Bone Fracture Detection in X-ray images using Convolutional Neural Networks

[1]Pilla Sanjana, [2]Dr. M. Ramjee
[1]Student, [2]Professor
[1]Department of Computer Science and Systems Engineering,
[2]Department of Information Technology and Computer Applications
Andhra University College of Engineering, Visakhapatnam, Andhra Pradesh, India.

Abstract: Bone fractures are a prevalent form of musculoskeletal injury that require timely and accurate diagnosis for effective treatment. Radiographic imaging, particularly X-ray analysis, remains the primary diagnostic tool. However, manual interpretation by radiologists is subject to human error, fatigue, and variability in judgment. This research presents a deep learning-based approach for the automated detection of bone fractures in X-ray images using Convolutional Neural Networks (CNNs). The proposed system is trained on a publicly available dataset comprising labeled images of fractured and non-fractured bones. Preprocessing techniques such as resizing, normalization, and data augmentation were applied to improve model robustness and generalization. The CNN architecture was designed and optimized to learn distinguishing features from input images without manual feature engineering. The model achieved a high classification accuracy of over 93% on test data, demonstrating strong potential for assisting clinical diagnosis. Evaluation metrics including precision, recall, and F1-score indicate that the model can reliably differentiate between fractured and healthy bone structures. The system is scalable, cost-effective, and suitable for integration into computer-aided diagnostic tools, particularly in resource-limited settings. This study contributes toward the development of intelligent diagnostic systems that can support healthcare professionals by reducing diagnostic delays and enhancing patient outcomes.

Keywords: X-ray images, Convolutional Neural Networks (CNN), Deep Learning, Computer-aided detection (CAD).

I. INTRODUCTION

Bone fractures represent one of the most frequently encountered injuries in both emergency and orthopedic care settings. Timely and precise detection of fractures is essential to ensure effective treatment and to prevent complications such as chronic pain, deformities, or impaired mobility. The primary diagnostic modality for bone fractures remains radiographic imaging, particularly X-ray scans, due to their accessibility, speed, and cost-effectiveness. However, the manual interpretation of X-ray images by radiologists is not without limitations. Factors such as image complexity, subtle fracture patterns, fatigue, and varying levels of clinical experience can contribute to diagnostic errors or delays.

In recent years, artificial intelligence (AI) and deep learning techniques have shown considerable promise in the field of medical image analysis. Among these, Convolutional Neural Networks (CNNs) have demonstrated exceptional performance in visual recognition tasks, including object detection, classification, and segmentation. CNNs have the distinct advantage of automatically learning spatial hierarchies and patterns from raw image data without the need for handcrafted feature extraction, making them particularly suitable for analyzing complex medical images such as X-rays.

This study aims to harness the power of CNNs to develop an automated system for detecting bone fractures from X-ray images. The proposed system is trained and evaluated on a curated dataset containing labeled images of fractured and non-fractured bones. The methodology involves image preprocessing, CNN model design, training using GPU-accelerated computing, and performance evaluation using standard classification metrics. By automating fracture detection, the system can provide rapid, consistent, and objective assessments, thereby reducing the burden on radiologists and enhancing diagnostic workflows. Furthermore, such AI-driven tools have significant potential for deployment in rural or underserved healthcare environments where access to specialized radiological expertise may be limited. The integration of deep learning into clinical decision support systems offers an opportunity to bridge the gap between demand and availability of diagnostic services, ultimately contributing to improved patient care and outcomes.

II. Literature Review

The integration of artificial intelligence (AI) into medical diagnostics has transformed healthcare by enabling faster and more accurate interpretation of clinical data. In the domain of radiology, computer vision—especially deep learning—has shown promising results in automating the detection of various abnormalities, including bone fractures.

Earlier computer-aided detection (CAD) systems for bone fractures relied on classical image processing techniques such as edge detection, thresholding, and histogram analysis. While these approaches offered some degree of automation, they were highly sensitive to noise, illumination, and image quality, often resulting in suboptimal accuracy and limited adaptability. Moreover, such methods required expert-defined features, which posed challenges when applied to large-scale or diverse datasets.

