

## LONG BONES X-RAY FRACTURE CLASSIFICATION USING MACHINE LEARNING

Soaad N. Kadry<sup>1,3\*</sup>, Hala Maghraby<sup>2</sup>, Sabry Abd El-Moetty<sup>3</sup>, Ashraf Al-Marakeby<sup>1</sup>

<sup>1</sup>Systems and Computers Engineering Department, Faculty of Engineering, Al-Azhar University, Cairo, Egypt

<sup>2</sup>Departments of Radiology, Faculty of Medicine, Al Azhar University, Cairo, Egypt

<sup>3</sup>Computer Engineering and Information Technology Department, Modern Academy For Engineering & Technology, Cairo, Egypt

\*Correspondence: [zo3adnasser@gmail.com](mailto:zo3adnasser@gmail.com)

### Citation:

S.N. Kadry, H. Maghraby, S. Abd El-Moetty and A. Al-Marakeby, "Long bone X-ray fracture classification using machine learning", Journal of Al-Azhar University Engineering Sector, vol. 19, pp. 121 - 133, 2024.

Received: 31 December 2023

Revised: 01 March 2024

Accepted: 30 March 2024

DoI:10.21608/auj.2024.259630.1577

Copyright © 2024 by the authors. This article is an open-access article distributed under the terms and conditions of Creative Commons Attribution-Share Alike 4.0 International Public License (CC BY-SA 4.0)

### ABSTRACT

Accurate long bone fracture diagnosis is essential to prevent permanent deformities resulting from misdiagnosis. This study uses machine learning to introduce a multi-class classification and detection system for long bone fractures. In this study, two image classifications are applied Binary classification and Multi-class classification, and an image detection model. Binary classification to distinguish normal and fractured bone X-ray images. Three models are used for this classification, Model A and Model B are used for grayscale images, and a ResNet50 pertained model for RGB images. Multi-class classification to identify fracture type using ResNet50 fine-tuned model And a Faster RCNN detection model to classify and detect the fracture type and its location in the X-ray images. The dataset was collected from various resources and labeled and annotated following Müller AO classification for bone fracture types. Binary classification achieved a 90.2% accuracy rate for Model A, 90.85% for Model B, and 96.5% for ResNet50, While the multi-class classification model achieved 87.7% accuracy in identifying fracture types for ResNet50 and 80% for Faster RCNN in fracture detection.

**KEYWORDS:** Radiology, Image Classification, Image Detection, CNN, ResNet50, Faster RCNN.

### تصنيف كسور العظام الطويلة في الأشعة السينية باستخدام التعلم الآلي

سعاد ناصر قذري<sup>1,3\*</sup>، هالة مغربي<sup>2</sup>، صبري عبد المعطي<sup>3</sup>، أشرف المراكبي<sup>1</sup>

<sup>1</sup>قسم هندسة النظم والحاسبات، كلية الهندسة، جامعة الأزهر، القاهرة، مصر،

<sup>2</sup>قسم الأشعة، كلية الطب، جامعة الأزهر، القاهرة، مصر،

<sup>3</sup>قسم هندسة الحاسبات وتكنولوجيا المعلومات، الأكاديمية الحديثة للهندسة والتكنولوجيا، القاهرة، مصر.

\*البريد الإلكتروني للباحث الرئيسي : [zo3adnasser@gmail.com](mailto:zo3adnasser@gmail.com)

### المخلص

يعد التشخيص الدقيق لكسور العظام الطويلة أمرًا مهمًا لمنع التشوهات الدائمة الناتجة عن التشخيص الخاطئ. تقدم هذه الدراسة نظام تصنيف وكشف متعدد الفئات لكسور العظام الطويلة باستخدام التعلم الآلي. تم في هذه الدراسة تطبيق تصنيفين للصور: التصنيف الثنائي والتصنيف متعدد الفئات، ونموذج كشف الصور. التصنيف الثنائي للتمييز بين صور الأشعة السينية للعظام السليمة والمكسورة. يتم استخدام ثلاثة نماذج لهذا التصنيف، النموذج A والنموذج B يستخدمان للصور ذات التدرج الرمادي،

ونموذج ResNet50 الخاص بـ صور RGB. تصنيف متعدد الفئات لتحديد نوع الكسر باستخدام نموذج ResNet50 الدقيق ونموذج كشف Faster RCNN لتصنيف وكشف نوع الكسر وموقعه في صور الأشعة السينية. تم جمع مجموعة البيانات من موارد مختلفة وتم تصنيفها وشرحها وفقاً لتصنيف Müller AO لأنواع كسور العظام. حقق التصنيف الثنائي نسبة دقة 90.2% للنموذج A، و 90.85% للنموذج B، و 96.5% لـ ResNet50، بينما حقق نموذج التصنيف متعدد الفئات دقة 87.7% في تحديد أنواع الكسور لـ ResNet50 و 80% لـ Faster RCNN في اكتشاف الكسور وتحديد مكانها.

**الكلمات المفتاحية:** الأشعة، تصنيف الصور، اكتشاف الصور، ResNet50، CNN، Faster RCNN.

## 1. INTRODUCTION

In today's healthcare system, medical imaging is essential for diagnosing and treating a wide range of illnesses. In particular, X-ray pictures offer insightful information about the interior workings of the human body. The primary purposes of X-rays are to examine the bones and joints and identify issues with interior organs, teeth, and bone fractures, among other things [1].

