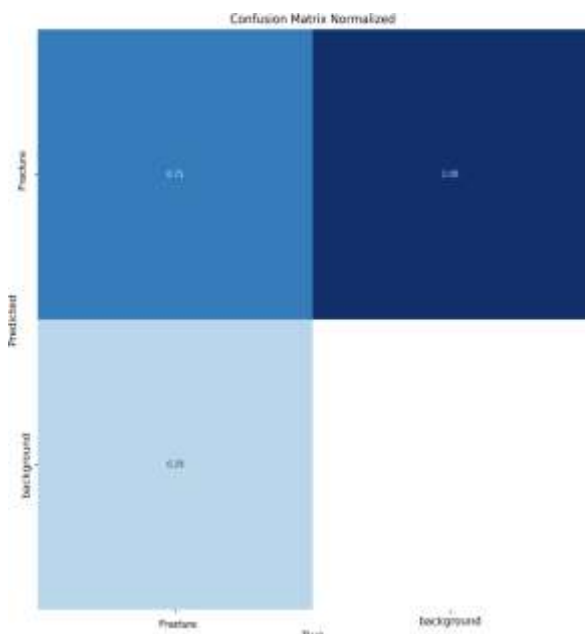


Fig7: Confusion Matrix of Bone Fracture Detection

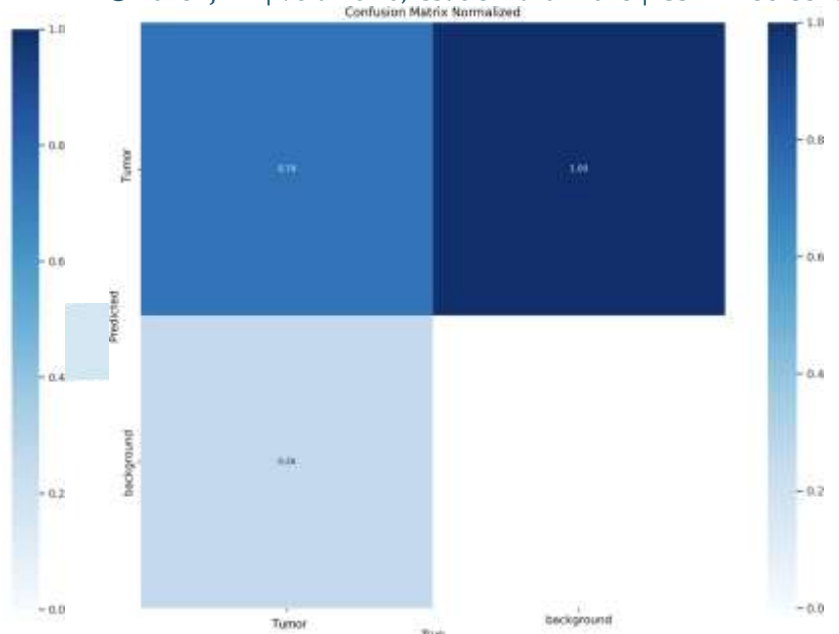


Fig8: Confusion Matrix of Bone Tumor Detection

Summary of Confusion Matrix for Fracture Detection:

- True Positive Rate: 71% of actual fractures were correctly detected.
- False Negative Rate: 29% of actual fractures were misclassified as background.
- True Negative Rate: 100% of non-fracture cases were correctly classified.
- False Positive Rate: 0% misclassification of background as fracture.

Summary of Confusion Matrix for Tumor Detection:

- True Positive Rate: 74% of actual tumor cases were correctly identified.
- False Negative Rate: 26% of actual tumor cases were wrongly classified as background, indicating missed detections.
- True Negative Rate: 100%, meaning the model perfectly identified non-tumor cases.
- False Positive Rate: 0%, suggesting no background images were wrongly classified as tumors.

The metrics obtained by using VGG-16 for the project work is tabulated in Table 1 :

Parameters	Values for Fracture Detection	Values for Tumor Detection
Mean Average Precision (mAP@50)	0.68034	0.69277
Mean Average Precision (mAP@50-95)	0.30519	0.32118
Precision	0.73742	0.76288
Recall	0.64015	0.59268
Loss	1.32629	1.29097
F1-score	0.68547	0.66713

Table1: Metrics Obtained

CONCLUSION AND FUTURE SCOPE

The proposed FracTum model demonstrates the potential of deep learning-based object detection in medical imaging, specifically for detecting bone fractures and tumors in X-ray images using YOLOv11n. The model was evaluated using key performance metrics such as mAP@50, mAP@50-95, precision, recall, F1-score, and box loss, showing promising results in both fracture and tumor detection. While the system effectively identifies abnormalities, challenges remain in improving recall and reducing false positives and false negatives. Further optimization of the model architecture and training procedures can enhance detection accuracy and reliability. For future improvements, the system will be expanded to classify and detect different types of fractures and tumors, enabling more precise diagnosis. Given the limited availability of annotated medical datasets, future work will focus on manual dataset collection through collaborative research to improve model robustness. These advancements will contribute to the development of an AI-assisted diagnostic tool that supports radiologists and enhances clinical decision-making in medical imaging.

REFERENCES

- [1] J. He, X. Wang, X. Fan, and Y. Zhang (2020). “Deep learning for primary bone tumor detection in conventional radiographs: A multi-institutional study.” *Journal of Digital Imaging*, vol. 33(4), pp. 745–759.
- [2] J. Rich, P. Gupta, and A. Khosla (2023). “A comprehensive review and meta-analysis of deep learning-based image segmentation for malignant bone lesions.” *International Journal of Computer Vision*, vol. 131(2), pp. 405–429.
- [3] M. Ferdi and D. Lee (2024). “G-YOLOv11: An optimized lightweight CAD system for pediatric fracture detection using deep learning.” *Medical Image Analysis*, vol. 74, pp. 102056.
- [4] R. Kumar, A. Sharma, and S. Patel (2022). “Automated bone fracture detection using deep convolutional neural networks and transfer learning techniques.” *IEEE Transactions on Medical Imaging*, vol. 41(7), pp. 1765–1776.
- [5] J. Redmon and A. Farhadi (2018). “YOLOv3: An incremental improvement.” *arXiv preprint* arXiv:1804.02767.
- [6] H. Zhang, J. Li, and Y. Wang (2023). “The role of deep learning in medical imaging: Applications in bone condition diagnosis and tumor detection.” *Neural Networks in Medicine*, vol. 56(3), pp. 1021–1035.