With the advent of deep learning, particularly Convolutional Neural Networks (CNNs), significant improvements have been realized in the field of medical image classification. CNNs are capable of learning hierarchical representations directly from raw pixel data, thereby eliminating the need for manual feature engineering. This makes them particularly suitable for analyzing complex and high-dimensional data such as X-ray images.

Several studies have explored the use of CNNs in fracture detection. For instance, Olczak et al. demonstrated that deep learning models could match or even exceed human radiologists in detecting wrist and hand fractures. Similarly, Rajpurkar et al. introduced CheXNet, a deep CNN that achieved radiologist-level performance in identifying pneumonia from chest X-rays—paving the way for broader applications of CNNs in radiographic analysis. Other researchers have implemented transfer learning techniques by leveraging pre-trained models such as VGG16, ResNet, and InceptionV3. These architectures, when fine-tuned on medical datasets, have yielded high classification accuracies with relatively fewer training images. Badgeley et al. further extended the use of deep learning to predict not just the presence of fractures but also patient-specific variables such as bone age.

More recent works have incorporated attention mechanisms and heatmap visualizations (e.g., Grad-CAM) to make AI predictions more interpretable to clinicians. This has addressed one of the critical challenges in clinical AI adoption: trust and explainability. These interpretability techniques help localize the fracture regions on the image, providing radiologists with visual cues and enhancing confidence in the system's outputs.

Despite the advancements, there are still challenges related to dataset quality, image imbalance, generalizability, and integration into clinical workflows. Nonetheless, the collective evidence from the literature confirms that CNNs provide a powerful framework for building scalable, accurate, and robust fracture detection systems.

The present study builds upon these advancements by designing a customized CNN architecture trained on a balanced dataset of X-ray images categorized into fractured and non-fractured classes. The model is evaluated across various metrics to validate its clinical applicability and potential for deployment in real-world settings.

III. Problem Statement

Bone fractures are among the most common medical conditions encountered in emergency and trauma care. Accurate and timely diagnosis is essential to ensure effective treatment and reduce the risk of long-term complications. In current clinical settings, fracture detection relies heavily on the manual examination of radiographic (X-ray) images by trained radiologists. While radiographs are cost-effective and widely available, their interpretation can be subjective and prone to error, particularly in high-volume or resource-limited environments. Factors such as poor image quality, subtle fracture patterns, and clinician fatigue further increase the likelihood of misdiagnosis or delayed diagnosis.

The need for a reliable, automated system that can assist healthcare professionals in interpreting X-ray images has become increasingly evident. Traditional image processing techniques and rule-based systems offer limited generalization and lack the flexibility to adapt to diverse anatomical regions and varying image conditions. Moreover, manual feature extraction approaches are not only time-consuming but also require significant domain expertise, making them unsuitable for large-scale deployment.

In light of these limitations, deep learning—particularly Convolutional Neural Networks (CNNs)—has emerged as a promising solution. CNNs have the ability to learn complex visual patterns directly from raw images, reducing the dependency on handcrafted features and enhancing classification performance. However, challenges remain in designing an efficient model that performs well across diverse fracture types, manages class imbalance, and integrates seamlessly into clinical workflows.

This research aims to address the problem by developing a CNN-based model for binary classification of X-ray images into fractured and not fractured categories. The goal is to create a system that not only achieves high diagnostic accuracy but also maintains computational efficiency, interpretability, and scalability for real-world healthcare applications.

IV. Proposed System

The system commences by meticulously compiling a diverse dataset comprising X-ray images depicting normal bone structures alongside various fracture types. Careful curation and augmentation techniques ensure dataset uniformity and bolster the model against potential biases. This step aims to create a robust dataset representative of real-world scenarios, essential for the model's versatility across different fractures and imaging conditions.