Patients may experience serious consequences from an incorrect diagnosis of the type of bone fracture, including possible long-term disabilities [2]. The body's bones are shaped and sized differently. Long, short, flat, and irregular bones are the four primary forms of bones. Bones whose length is greater than their width are called Long bones. These bones include the tibia/fibula, femur, humerus, radius/ulna, metatarsals, and phalanges, which have long shafts with two bulky ends. Long bone fractures were the most frequent kind of fracture in 2019, according to an analysis study conducted in 204 nations [3].

There are numerous ways to categorize the various types of shaft-long bone fractures. The Müller AO classification is one of them. it categorizes fractures based on the fracture line, the degree of fragmentation, and the anatomical location of the fracture [4]. long bone fractures were classified into three groups based on the Müller AO system: type A, B, and C according to the fracture's complexity. Type A fractures consist of two fracture pieces and a single fracture line. Type B fractures are wedge-shaped, meaning that multiple intermediate fracture blocks exist, but the main cortical bone can still make direct contact Type C is a complex fracture, containing multiple intermediate fracture blocks that after reduction there is no contact between the main fracture blocks.

With advancements in machine learning and artificial intelligence, there has been an increasing interest in automating the analysis of X-ray images. AI models [5] have demonstrated their efficacy in image classification [6] and object detection [7] in medical image diagnosing including diagnosis of bone fracture [8]. In the last few years, deep learning has been applied in the field of traumatology and orthopedics. Most of them focus on classifying bone as normal or fractured. For example, A deep Neural Network (DNN) model for detecting and classifying bone fractures was developed that overcomes the limitations of the small dataset by using data augmentation techniques on 100 X-ray images to reach 4000 images [9], It achieves 92.44% classification accuracy in differentiating between healthy and fractured bones. Using a deep learning algorithm, A model to detect rib fractures in chest CT images was developed, It was trained over 1,700 patients and used a three-stage algorithm, bone segmentation, rib location detection, and fracture classification [10]

Some used pre-processing, edge detection, and feature extraction for detecting lower leg bone fracture in X-ray images, the research showcases significant improvements in fracture detection accuracy using various machine learning algorithms [11]. Some papers also used a detection model to locate and classify the fracture. A study presented a Faster RCNN detection

model to detect femoral shaft fractures and classify the fracture type into 9 different types using 2333 annotated X-ray images [12].

Therefore, in this study, several machine learning models were applied to classify and detect long bone fractures in 3000 X-ray images that were manually labeled and annotated According to the AO classification and compared the results.

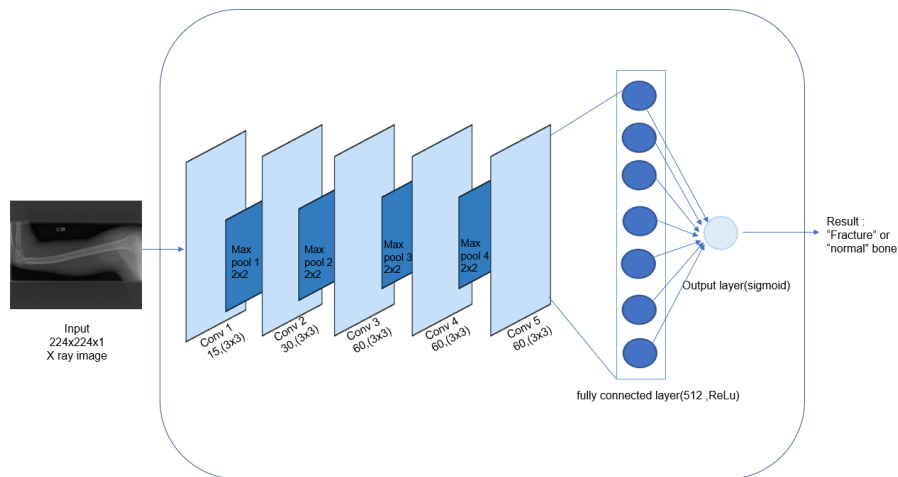
## 2. MACHINE LEARNING FOR CLASSIFICATION AND DETECTION

### 2.1. Binary classification

First, binary classification models were applied to X-ray images to classify the dataset into normal or fractured bone

#### 2.1.1. Model A

The model takes a 224x224 grayscale image as input and outputs whether the image contains normal or fractured bone. As shown in **Fig. 1**, the model consists of 5 convolution layers with activation='relu', 4 max-pooling layers size 2, and 2 fully connected layers with activation functions of 'relu' and 'sigmoid' Each convolutional layer has a certain number of filters, which are used to extract features from the input image. Each convolutional layer is followed by a 'relu' activation function, which helps the model learn non-linear relationships between the input data. Each max-pooling layer reduces the spatial dimensions of the feature maps by half. This helps to reduce the model's computational complexity and makes the model more robust to noise and variations in the input image.



**Fig. 1.** Pipeline of Model A

To reduce overfitting, several strategies were implemented, each targeting different aspects of the training process to ensure the model generalizes well to unseen data, A Dropout was added between convolution layers that randomly disables up to 20% and 30% of neurons and encourages the network to learn more generalized features. L2 Regularization was applied with a factor of 0.001 in the Dense layer that penalizes large weights to prevent overfitting, encouraging simpler models by making the weights smaller. Early Stopping was added, it halts training if validation loss doesn't improve for 5 epochs, restoring weights to the best state to prevent overfitting and ensure optimal generalization.

### 2.1.2. Model B

This model takes a 224x224 grayscale image as input and outputs whether the image contains normal or fractured bone, it consists of 11 convolution layers with activation=, 5 max-pooling layers (2x2), and 2 fully connected layers with activation functions of 'relu' and 'sigmoid'. In an attempt to increase the accuracy, the number of layers was increased compared to model A but because of the small size of the image, only one max pooling layer was added to the previous model structure as shown in Fig. 2.

