Automated Bone Fracture Detection For X-Ray Images

Prof. Pallavi Narkhede[1], Bansi Bhimraj Thorat[2], Vrindha Trakroo[3], Riya Shivaji Patil[4], Priti Birappa Vhanzende[5]

[1,2,3,4,5]*Department of Information Technology, Bharati Vidyapeeth's College of Engineering for Women Pune, Maharashtra, India*

***Abstract*—Detection of bone fractures is an important aspect of medical diagnostics; however, the detection process is done manually by analyzing images acquired through X-rays. This task consumes a lot of time and might be challenging in some instances. In this study, a deep learning system for the automatic detection of bone fractures from X-rays is proposed. Transfer learning using DenseNet121 network is incorporated to classify input images as either fractured or non-fractured. Resizing, normalization, converting grayscale image to RGB color space, and applying data augmentation such as rotating, flipping, affine transformation, and adjusting the brightness of the input image is done for improved model performance. In addition to classifying input images, the YOLO model is used to locate the fracture on the image accurately, whereas, Grad-CAM is implemented to explain the prediction made by the model. Experiments show that the proposed model is capable of detecting fractured bones with accuracy around 80% and very high F1-score and ROC-AUC metrics.**

Index Terms—Bone Fracture Detection, X-ray Imaging, Deep Learning, Transfer Learning, DenseNet121, YOLO, Fracture Localization, Grad-CAM, Medical Image Analysis, Data Augmentation.

I. INTRODUCTION

Fractures are among the commonest ailments that affect individuals all around the globe. Such fractures may be caused by internal and external reasons. Internal reasons include bone diseases, while external reasons involve physical activities and accidents. [1][2]. The extent of fractures is diverse, starting from tiny cracks in the bones to full displacement of the bones, which makes the need for proper diagnosis imperative to ensure effective treatment. The prevalence rate of fractures according to contemporary research is steadily rising around the world, posing heavy burdens to health care facilities [6]. If fractures are not diagnosed correctly at an early stage, they may lead to complications such as improper healing, long-term disability, or increased treatment costs.

A number of medical imaging techniques like X-rays, CT scans, MRIs, and ultrasounds can be used to diagnose fractures. Out of all the medical imaging techniques, the most common one is X-ray imaging because of its high availability and relatively low cost [2][3]. X-rays can offer enough resolution to detect bone breaks and deformations, which makes them the most commonly used method for diagnosing fractures. Nevertheless, conventional techniques of fracture recognition are based on manual inspection carried out by radiologists. This process takes a lot of time and depends on the skills of a specialist. Moreover, under complicated conditions, such as superimposed bone structures, noise, or poor image quality, fractures can be overlooked [1][3][4]. Additionally, the increasing workload on medical professionals further increases the chances of diagnostic errors.

Because of the fast pace at which artificial intelligence (AI) is evolving, there has been much emphasis on the use of machine learning and deep learning in medical imaging. The technique makes it possible to extract features automatically from X-rays images [1][6]. CNNs have been found to be quite successful at pattern recognition like edges, texture, and fracture detection on medical imagery [1]. Furthermore, transfer learning, which involves the use of pre-trained networks, has greatly contributed to the success of CNNs in image analysis tasks, particularly in small-sized medical datasets [6]. Pre-trained models can leverage existing feature extraction knowledge from large datasets and apply them to medical imagery problems.

Another recent advancement is the integration of algorithms like YOLO that can be used in detecting where exactly in an X-ray image the fracture lies, thereby helping classify and localize the problem area in question [2]. This is helpful not only to detect the presence of a fracture but to find out where exactly it is located in the image. Another technique used to enhance the interpretability of deep learning algorithms is the use of Explainable AI (XAI) through methods like Grad-CAM, which show important parts of the image that play a significant role in making the decision [6].

This paper proposes the design of an intelligent bone fracture detection framework based on deep learning that utilizes X-ray images. This includes using the DenseNet121 architecture for classification, the YOLO algorithm for fracture localization, and Grad-CAM for visual explanation. Pre-processing steps involving resizing, normalization, grayscale to RGB transformation, and augmentation are employed to enhance the performance and generalizability of the model. The proposed system can be considered a complete solution as it not only detects but also localizes fractures, and gives their visual explanations.

