International Journal of Artificial Intelligence in Medical Issues

Volume 3 Issue 1
ISSN 3025-4167
https://doi.org/10.56705/ijaimi.v3i1.249

Research Article

Classification of Multi-Region Bone Fractures from X-ray Images Using Transfer Learning with ResNet18

Rasni Alex [1,*]; Rosmasari [2]
[1] Universitas Mulawarman, Samarinda, Kalimantan Timur 75119, Indonesia, rasnialex@ft.unmul.ac.id
[2] Universitas Mulawarman, Samarinda, Kalimantan Timur 75119, Indonesia, rosmasari@unmul.ac.id

Correspondence should be addressed to Rasni Alex; rasnialex@ft.unmul.ac.id

Received 02 February 2025; Revised 06 February 2024; Accepted 25 April 2024; Published 30 May 2025

Abstract:
Fracture detection in radiographic images is a critical task in orthopaedic diagnostics, often requiring timely and accurate interpretation by medical professionals. However, manual evaluation of X-rays is time-consuming and prone to subjective bias. This study proposes an automated deep learning approach for binary classification of bone fractures using a pre-trained ResNet18 architecture. The model was trained and validated on a multi-region X-ray dataset consisting of 10,580 images categorized into fractured and non-fractured classes. To improve generalization, data augmentation techniques such as rotation and horizontal flipping were applied during pre-processing. The final model achieved a validation accuracy of 97.59%, with high true positive and true negative rates as confirmed by the confusion matrix analysis. The results demonstrate the effectiveness of transfer learning in handling radiographic image classification tasks while maintaining computational efficiency. This research contributes to the development of reliable and scalable computer-aided diagnostic tools that can support clinical decision-making, especially in environments with limited resources.

Keywords: Bone fracture classification, Deep learning, Medical imaging, ResNet18, Transfer learning.

Dataset link: https://www.kaggle.com/datasets/bmadushanirodrigo/fracture-multi-region-x-ray-data

1. Introduction

Bone fractures are among the most common orthopaedic conditions requiring timely and accurate diagnosis for effective treatment. Manual interpretation of radiographic (X-ray) images by radiologists can be time-consuming and subject to inter-observer variability, especially when processing large volumes of data. In response to these challenges, the integration of artificial intelligence—particularly deep learning—has emerged as a promising approach to automate and enhance fracture detection from radiographic images with consistent performance.

Recent advancements in deep learning technologies have significantly transformed diagnostic radiology, particularly in the classification of bone fractures using X-ray images. Numerous studies have demonstrated the effectiveness of transfer learning models such as ResNet, DenseNet, and Vision Transformers in detecting and classifying multi-location fractures with high accuracy [1]. For instance, a lightweight DenseNet-based framework has achieved an accuracy of 90.3% in sports fracture detection while maintaining computational efficiency [2]. Additionally, models such as ResNet152V2 and DenseNet201 have shown promising results when applied to multi-region X-ray datasets, demonstrating strong capability in distinguishing various fracture types [1]. Furthermore, hybrid optimization strategies like particle swarm optimization combined with transfer learning have been employed in medical image classification—including X-ray and CT scans—to boost classification accuracy while reducing false-positive rates [3].

Despite these advances, most existing studies are limited to specific anatomical regions or fracture types, such as the wrist, hip, or knee. Moreover, complex model architectures often require significant computational resources, making them impractical for deployment in real-time clinical environments with constrained infrastructure. These limitations highlight a research gap in developing efficient yet accurate models for broad, multi-region fracture classification from radiographs.

To address this gap, the present study proposes a deep learning approach based on a pre-trained ResNet18 model for binary classification of fractured and non-fractured conditions across various anatomical regions [4], [5]. Leveraging a large publicly available multi-region X-ray dataset, along with data augmentation techniques and transfer learning, this research aims to evaluate the model's performance in terms of accuracy, generalizability, and computational efficiency. The goal is to offer a practical and scalable solution for automated fracture classification in clinical decision support systems.

2. Method

Research Design:

This study employs an experimental research design that utilizes a supervised learning approach through convolutional neural networks (CNNs) for binary classification of bone fractures from X-ray images. The primary focus is on evaluating the effectiveness of the ResNet18 architecture using transfer learning to distinguish between fractured and non-fractured bone conditions across multiple anatomical regions [6], [7].