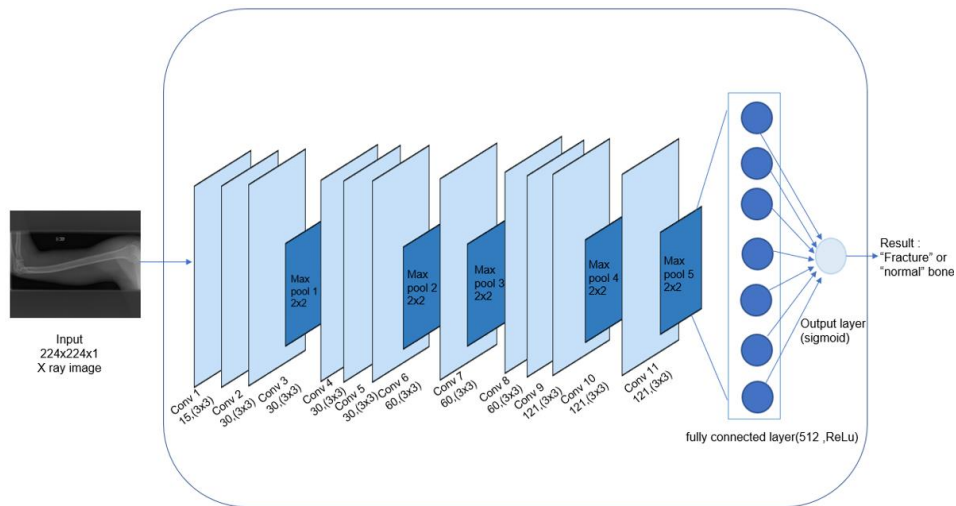


Fig. 2. Pipeline of Model B

### 2.1.3. ResNet50 Fine Tuned

To get more accurate results a transfer learning approach was used[13], this model architecture combines a powerful pre-trained ResNet50 short for Residual Networks (a classic neural network used as a backbone for many computer vision tasks ) [14] with additional layers for fine-tuning and classification, shown in Fig. 3. It will be used as a feature extractor to extract meaningful features from X-ray images. the top (classification) layers of the ResNet50 model were not included. These top layers are responsible for classifying objects into 1,000 categories, which is unnecessary for this binary classification task. the model is initialized with pre-trained weights from the ImageNet dataset. These weights provide a good starting point for feature extraction.

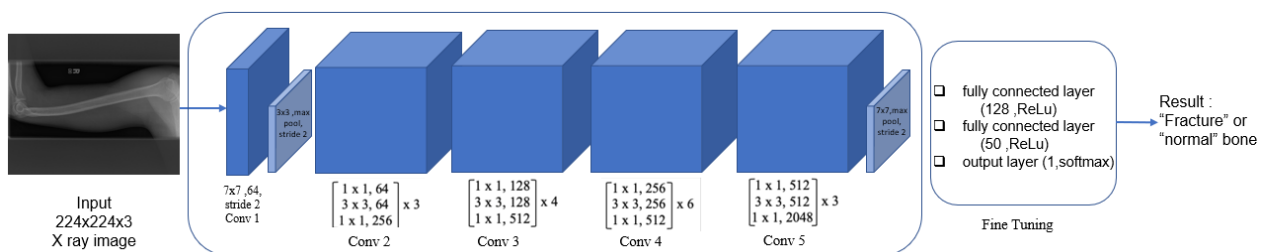


Fig. 3. Pipeline of the Fine-Tuned ResNet50 binary classification model

Global Average Pooling is applied to the output of the ResNet50 feature extractor also data augmentation layers are added to the model to artificially increase the diversity of the training dataset, and random transformations were applied to the input images. These transformations

include random horizontal and vertical flips, random rotations (up to 20% rotation), and random contrast adjustments

two dense (fully connected) layers are added on top of the output of the pre-trained ResNet50 feature extractor. These layers are responsible for learning the classification features from the extracted image features. The output layer consists of two units with softmax activation. Dropout, L2 Regularization, and Early Stopping were added to reduce overfitting. The softmax activation function computes the probability distribution over the two classes (fracture and normal), and the model predicts the class with the highest probability.

## 2.2. Multi-class classification

### 2.2.1. ResNet50 Fine Tuned

The model uses the ResNet50 pre-trained model as the base architecture. The pre-trained model's weights are frozen, and two additional fully connected layers with 'relu' activation are added for fine-tuning. The model's weights are loaded from ImageNet. Two fully connected layers with 128 and 50 units, respectively, are added to the output of the pre-trained model, followed by a Dense layer with a softmax activation function that contains 4 output neurons to output 4 classes to classify the image into normal, type A, type B or type C, as shown in Fig. 4.