II. RELATED WORK

Over the past few years, there has been great progress achieved in medical imaging analysis because of the use of advanced machine learning and deep learning. This approach has proven to be effective when used in bone fracture detection using X-rays. Various researches conducted in this area were aimed at designing computer algorithms that help medical experts detect fractures and minimize errors made by humans [1][6]. With the development of artificial intelligence, more and more applications started being used in medical diagnostics.

Previous works mostly made use of conventional image processing methodologies to detect bone fractures. Conventional image processing methods usually require several processes, including pre-processing, feature extraction, and classification. Methods such as edge detection based on the Canny operator, shape detection through the Hough transform, and feature extraction through Harris Corner Detection were popular in conventional image processing. The extracted features would be classified using machine learning methodologies such as Support Vector Machine (SVM), Decision Tree, Random Forest, and K-Nearest Neighbours (KNN) [3]. Despite their effectiveness in attaining accuracies within the range of 64% to 92%, they faced challenges such as the need for manual feature extraction and sensitivity to noise.

Following the development of deep learning, CNNs have become the go-to method for analyzing medical images. CNNs learn hierarchical features automatically without requiring feature extraction, making them highly accurate and efficient. Multiple experiments have proven the viability of CNN models for detecting fractures [1][6]. Moreover, transfer learning methods that use pre-trained models such as DenseNet, ResNet, and MobileNet have become popular for enhancing the accuracy of CNN models, particularly in the case of small medical datasets [6]. Pre-trained models employ features learned from extensive data sources to facilitate specific medical imaging applications.

The DenseNet architecture has been proven to be more efficient because of the dense connections that exist between each layer and all other layers. It makes the training process more efficient, and helps propagate features well and eliminate the issue of vanishing gradients [6]. In addition, numerous papers have employed DenseNet121 in classifying bone fractures, and it has been successful in its applications.

Classification alone is not the only area that researchers have investigated lately; there has also been work done in the area of using object detection algorithms such as YOLO and Faster R-CNN for localization of fractures. Such object detection models are able to detect the presence of fractures in X-ray images and can precisely locate them through the use of bounding boxes [1][2].

Explanation of artificial intelligence (XAI) has become another area of focus when using AI in medical imaging. Methods such as Grad-CAM have been commonly employed in visualizing the important parts responsible for predictions made by the AI models. The heatmaps generated by Grad-CAM identify areas of importance within X-rays, and this makes decisions made by the AI model easy to understand [6].

Additionally, image pre-processing and augmentation techniques are extremely important in increasing model accuracy. Image scaling, normalization, rotation, flipping, affine transformations, and contrast adjustments are examples of techniques used to enhance diversity within the dataset and decrease overfitting [1][3]. These processes are particularly helpful in fields like medical imaging where the dataset size is usually small and unbalanced.

Even with all these improvements, there are still various difficulties encountered in automated detection of fractures. The presence of different kinds of fractures, noise in X-rays, and limited availability of annotated data pose serious threats to model effectiveness. In some cases, classifying whether a bone is fractured or not can be very challenging due to similar characteristics of their corresponding images [2][5].

From the literature reviewed, the incorporation of deep learning, especially DenseNet Transfer Learning combined with object detection and explainable approaches like YOLO and Grad-CAM respectively, is expected to improve bone fracture detection systems. Nevertheless, there is still a need for further improvement in order to cope with different medical problems and obtain higher accuracy.

Recent studies have also developed deep learning fracture detection systems with enhanced performance and accuracy [7][8].

III. METHODOLOGY

A. Data Resource

In this study, an X-ray image dataset is used, which is a publicly available medical imaging dataset available online through Kaggle, for instance. The X-ray images are split into two categories: the first category is that of bone fractures (positive category), while the second one is that of no bone fractures (negative category). In order to facilitate the process of working with the dataset and its compatibility with machine learning libraries like PyTorch, the dataset is placed in different folders according to the category label.