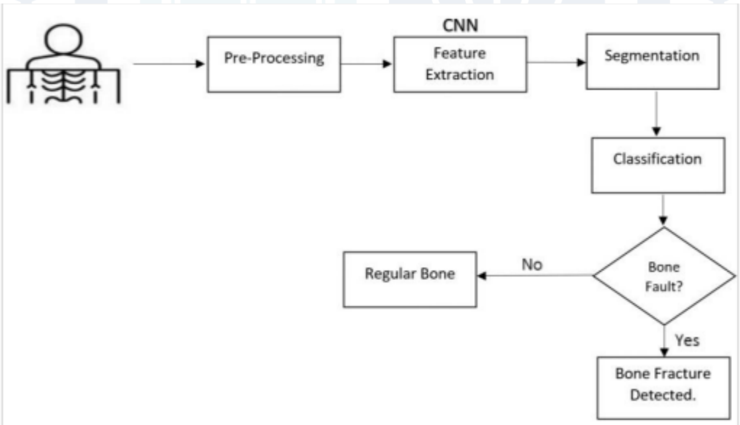


Fig. Block Diagram for Bone Fracture Detection System

V. Methodology

The proposed methodology follows a systematic approach to develop and evaluate a deep learning model for detecting bone fractures in X-ray images. The entire workflow consists of multiple stages: dataset preparation, preprocessing, model architecture design, training, and evaluation.

5.1. Dataset Collection

The dataset used in this study was obtained from Kaggle's publicly available repository, comprising X-ray images of fractured and non-fractured bones. The dataset is organized into three subsets:

- Training set: 4,606 fractured, 4,640 non-fractured images
- Validation set: 337 fractured, 492 non-fractured images
- Test set: 238 fractured, 268 non-fractured images

This split ensures a balanced distribution across classes and enables unbiased model evaluation. The images were categorized into separate directories based on their class labels to facilitate use with Keras' flow_from_directory() function.

5.2. Data Preprocessing

To ensure model efficiency and generalization, several preprocessing steps were applied:

- Resizing: All images were resized to 224×224 pixels to match the input requirements of the CNN.

- Normalization: Pixel values were rescaled to the [0, 1] range using rescale=1./255.
- Data Augmentation: Applied to the training data to improve robustness and reduce overfitting. Techniques included:
    - Random rotations (up to 20°)
    - Horizontal flipping
    - Width and height shifts (10%)
    - Zoom (15%)
    - Nearest neighbor filling for edge gaps

Separate data generators were created for training, validation, and test sets. Augmentation was applied only to the training generator, while the others were rescaled.

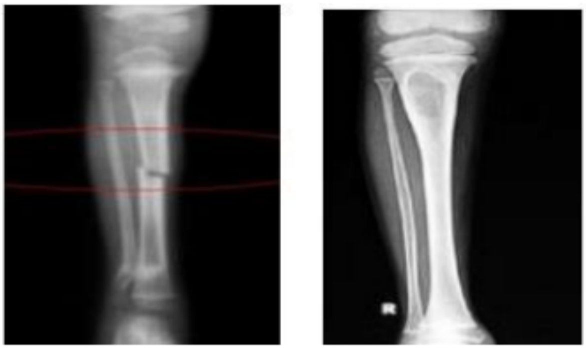


Fig. Normal image vs Preprocessed image

5.3. Model Architecture

The CNN model was built using the Keras Sequential API. The architecture consists of:

- Three Convolutional Blocks:
    - Each with a Conv2D layer followed by MaxPooling2D
    - Filter sizes: 32, 64, and 128 respectively
    - Kernel size: 3×3
    - Activation: ReLU
- Flatten Layer: Converts feature maps into a 1D vector
- Dense Layer: 128 neurons with ReLU activation
- Dropout Layer: 50% rate to prevent overfitting
- Output Layer: A single neuron with sigmoid activation for binary classification

The model accepts input of shape (224, 224, 3), reflecting RGB images.

5.4. Model Training

The model was compiled with the Adam optimizer, using:

- Loss function: Binary Crossentropy
- Evaluation metric: Accuracy

Training was conducted over 20 epochs with early stopping and model checkpointing enabled. These callbacks were used to monitor validation loss and save the best performing model weights automatically.

5.5. Evaluation metrics

After training, the model was evaluated on the test dataset using:

- Accuracy – overall classification performance
- Precision & Recall – performance on fractured and non-fractured cases
- F1-Score – balance between precision and recall
- Confusion Matrix – detailed class-wise prediction breakdown

These metrics were calculated using Scikit-learn's evaluation tools to ensure interpretability and standardization.