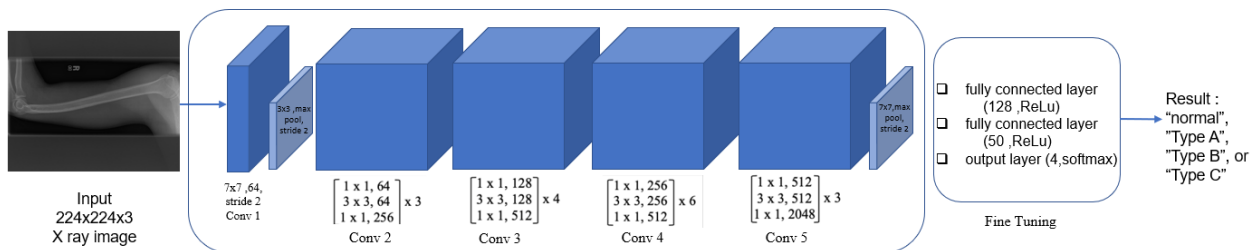


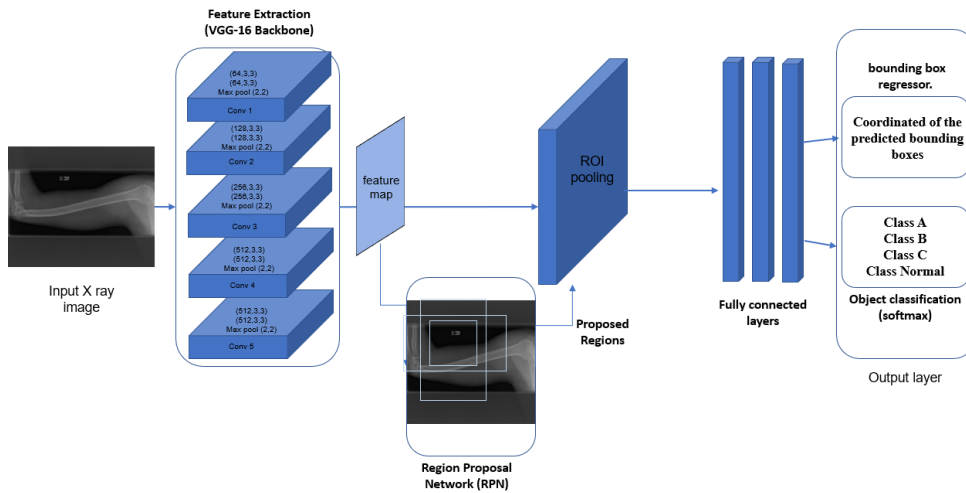
Fig. 4. Pipeline of the Fine-Tuned ResNet50 model multi-class classification model

## 2.3. X-ray bone Fracture Detection

### 2.3.1. Faster RCNN

Object detection is one of the fundamental tasks in computer vision. It refers to the recognition of both the positions and classes of objects in the image.[15] The research aims to detect the fracture's type and location on the X-ray image. For this purpose, Faster RCNN object detection model [16] has been trained to find the shaft fractures in the X-ray images of long bones and classify them according to their AO classification. This model uses RPN (region proposal network) to generate detection boxes directly. it is a convolutional network used to generate detection boxes for region proposals.

As shown in Fig. 5, The Faster RCNN model contains a convolution neural network backbone model (VGG-16) as a feature extractor, the VGG16 backbone is applied without a fully connected layer and generates a feature map. The RPN layer is responsible for proposing regions of interest (ROI). Finally, a classification regression network is used to classify predicted bounding boxes and produce the coordinated regression and classification values.



**Fig. 5.** Pipeline of the Faster RCNN model with VGG-16 as the backbone

### 3. DATASET

#### 3.1 Data collection

The dataset collected contains 3000 X-ray images of long bones, including normal and fractured ones. Fractures are composed of three groups: simple fracture (type A) 332 images, wedge-shaped fracture (type B) 123 images, and complex fracture (type C) 124 images.

Humerus and Forearm X-ray images were collected from the MURA dataset, which contains 727 radiographic studies of the humerus bone and 1010 studies on the Forearm bone(radius/ulna) that were labeled into Normal -Abnormal ie. (tumor –fracture –wires ..etc.) After filtering to normal and fractured it reached 2339 normal and 275 fractured then the rest of the long bone X-ray images were collected from various resources as shown in **Table 1**.

**Table 1.** Number of X-ray images used in the dataset with its sources

No of images	type	bone	source
2339	normal	Humerus radius/ulna	Mura dataset
82	normal	Femur tibia/fibula Humerus radius/ulna	<a href="http://bones.getthediagnosis.org/">http://bones.getthediagnosis.org/</a>
275	fracture	Humerus radius/ulna	Mura dataset
304	fracture	Femur tibia/fibula Humerus radius/ulna	<a href="https://www.google.com/imghp?hl=ar&amp;ogbl">https://www.google.com/imghp?hl=ar&amp;ogbl</a>