There are numerous X-ray images available in this dataset along with variations in terms of fracture types, bone alignment, quality, and body part positions. This variation is essential in order for the model to be able to generalize and learn useful features, thus enhancing the accuracy of the model when applied to new, unseen images. Moreover, the labeled dataset enables the model to adopt a supervised learning strategy, which is highly effective in discriminating between images of fractured and intact bones. The variations within the dataset will allow the model to detect both lower-level features, such as edges and texture information, as well as higher-level features, including fractures and bone abnormalities.

In order to have an adequate training and evaluation process, the data set is split into three sets: Training, Validation, and Testing. Training set includes 8450 images (8448 labeled examples), which is used for learning the essential features by the model, including bone structure and fractures. Validation set is represented by 1076 images, which is used to control and prevent overfitting during training process. Testing set includes 1110 images, and it is used to measure the performance of the system at the end of the training. The dataset allows performing two types of experiments: Classification experiment using DenseNet121, and Localization using YOLO model.

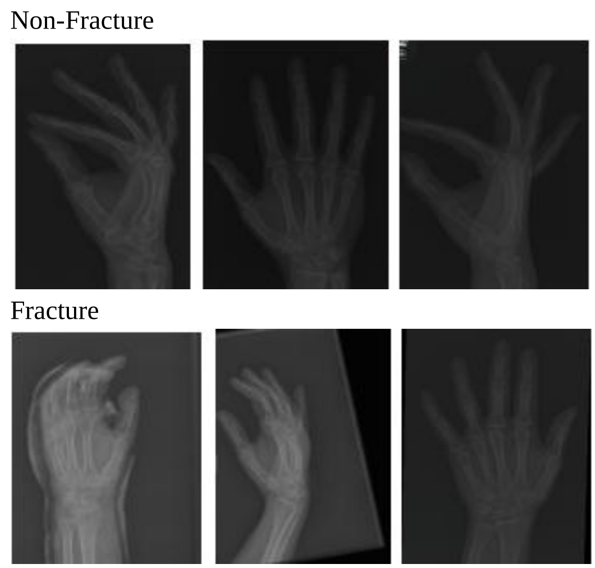


Figure 1: Sample X-ray Images from Dataset

Table 1: Dataset Distribution

Dataset	Total Images
Training Set	8450 images (8448 labeled)
Validation Set	1076 images (1076 labeled)
Testing Set	1110 images (1110 labeled)

Each image has been assigned one of two categories: fractured or non-fractured. This makes it a classification of binary nature. Binary classification helps make the job easier and also ensures that the machine makes the right decisions. The class labels are handled using the class mapping technique in which label data is stored in a JSON file. In this class mapping, the category fractured is assigned label 0, and non-fractured is assigned label 1.

B. Preprocessing

Preprocessing step is extremely important in the proposed approach as it guarantees that input X-rays are adequately processed prior to passing them through any deep learning architecture. The size, orientation, brightness, and contrast of medical images tend to differ, and thus, it is important to ensure adequate preprocessing in order to achieve consistent results from deep learning models. Preprocessing of X-ray images will also guarantee that the level of noise is minimized as well as the presence of important features such as bone edges and fracture patterns.

Firstly, input X-rays should be transformed to the grayscale version as a way of simplifying the image. Since DenseNet121, a pre-trained transfer learning model, needs input images to be processed through three channels, i.e., RGB, input X-rays need to be transformed to three channels by duplicating the same channel thrice.

The images are further resized to have dimensions of 224 by 224 pixels after channel conversion. Resizing of the image is necessary as the deep learning model requires input size consistency in order to optimize computational costs. It further reduces the computational complexity of the problem and increases efficiency during training. After resizing, normalization of the image is done based on the use of standard ImageNet parameters to normalize the pixel values.

Other than simple processing of images, augmentation methods are also applied to training data to enrich the dataset for better generalization. The augmented data will increase the variance and provide different perspectives, which will minimize the risk of overfitting. In light of limited datasets from the medical field, augmentation of the dataset will ensure that the model can generalize. Types of augmentation to be used include horizontal flipping, rotation, affine transforms like scaling and translation, and adjustments of brightness.