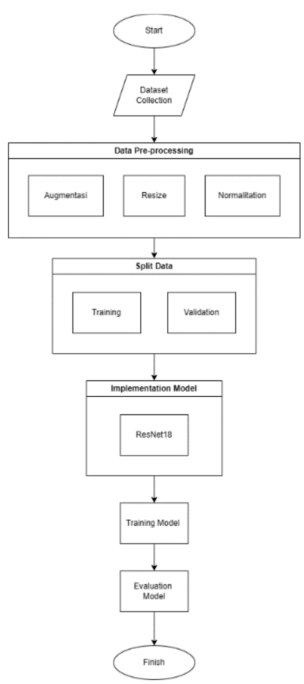

Figure 1: Research Workflow

Figure 1 illustrates the workflow of the proposed fracture classification system, starting from data collection to model evaluation. The process begins with the acquisition of X-ray image data, followed by a series of pre-processing steps including image resizing, normalization, and data augmentation to enhance model generalization. The dataset is then divided into training and validation subsets. The core of the model implementation involves the use of a pre-trained ResNet18 architecture, which is fine-tuned on the training data [8]–[10]. Finally, the model is evaluated based on classification performance metrics such as accuracy and confusion matrix analysis.

Dataset and Pre-processing:

The dataset used in this study consists of 10,580 radiographic images sourced from a publicly available dataset on Kaggle. **Figure 2** presents a selection of sample X-ray images from the dataset, illustrating both fractured and non-fractured conditions. These images represent various anatomical regions and highlight the visual variability and complexity inherent in the dataset. The distinction between classes is not always visually apparent, which emphasizes the need for robust deep learning techniques to accurately classify fracture status across diverse bone structures.

Fractured | Fractured | Not Fractured | Not Fractured

Figure 2. Sample Images from the Dataset

Each subset includes two categories: *fractured* and *non-fractured*. The data pre-processing pipeline includes:

a Resizing images to 224×224 pixels [11].

b Data Augmentation applied only to the training set to improve model generalization [12]:

- Random horizontal flipping
- Random rotation up to 10 degrees

c Normalization using ImageNet mean and standard deviation values:

$$mean = [0.485, 0.456, 0.406], \quad \text{std} = [0.229, 0.224, 0.225] \tag{1}$$

Model Architecture:

We utilized the ResNet18 architecture pre-trained on ImageNet [13]–[15]. The final fully connected (FC) layer of the network was replaced to match the number of target classes (2 classes: fractured and non-fractured). The modified FC layer is defined as:

$$fc_{out} = Linear(512,2) \tag{2}$$

where 512 is the number of features from the last convolutional layer [16], [17].

Training Procedure

The model was trained for 10 epochs using the Adam optimizer with a learning rate of 0.001. The loss function used was Cross-Entropy Loss [18], [19], suitable for multi-class classification tasks:

$$\mathcal{L}(y, \hat{y}) = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \quad (2)$$

Where y_i is the true label, $\hat{y}_i$ is the predicted probability, and C is the number of classes (2 in this case). Training was conducted using a batch size of 32 and a GPU-enabled environment for computational efficiency. The best-performing model during validation was saved based on the highest validation accuracy.

Evaluation Metrics

To evaluate the model's performance, several metrics were employed:

Accuracy: The proportion of correct predictions over the total number of samples, calculated as [20], [21]:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (3)$$

where TP: True Positive, TN: True Negative, FP: False Negative, FN: False Negative.

Confusion Matrix: Used to visualize the distribution of predictions and identify misclassification patterns between the two classes.

Validation Accuracy per Epoch: Tracked during training to monitor learning progress and avoid overfitting.

3. Result and Discussion

Results

The performance of the ResNet18-based model was evaluated across 10 training epochs using a binary classification approach. During the training process, the model consistently showed improved performance in both training and validation stages. The training loss decreased significantly from the first epoch to the last, indicating effective learning and convergence. The training and validation accuracy metrics across each epoch are summarized in **Table 1**, which provides a comprehensive view of the model's learning progression.

Table 1. Training and Validation Performance per Epoch

Epoch	Training Loss	Training Accuracy	Validation Accuracy
1	0.2619	0.8997	0.8721
2	0.0896	0.9697	0.9385
3	0.0713	0.9756	0.9361
4	0.0508	0.9822	0.9373
5	0.0407	0.9849	0.9192
6	0.0442	0.9846	0.9505
7	0.0369	0.9875	0.9493
8	0.0327	0.9879	**0.9759**
9	0.0315	0.9882	0.9505
10	0.0336	0.9871	0.9723

Table 1 presents the training loss, training accuracy, and validation accuracy recorded during each epoch. The model achieved a rapid increase in accuracy during the early epochs, with the highest validation accuracy of 97.59% recorded in epoch 8, which was subsequently selected as the best model for evaluation. These results demonstrate that the model was able to generalize well without overfitting, as indicated by the consistency between training and validation accuracies.

To further evaluate the performance of the model on the validation set, a confusion matrix was constructed. **Figure 3** illustrates the confusion matrix for the best-performing model. This matrix reveals the number of correctly and incorrectly classified samples across the two classes: fractured and non-fractured. The model correctly identified 327 fractured and 479 non-fractured images, with only 23 total misclassifications. These results reflect a high level of predictive accuracy.