The CNN model is as follows:

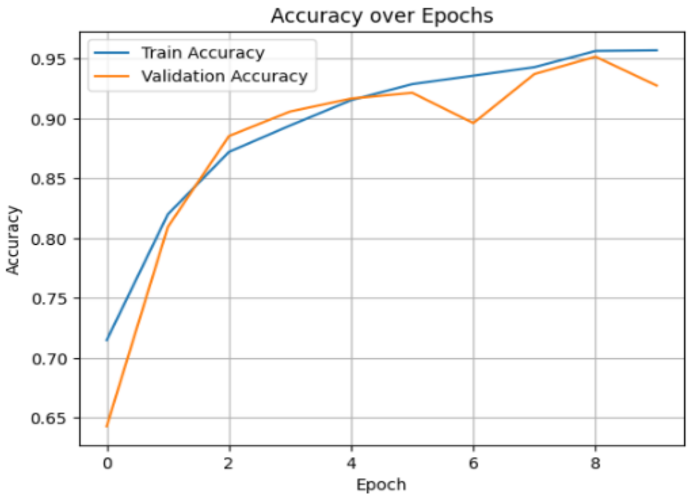
Model: "sequential"

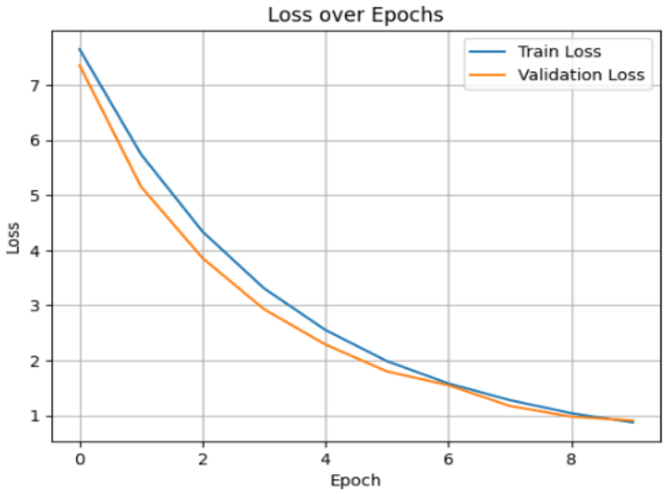
Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 222, 222, 32)	320
batch_normalization (BatchNormalization)	(None, 222, 222, 32)	128
max_pooling2d (MaxPooling2D)	(None, 111, 111, 32)	0
conv2d_1 (Conv2D)	(None, 109, 109, 64)	18,496
dropout (Dropout)	(None, 109, 109, 64)	0
max_pooling2d_1 (MaxPooling2D)	(None, 54, 54, 64)	0
conv2d_2 (Conv2D)	(None, 54, 54, 128)	73,856
batch_normalization_1 (BatchNormalization)	(None, 54, 54, 128)	512
max_pooling2d_2 (MaxPooling2D)	(None, 18, 18, 128)	0
flatten (Flatten)	(None, 41472)	0
dense (Dense)	(None, 256)	10,617,088
dropout_1 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 128)	32,896
dropout_2 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 1)	129

Total params: 10,743,425 (40.98 MB)
Trainable params: 10,743,105 (40.98 MB)
Non-trainable params: 320 (1.25 KB)

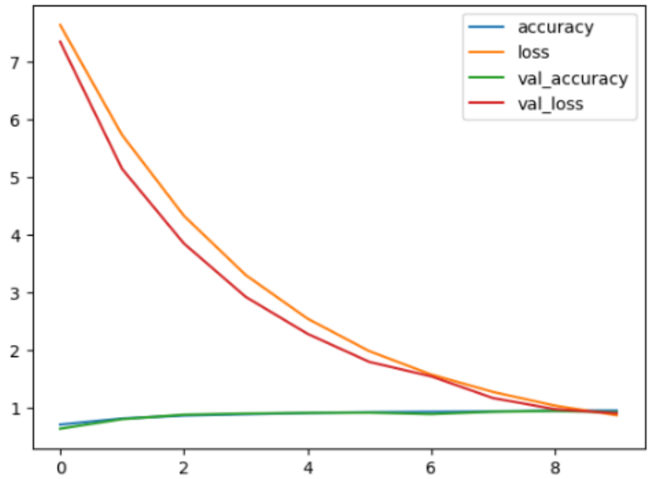
VI. Results and Discussion

The final results are:
Final Train Accuracy: 95.71%
Final Train Loss: 0.8763
Final Validation Accuracy: 92.76%
Final Validation Loss: 0.9100