In addition, efficient data management is achieved through the utilization of the DataLoader from the PyTorch library. The DataLoader allows the batch-wise processing of the data, random shuffling of the data, as well as data loading. The latter will help to improve the efficiency of training and optimize memory usage. Next, the preprocessed images are fed into a model called DenseNet121 that performs classification based on whether the image depicts a fracture or not. Meanwhile, the same set of prepared images is also fed into the YOLO model to perform localization of fractures.

It should be noted that proper data preprocessing contributes to the balanced ratio between positive and negative samples. Thus, biases in the model prediction will be minimized. Proper organization of the dataset also ensures that no problems occur during the data loading. Overall, data preprocessing and data preparation enable converting raw images into a convenient input form that facilitates effective learning and generalization by the neural networks.

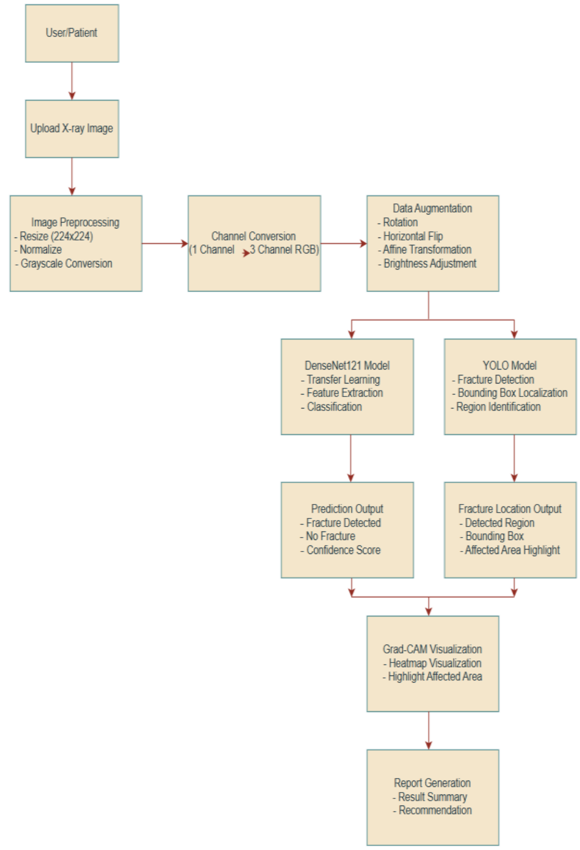


Figure 2: Proposed System Architecture for Bone Fracture Detection

C. Model Architecture (DenseNet121 Based Approach)

In this research work, the current study involves the use of the DenseNet121 architecture as the main deep learning framework for extracting features from X-ray images and their classification. DenseNet121 is a Convolutional Neural Network famous for having dense connectivity, which means that any layer in the dense block is directly connected to the subsequent layers of the block. Such connectivity allows for efficient use of features.

Figure 3 shows that the input to the model is an image of dimensions 224 × 224 × 3, which results from preprocessing tasks like resizing and channel transformation. The structure of the DenseNet121 network is made up of several dense blocks connected by transition blocks. Inside each dense block, the convolution layer is followed by the batch normalization layer and the activation function layer, and the model is able to extract both low-level features, such as edges and textures, and high-level features, such as fracture and bone patterns.