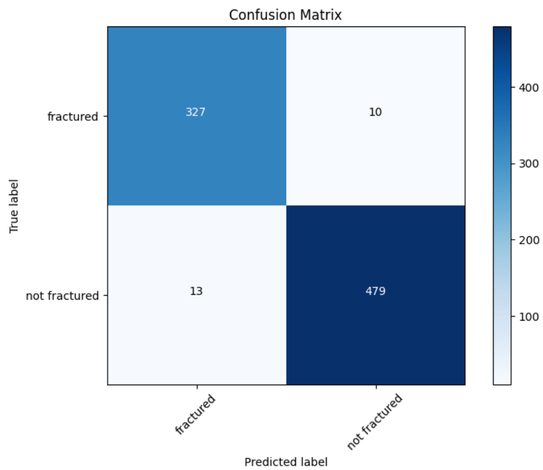

Figure 3. Confusion Matrix of the Best Model.

Discussion

The results of this study confirm that transfer learning using the ResNet18 architecture is highly effective for binary classification of bone fractures in multi-region X-ray images. The model demonstrated excellent performance across all training epochs, with a final validation accuracy of 97.59%, placing it among the top-performing approaches in recent literature. The decreasing trend in training loss, paired with consistent validation accuracy, suggests that the model learned meaningful features from the radiographic data without overfitting.

The confusion matrix, as shown in Figure 1, provides deeper insight into the model's classification capabilities. The high number of true positives and true negatives indicates that the model was effective in distinguishing between fractured and non-fractured cases. The low number of false positives (13) and false negatives (10) further supports the reliability of the classification system. These results are particularly notable given the anatomical variability present in the multi-region dataset used.

Compared to previous studies using deeper and more complex models such as ResNet152V2 or DenseNet201, the proposed approach using ResNet18 achieved comparable or even superior accuracy while maintaining computational efficiency. This is especially beneficial for practical deployment in clinical environments with limited processing power. Nevertheless, the presence of some misclassified images indicates that further improvements are possible. These misclassifications may have occurred due to overlapping visual features between classes, low image quality, or ambiguous fracture patterns. Future research may explore the integration of attention mechanisms, localization-based pre-processing, or multi-model ensemble techniques to enhance accuracy and interpretability.

4. Conclusion

This study demonstrated the effectiveness of a transfer learning approach using the ResNet18 architecture for binary classification of bone fractures in multi-region radiographic images. By leveraging a publicly available dataset comprising over 10,000 X-ray images from various anatomical regions, the model achieved a high validation accuracy of 97.59%, with low false-positive and false-negative rates. The model was trained using minimal architectural modifications and benefited from standard data augmentation techniques, resulting in both accurate and computationally efficient performance.

The experimental results confirmed that even with a relatively lightweight architecture such as ResNet18, high classification performance can be achieved without sacrificing generalization capabilities. The confusion matrix analysis further reinforced the model's ability to correctly distinguish fractured from non-fractured cases across diverse anatomical areas, supporting its applicability in real-world diagnostic workflows.

In addressing the initial research objective—to develop an efficient and accurate automated fracture classification system—this study successfully established a foundation for clinical decision support tools that can assist radiologists in expediting diagnostic processes. The proposed approach contributes to the growing body of research advocating the use of deep learning in medical imaging, particularly in fracture detection tasks.

For future work, enhancements may focus on expanding the classification task to multiple fracture types or anatomical regions, incorporating explainable AI methods to improve model transparency, and evaluating the model's performance in real-time clinical settings. Furthermore, integrating multimodal data sources, such as clinical reports or CT scans, may offer additional insights and improve diagnostic reliability.

References:

[1] T. Aldhyani *et al.*, "Diagnosis and detection of bone fracture in radiographic images using deep learning approaches," *Front. Med.*, vol. 11, Jan. 2025, doi: 10.3389/fmed.2024.1506686.

[2] A. Abdusalomov *et al.*, "Lightweight Deep Learning Framework for Accurate Detection of Sports-Related Bone Fractures," *Diagnostics*, vol. 15, no. 3, p. 271, Jan. 2025, doi: 10.3390/diagnostics15030271.

[3] H. Liu, M. Zhao, C. She, H. Peng, M. Liu, and B. Li, "Classification of CT scan and X-ray dataset based on deep learning and particle swarm optimization," *PLoS One*, vol. 20, no. 1, p. e0317450, Jan. 2025, doi: 10.1371/journal.pone.0317450.