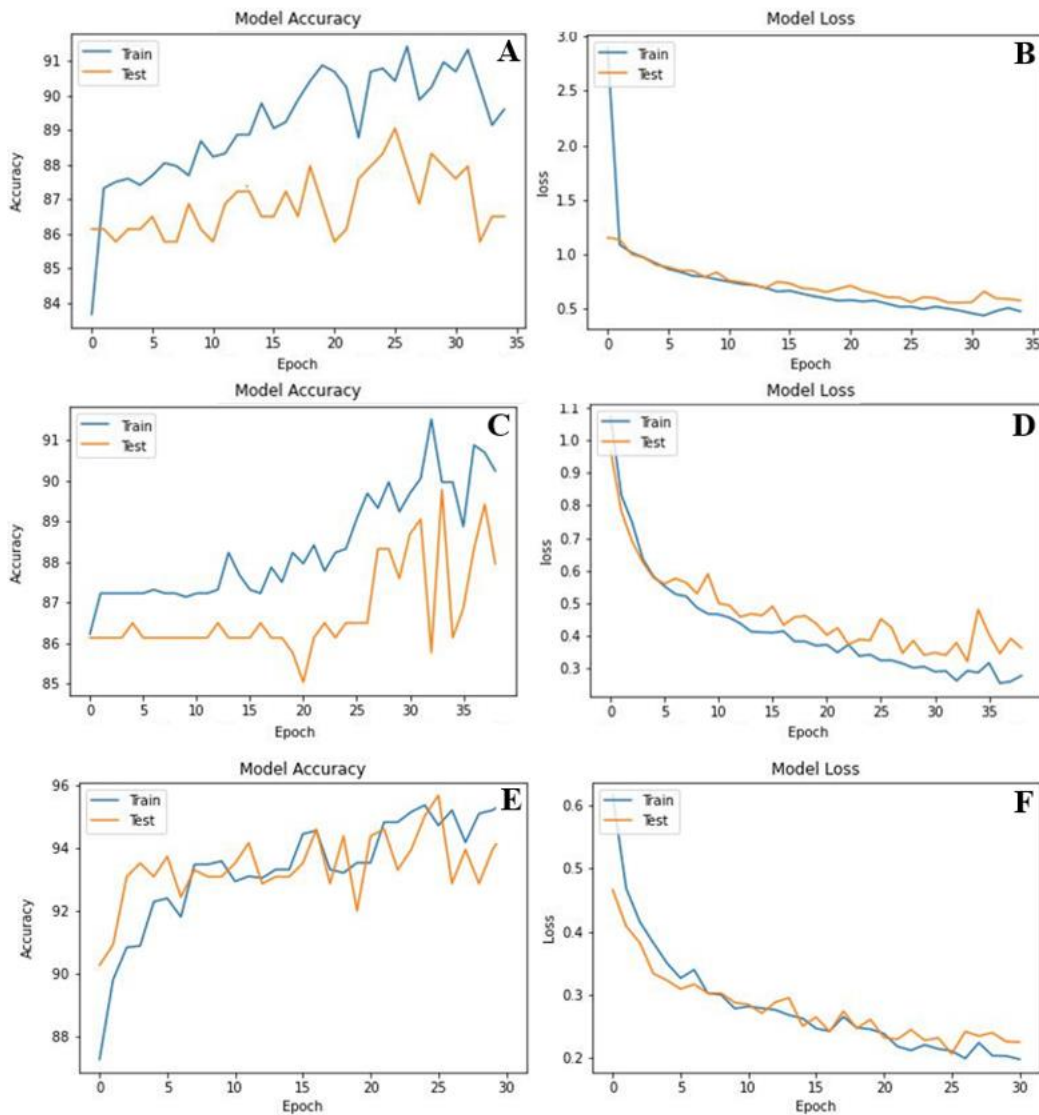
#### 3.2 Data preprocessing

In this research some rules were applied to the collected images for example people with intertrochanteric or neck fractures are not included. Blurry pictures and patients with tumors or other pathological fractures are not accepted. However, X-rays from the same patient taken at various times are acceptable. Multiple image enhancements were applied (such as adjusting contrast, sharpening, Histogram equalization, cropping, and a few more) as each image needed,

Then it was annotated by the Roboflow online tool [17] into 4 classes (A, B, C, and normal) and saved as a PASCAL-VOC dataset form that contained 3670 annotation files.

#### 4. RESULTS AND DISCUSSION

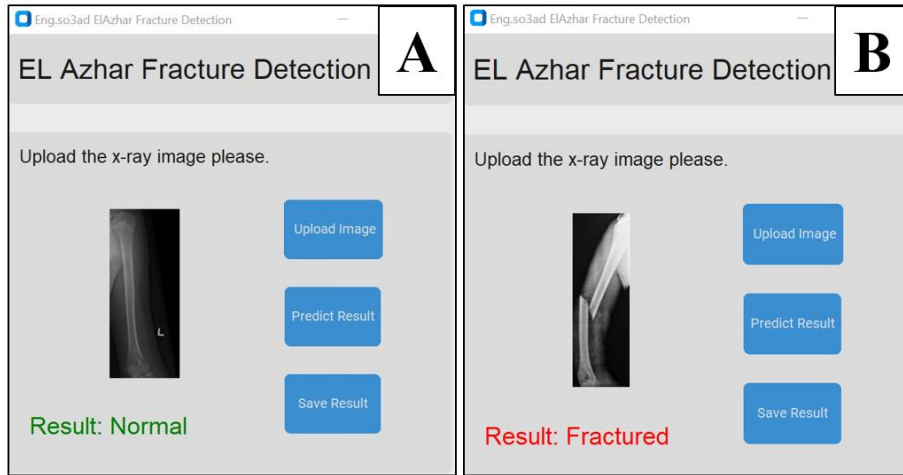
The dataset was split into 70 % train(2100 images), 20% validation (600 images), and 10% test (300 images), the models were implemented in Spider and Colab framework. **Fig. 6** shows the training and testing accuracy and losses over epochs for binary classification models applied, Model A achieved 90.2% test accuracy with losses of 0.26, Model B achieved 90.85% test accuracy with a loss of 0.25, And the ResNet50 Fine Tuned model achieved 96.5% test accuracy with losses of 0.16.