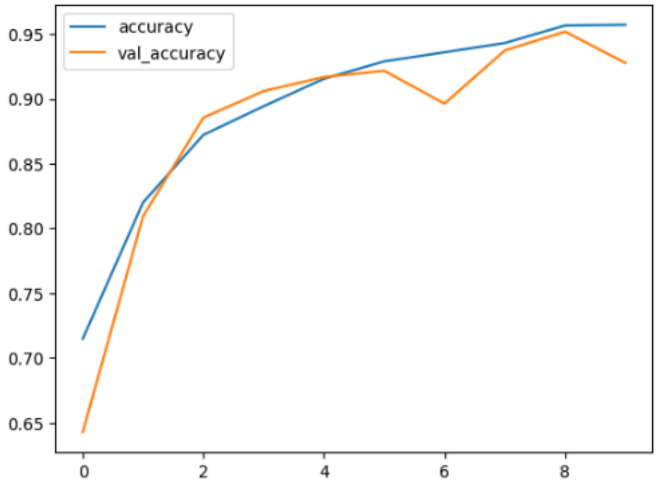




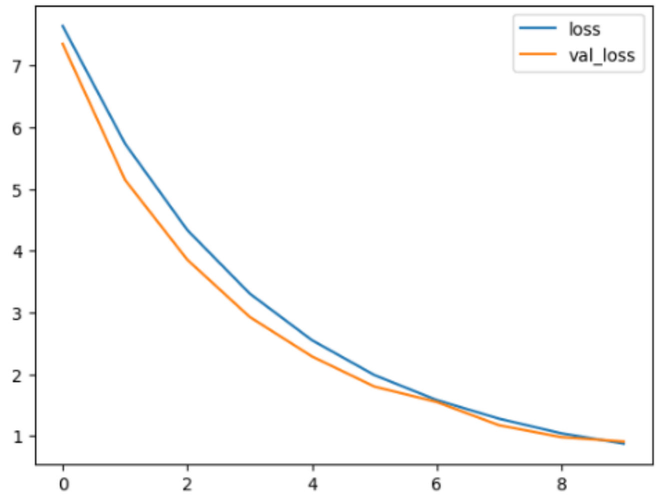
<Axes: >



<Axes: >



<Axes: >




```
16/16 ━━━━━━━━━━━━━━━━━━━━ 4s 222ms/step - accuracy: 0.9589 - loss: 0.8042
Test Accuracy: 93.68% , Test Loss: 0.8647
```

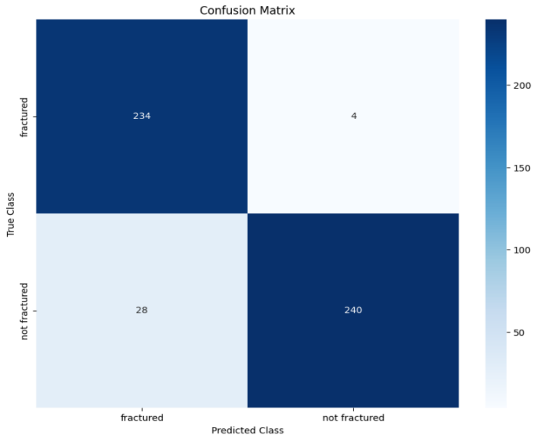
Classification report and confusion matrix:

```
              precision    recall  f1-score   support

   fractured       0.89      0.98      0.94       238
not fractured      0.98      0.90      0.94       268

    accuracy                           0.94       506
   macro avg       0.94      0.94      0.94       506
weighted avg       0.94      0.94      0.94       506
```