[4] X. Ye, "Image Binary Classification of Coronary Artery Stenosis Based on Resnet18 and Transfer Learning,"

Proceedings of the 35th Chinese Control and Decision Conference, CCDC 2023. pp. 2721–2726, 2023, doi: 10.1109/CCDC58219.2023.10326590.

[5] R. A. Al-Falluji, “Automatic detection of COVID-19 using chest X-ray images and modified resnet18-based convolution neural networks,” *Comput. Mater. Contin.*, vol. 66, no. 2, pp. 1301–1313, 2020, doi: 10.32604/cmc.2020.013232.

[6] M. Odusami, “Analysis of features of alzheimer’s disease: Detection of early stage from functional brain changes in magnetic resonance images using a finetuned resnet18 network,” *Diagnostics*, vol. 11, no. 6, 2021, doi: 10.3390/diagnostics11061071.

[7] Y. Zhang, “Classification of mutton location on the animal using improved ResNet18 network model and mobile application,” *Nongye Gongcheng Xuebao/Transactions Chinese Soc. Agric. Eng.*, vol. 37, no. 18, pp. 331–338, 2021, doi: 10.11975/j.issn.1002-6819.2021.18.038.

[8] S. Liu, “Optimized Modified ResNet18: A Residual Neural Network for High Resolution,” *2024 IEEE 4th International Conference on Electronic Technology, Communication and Information, ICETCI 2024*. pp. 1–5, 2024, doi: 10.1109/ICETCI61221.2024.10594672.

[9] S. Liawatimena, “Fish Classification System Using YOLOv3-ResNet18 Model for Mobile Phones,” *CommIT J.*, vol. 17, no. 1, pp. 71–79, 2023, doi: 10.21512/COMMIT.V17I1.8107.

[10] K. A. Zakry, “Classification of dances using AlexNet, ResNet18 and SqueezeNet1_0,” *IAES Int. J. Artif. Intell.*, vol. 12, no. 2, pp. 602–609, 2023, doi: 10.11591/ijai.v12.i2.pp602-609.

[11] I. Tahyudin, “High-Accuracy Stroke Detection System Using a CBAM-ResNet18 Deep Learning Model on Brain CT Images,” *J. Appl. Data Sci.*, vol. 6, no. 1, pp. 788–799, 2025, doi: 10.47738/jads.v6i1.569.

[12] X. Ma, “A bearing fault diagnosis model with convolutional cross transformer and ResNet18,” *Meas. Sci. Technol.*, vol. 36, no. 1, 2025, doi: 10.1088/1361-6501/ad8a7b.

[13] S. She, “Evaluation of Defects Depth for Metal Sheets Using Four-Coil Excitation Array Eddy Current Sensor and Improved ResNet18 Network,” *IEEE Sens. J.*, vol. 24, no. 12, pp. 18955–18967, 2024, doi: 10.1109/JSEN.2024.3367816.

[14] S. T. Prasath, “Deep Learning Framework for Colorectal Cancer Classification Using ResNet18 Based on Dietary Habits Related to Meat Intake and Cooking Methods,” *IEEE Access*, vol. 12, pp. 99453–99468, 2024, doi: 10.1109/ACCESS.2024.3430036.

[15] Y. Jia, “Research on Improving ResNet18 for Classifying Complex Images Based on Attention Mechanism,” *Communications in Computer and Information Science*, vol. 2139. pp. 123–139, 2024, doi: 10.1007/978-981-97-3948-6_13.

[16] H. Liu, “ResNet18 Performance: Impact of Network Depth and Image Resolution on Image Classification,”

ICAAI 2024 - Conference Proceedings of the 2024 8th International Conference on Advances in Artificial Intelligence. pp. 351–356, 2025, doi: 10.1145/3704137.3704173.

[17] Y. Mei, "ResNet18 facial feature extraction algorithm improved based on hybrid domain attention mechanism," *PLoS One*, vol. 20, no. 3, 2025, doi: 10.1371/journal.pone.0319921.

[18] M. Alruwaili, "LSTM and ResNet18 for optimized ambulance routing and traffic signal control in emergency situations," *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-89651-4.

[19] S. Cheng, "Efficient acne classification based on Resnet18," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 13545. 2025, doi: 10.1117/12.3060304.

[20] S. R. Kempanna *et al.*, "Revolutionizing brain tumor diagnoses: a ResNet18 and focal loss approach to magnetic resonance imaging-based classification in neuro-oncology," *Int. J. Electr. Comput. Eng.*, vol. 14, no. 6, p. 6551, Dec. 2024, doi: 10.11591/ijece.v14i6.pp6551-6559.

[21] H. Jin, K. Han, H. Xia, B. Xu, and X. Jin, "Detection of weeds in vegetables using image classification neural networks and image processing," *Front. Phys.*, vol. 13, Jan. 2025, doi: 10.3389/fphy.2025.1496778.