**Fig. 6.** Accuracy and losses Vs. epochs for model A (A,B), model B (C,D), and ResNet50 model (E,F)

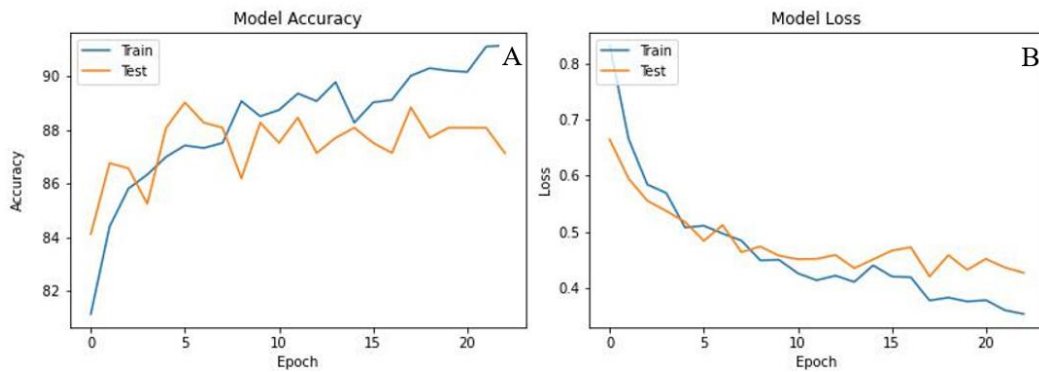
As the ResNet50 model gave the best accuracy and lowest losses among the binary classification models applied, A user interface was developed to help doctors use this model with ease as shown in **Fig. 7**.





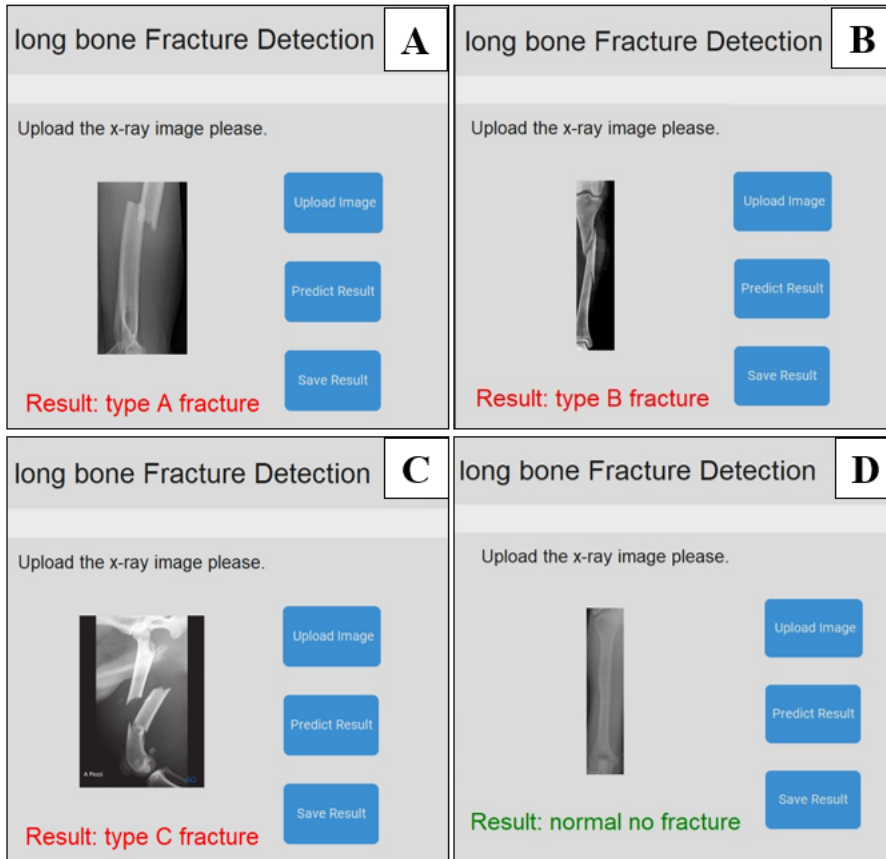
**Fig. 7.** Testing ResNet50 model on normal (A) and fractured (B) X-ray images

Second is the multi-class classification and detection models. ResNet50 achieved an accuracy of 87.7 % and losses of 0.5 in classifying the type of fracture in the X-ray images as shown in **Fig. 8**. Test samples for the ResNet50 model for multi-class classification are presented in **Fig. 9**.



**Fig. 8.** Accuracy Vs. epochs (A) and losses Vs. epochs (B) for multiclass ResNet50 model

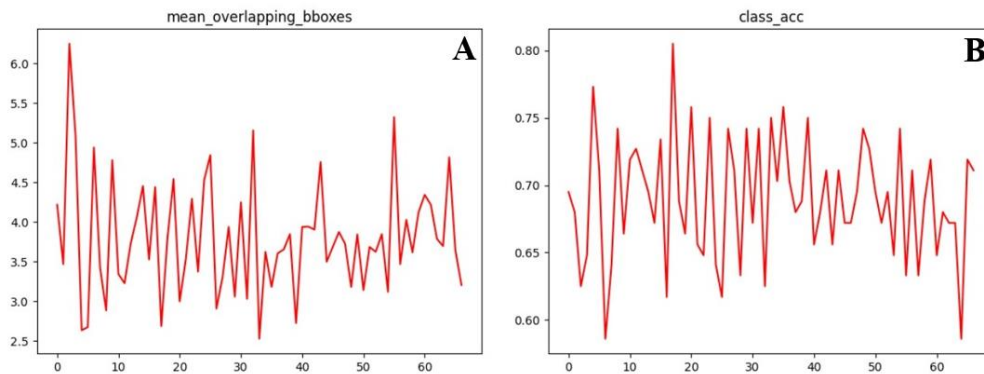




**Fig. 9.** Testing multi-class ResNet50 model on Type A (A), Type B (B), Type C (C), and normal (D) X-ray images

A Faster R-CNN detection model with VGG -16, has been applied using the Google Colab environment, it takes the ground truth bounding boxes -that were manually annotated in the training dataset as input. It contains 3670 annotation files with four classes: (A, B, C, and normal) with their actual bounding box coordinates and output the predicted class along with its bounding box coordinates.