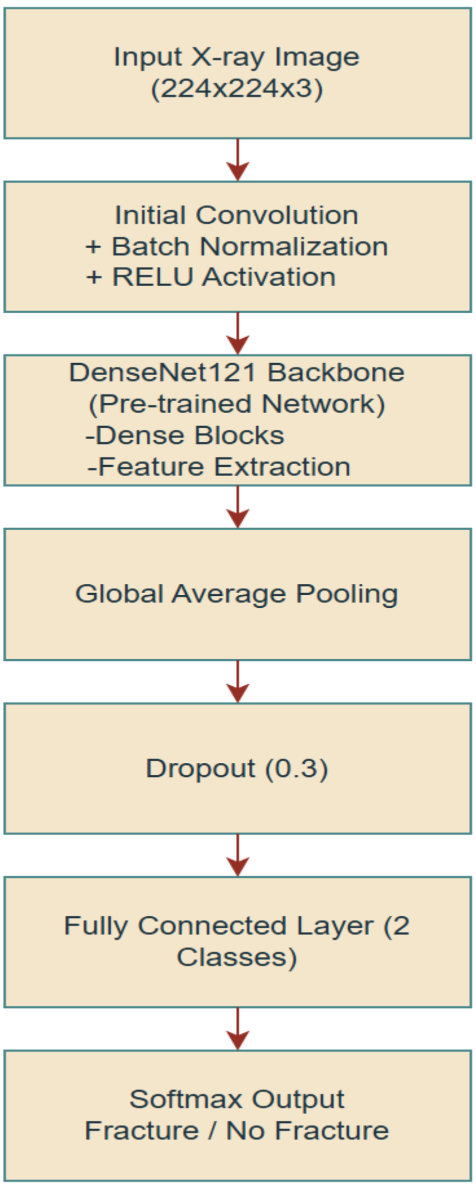


Figure 3: Architecture of the proposed DenseNet121-based model for Bone fracture Detection

Once the feature extraction is done, a global average pooling layer is added to minimize the spatial dimensions of the feature map and transform it to a feature vector. A dropout layer is added to mitigate overfitting and enhance generalization. At last, a linear layer is utilized to classify the model into two classes—fracture and non-fracture. The output layer applies the softmax activation function to derive class probabilities and picks up the class having the maximum probability.

In this work, a pre-trained DenseNet121 architecture is employed, trained on the ImageNet data set, which illustrates transfer learning. This architecture can be fine-tuned for use in bone fracture detection, hence efficiently learning relevant features from the domain. Generally, the DenseNet121 architecture gives a good performance in bone fracture classification, since it utilizes the concept of feature reusability and learning from data patterns

D. Fracture Detection Using YOLO

The fracture detection component of the proposed system is implemented using the YOLO model, which is a real-time object detection algorithm. Unlike classification models that only predict the presence of a fracture, YOLO is capable of identifying the exact location of the fracture within the X-ray image.

YOLO works by dividing the input image into a grid and predicting bounding boxes along with confidence scores for each region. In this work, the model is trained to detect fracture regions and generate bounding boxes around the affected areas. This enables the system to provide precise spatial information about the fracture.

The YOLO model outputs the bounding box coordinates, confidence score, and localized fracture region. All this contributes to the efficient localization of the fracture regions within the x-ray image. The combination of the YOLO and classification model improves the functionality of the system.

E. Explainability Using Grad-CAM

To increase the explainability of the output of the model, the Grad-CAM method for visualizing activations in the form of a heatmap of relevant pixels is employed. Grad-CAM will be able to illustrate important regions of the X-ray image that affected the model prediction.

In general, Grad-CAM involves finding gradients of the target class with regard to feature maps produced by the last convolutional layer of the model. These gradients help create a heatmap illustrating regions of importance of the image related to the target class. In the proposed solution, the Grad-CAM technique will be employed for the DenseNet121 model in order to show relevant fracture regions in the X-ray image.
Grad-CAM provides visual explanations by generating heatmaps that highlights important regions in the X-ray images.

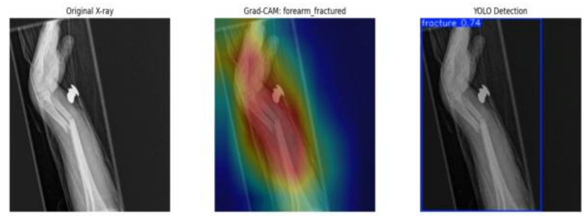
IV. RESULT AND DISCUSSION

The effectiveness of the proposed system for the detection of bone fracture was analyzed by testing the X-ray images. The system incorporates DenseNet121 for classification, YOLO for detection, and Grad-CAM for localization and visualization of important areas.

Figure. 4 represents some output of the proposed system. With DenseNet121, the X-ray images were classified based on whether the image contains a fracture or not. For the fractured images, the classification results show that there is a presence of bone fracture. In contrast, the normal images indicate the absence of bone fractures.
The visualization technique of Grad-CAM depicts the important areas in the X-ray image based on which the prediction was made. It is seen from the output that the heat map emphasizes the fracture part of the X-ray image.