Test Accuracy: 93.68%
Test Loss: 0.8647

For a sample image:

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 29ms/step
Prediction: ✅ Fractured (Confidence: 0.6569)
```



VII. Conclusion

This study presents an effective and automated approach for detecting bone fractures in X-ray images using Convolutional Neural Networks. By leveraging deep learning, the proposed system eliminates the need for manual feature extraction and achieves high diagnostic accuracy. Through careful preprocessing, balanced dataset usage, and a well-structured CNN architecture, the model demonstrated strong performance in identifying fractured and non-fractured cases. The results highlight the potential of CNN-based models to assist healthcare professionals by providing consistent, rapid, and reliable fracture detection. This work serves as a foundation for developing AI-driven diagnostic tools that can be integrated into clinical workflows, particularly in areas with limited access to radiological expertise.

VIII. Future Scope

While the current model demonstrates promising results, there is substantial scope for further enhancement. Future work can focus on incorporating multi-class classification to identify different types and severities of fractures. Integrating localization techniques such as Grad-CAM can improve interpretability by highlighting the exact fracture regions on the X-ray. Additionally, transfer learning using pre-trained models like ResNet or EfficientNet could boost performance on more complex datasets. The system can also be extended into real-time applications through deployment in mobile or web-based platforms, enabling rapid diagnostics in emergency care and rural clinics. Expanding the dataset to include pediatric and geriatric cases may also improve the model's generalizability across patient demographics.

References

[1] Edward V, Cephas Paul. and Hepzibah S, Hilda. (2015) , ―A Robust Approach For Detection of the type of Fracture from XRay Images‖, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue No. 3, pp.479-482.

[2] Deshmukh, Snehal., Zalte, Shivani., Vaidya, Shantanu. and Tangade, Parag. (2015), ―Bone Fracture Detection Using Image Processing in Matlab‖, International Journal of Advent Research in Computer and Electronics.

[3] Rathode, Hs. and Ali, Wahid. (2015), ―MRI Brain Image Quantification Using artificial neural networks – A Review Report‖, ISOI Journal of Engineering and Computer science, Vol. 1, Issue No. 1, pp. 48-55.

[4] O. Bandyopadhyay, A. Biswas, and B. B. Bhattacharya, "Classification of long-bone fractures based on digitalgeometric analysis of X-ray images," Pattern Recognit. Image Anal., vol. 26, no. 4, pp. 742–757, 2016.

[5] U. Andayani et al., "Identification tibia and fibula bone fracture location using scanline algorithm," J. Phys. Conf. Ser., vol. 978, p. 012043, 2018.

[6] N. D. Deokar and D. A. G. Thakur, "Design , Development and Analysis of Femur Bone by using Rapid Prototyping," Int. J. Eng. Dev. Res., vol. 4, no. 3, pp. 881– 886, 2016.

[7] M. Khan, S. P. S. M. A. Sirdeshmukh, and K. Javed, "Evaluation of bone fracture in animal model using bioelectrical impedance analysis," Perspect. Sci., vol. 8, pp. 567–569, 2016.
[8] G. Narayana, "Automatic Detection of Radius of Bone Fracture," Int. Res. J. Eng. Technol., vol. 4, no. 6, pp. 4–7, 2017.
[9] I. Khatik, "A Study of Various Bone Fracture Detection Techniques," Int. J. Eng. Comput. Sci., vol. 6, no. 5, pp. 6– 11, 2017.
[10] H. Kaur and A. Jain, "Detection of Fractures in Orthopedic X-Ray Images," vol. 8, no. 3, pp. 545– 551, 2017.
[11] Zeelan Basha, C. M. A. K., Maruthi Padmaja, T., Balaji, G. N, "EAI Endorsed Transactions on Pervasive Health and Technology", Vol. 5 ,Issue 18, 1-6,2019.
[12] P.Ganesan, S.Sivakumar, and S.Sundar, A Comparative Study on MMDBM Classifier Incorporating Various Sorting Procedure, "Indian Journal of Science and Technology" Vol 8(9), 868–874, May 2015.