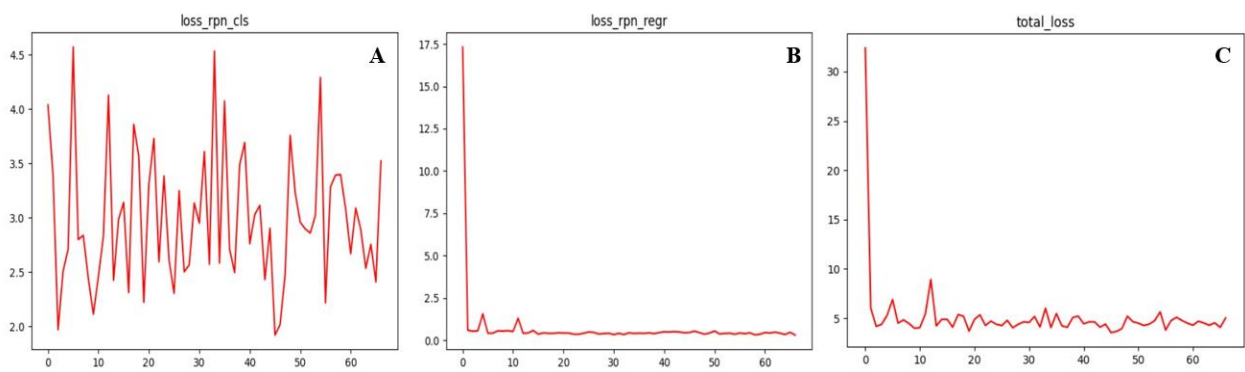
Initially, the Region Proposal Network (RPN) proposes preliminary bounding boxes that roughly outline potential fracture locations. the scales for anchor boxes used are [64, 128, 256] chosen relative to the input image size.



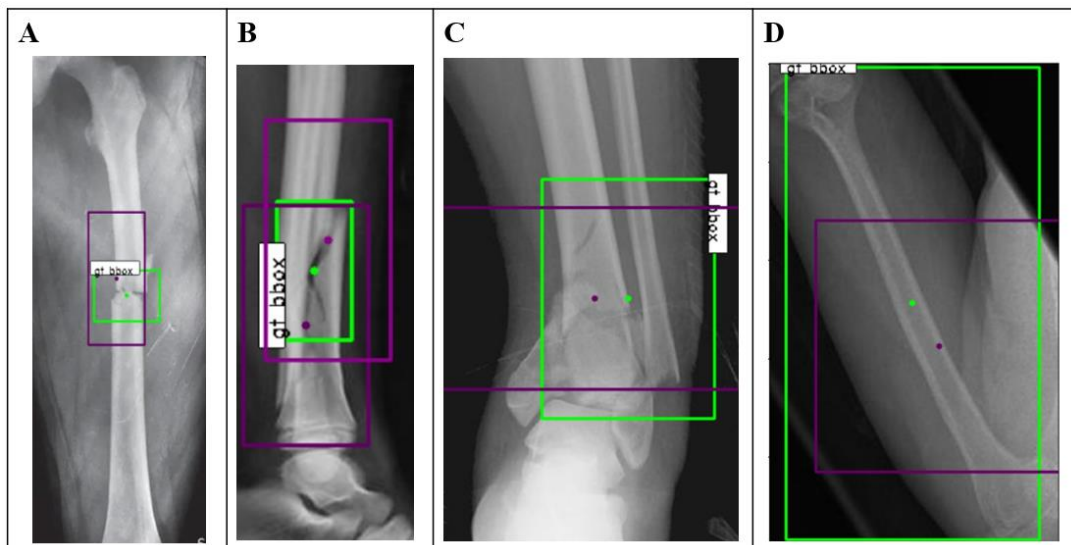
**Fig. 10.** Mean overlapping bounding bboxes Vs. epochs (A) and class accuracy Vs. epochs (B)

The intersection Over Union (IOU) metric calculates the overlap as a ratio of the area of intersection to the area of union between the predicted bounding box and the ground truth bounding box. An IOU threshold is often set at 0.5, meaning any predicted box with an IOU exceeding this threshold is considered accurate. **Fig. 10** shows the average IOU scores across all objects and images in a given epoch. it also shows the classification accuracy over epochs that achieved an average of 75%.

Mean average Precision (mAP) is the accurate metric for object detection, so we mainly used it to evaluate the model performance. It takes recall and precision into account. Precision illustrates the ratio of correctly predicted objects to all predicted objects, while recall quantifies the percentage of truly relevant objects that were found. The mAP shows the average of all AP scores for each class. The model's mAP was 79.8% on the validation set.



**Fig. 11.** RPN classification losses (A), RPN regression losses (B), and total losses(C) Vs. epochs



**Fig. 12.** Testing Faster RCNN model on Type A (A), Type B (B), Type C (C), and normal (D) X-ray images

**Fig. 11** shows the losses of RPN classification over epochs reached an average of 3.22, also the losses of RPN regression over epochs that concert the coordinates of the classes reached an average of 0.41, and the total losses over epochs were 4.65. **Fig. 12** shows samples of the predicted

fracture types, the green box presents the ground truth bounding box and the other is the predicted bounding box.