Moreover, the YOLO method is also able to localize the fracture region by creating a bounding box for the detected fracture. This gives spatial localization for the fracture, making the system more interpretable.
The experiment shows that the system designed is able to classify as well as detect the region where the fracture is located. Thus, this makes the system accurate, reliable, and feasible enough to assist in the diagnosis of bone fractures by health-care professionals.
As illustrated in Fig. 4, the designed system is able to detect fracture regions and provide a relevant explanation of it using the Grad-CAM and YOLO methods. The model demonstrates consistent performance across both fractured and non-fractured cases, indicating its robustness.

Fracture Case



Normal Case

Figure 4: Sample Results of the Proposed Bone Fracture Detection

The first row represents fractured cases showing the original X-ray, Grad-Cam visualization, and YOLO-based fracture localization. The second row represents a non-fractured case where no fracture is detected.

V. CONCLUSION

Bone fracture detection using a combination of deep learning models presents an opportunity to show the success of deep learning methods in medical image analysis. Through using the DenseNet121 model alongside transfer learning and fine-tuning, the detection system is able to classify X-ray images into fracture and no-fracture classes with considerable accuracy. Using the YOLO model also allows the model to localize fracture regions in the images, whereas the use of Grad-CAM makes it possible to get visual explanations of the decisions made by the model.

Preprocessing of the data is done through the application of various methods, including resizing and normalization. Data augmentation was also employed to improve model performance. Based on the experimental results, the proposed model achieved an accuracy of around 80%, and other metrics, including F1-score, sensitivity, and ROC-AUC, were also very favourable. This system can be further extended for real-time clinical applications.

VI. FUTURE WORK

In terms of future improvements of the presented system, one can increase the volume of the database and include more diverse data samples. Specifically, a dataset comprising various fracture cases, imaging differences, and other anatomical areas might allow enhancing the performance of the classifier. Furthermore, it is also reasonable to apply advanced neural networks, including ResNet and EfficientNet, instead of the currently used DenseNet121.

Another possible direction for improving the performance of the presented system is improving the detection algorithm. In particular, it is recommended to optimize the YOLO model and develop a method for identifying fractures in images using other approaches to object detection. In addition, it may be helpful to develop more sophisticated techniques of preprocessing and data augmentation to improve generalization on the test set. It would also be beneficial to integrate the system into an actual application, which could be run online, for example. Finally, one may consider using more advanced explainable AI approaches beyond Grad-CAM.

REFERENCES

[1] K. Sushma *et al.*, "Enhanced bone fracture detection through deep learning-based multi-scale feature fusion," in *Proc. Int. Conf. Advancements Power, Commun. Intell. Syst. (APCI)*, IEEE, 2025.

[2] H. P. Nguyen *et al.*, "A deep learning based fracture detection in arm bone X-ray images," in *Proc. Int. Conf. Multimedia Anal. Pattern Recognit. (MAPR)*, IEEE, 2021.

[3] K. D. Ahmed and R. Hawezi, "Detection of bone fracture based on machine learning techniques," *Measurement: Sensors*, vol. 27, 2023.

[4] F. Mohammad *et al.*, "Block-Deep: A hybrid secure data storage and diagnosis model for bone fracture identification of athlete from X-ray and MRI images," *IEEE Access*, vol. 11, 2023.

[5] V. Lam *et al.*, "Analyzing pediatric forearm X-rays for fracture analysis using machine learning," *Int. J. Comput. Assist. Radiol. Surg.*, 2025.

[6] Alam *et al.*, "Novel transfer learning-based bone fracture detection using radiographic images," *BMC Med. Imaging*, 2025.

[7] B. Pujitha *et al.*, "Detection of bone fracture using deep learning," in *Proc. ESIC*, 2024.

[8] Tieu *et al.*, "The role of artificial intelligence in the identification and evaluation of bone fractures," *Bioengineering*, vol. 11, 2024.