**Table 2.** The final results of each model applied with accuracy and losses

ML model	Task	Accuracy	Losses
Model A	Binary Classification	90.2 %	0.26
Model B	Binary Classification	90.85 %	0.25
ResNet50	Binary Classification	96.5 %	0.16
ResNet50	Multi-class Classification	87.7 %	0.5
Faster RCNN	Classification+ Detection + Localization	Map =0.79.8 accuracy of bbox from RPN:0.72	Loss RPN classifier: 3.22 Loss RPN regression: 0.41 Loss Detector classifier: 0.76 Loss Detector regression: 0.24 Total loss: 4.65

**Table 3.** Comparison with previous related papers

data type	no of images	Task	model	accuracy	ref
Thigh x-ray	3842	fracture detection	DCFPN	mAP=0.821	[19]
bone x-ray	100	binary classification (normal, fractured)	CNN	92.44%	[9]
femur x-ray	2453	multiclass classification (normal, A, B fracture)	InceptionV3	87%	[27]
			VGG16	82%	
			ResNet50	85%	
femoral x-ray	2333	fracture detection and classification	Faster RCNN with ResNet-50	71.50%	[12]
Arm X-ray	40000	fracture detection	R-CNN	mAP=0.62	[21]
spine CT	5134	fracture detection	Faster R-CNN	mAP=0.733	[25]
spine CT	5134	fracture detection	YOLOv2	mAP=0.753	[26]
chest CT	5000	fracture segmentation, detection, and classification.	U-Net + Dense Net	89%	[10]
Thigh X-ray	3842	fracture detection	R-CNN	mAP=0.878	[20]
bone X-ray	9040	fracture detection	FAMO	mAP=0.774	[22]
X-ray	1052	fracture detection	Faster R-CNN	mAP=0.884	[23]
hand trauma x-ray	3067	fracture detection	Faster R-CNN	mAP=0.7	[24]
lower leg bone X-ray	270	binary classification (normal, fractured)	Naïve Bayes	64%	[11]
			Decision Tree	80%	
			Nearest Neighbors	83%	
			Random Forest	85%	
			SVM	92%	
wrist x-ray	20327	fracture detection	YOLOv8	mAP=0.638	[18]

**Table 2** presents the final accuracy for each model used in this study, it shows that the best binary classification model was the fine-tuned ResNet50 model with 96.5% accuracy also compared to the recent related research presented in **Table 3** and when applied as a multi-class classification model it also did well.

## CONCLUSION

In conclusion, diagnosis of bone fracture is critical, so we presented a few models to classify and detect fractures on long bone X-ray images in this study. The dataset was collected from various resources and labeled following Müller AO classification for bone fracture types. A binary classification was applied to distinguish normal and fractured bone in X-ray images. Three models are used for this classification. Model A achieved 90.2% accuracy, Model B achieved 90.85% accuracy for grayscale images, and a ResNet50 FineTuned model for RGB images achieved 96.5% accuracy. Since the ResNet50 FineTuned model was trained on plenty of data, the initial weights of the model used aided in the learning process, leading to the highest accuracy among the binary classification models applied.

The Multi-class classification model was applied to identify fracture types using ResNet50 FineTuned model, it achieved 87.7% accuracy in detecting 4 classes. The reason for the accuracy to decrease from the binary classification may be due to the decrease of the number of images in each class after splitting the 3000 images into 4 classes. As for the Faster RCNN image detection model with VGG-16 as a backbone that classifies and detects the fracture type with its bounding box location. It achieved nearly 80% mean average precision. The future work is to apply the Transformer model on the same dataset and compare the results.

## REFERENCES

- [1] Ou, Xiangyu, et al., "Recent Development in X-Ray Imaging Technology: Future and Challenges," *Research* (Washington, D.C.), vol. 2021 9892152, Dec. 2021.
- [2] Pinto A., Berritto D., Russo A., Riccitiello F., Caruso M., Belfiore MP., Papapietro VR., Carotti M., Pinto F., Giovagnoni A., Romano L., Grassi R., "Traumatic fractures in adults: missed diagnosis on plain radiographs in the Emergency Department," *Acta Biomed*, vol. 89(1-S), pp. 111-123, 2018.
- [3] GBD 2019 Fracture Collaborators, "Global, regional, and national burden of bone fractures in 204 countries and territories, 1990-2019: a systematic analysis from the Global Burden of Disease Study 2019," *Lancet Healthy Longev*, vol. 2(9), pp. e580-e592, 2021.
- [4] Meinberg EG., Agel J., Roberts CS., Karam MD., Kellam JF., "Fracture and Dislocation Classification Compendium-2018," *J Orthop Trauma*, vol. 32 Suppl 1, S1-S170, 2018.
- [5] Panayides, A. S., et al., "AI in Medical Imaging Informatics: Current Challenges and Future Directions," *IEEE Journal of Biomedical and Health Informatics*, vol. 24(7), 1837-1857, 2020.
- [6] Kim, H.E., et al., "Transfer learning for medical image classification: a literature review," *BMC Medical Imaging*, vol. 22(69), 2022.
- [7] Wu, X., Sahoo, D., & Hoi, S.C.H., "Recent Advances in Deep Learning for Object Detection," *Neurocomputing*, 2020.
- [8] Kalmet, P. H. S., Sanduleanu, S., Primakov, S., Wu, G., Jochems, A., Refaee, T., Ibrahim, A., Hulst, L. V., Lambin, P., & Poeze, M. "Deep learning in fracture detection: a narrative review," *Acta orthopaedica*, vol. 91(2), pp. 215–220, 2020.
- [9] Yadav, D. P., & Rathor, S. "Bone Fracture Detection and Classification using Deep Learning Approach," 2020 International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC), GLA University Mathura UP India, February 28-29, 2020.
- [10] Yao, L., Guan, X., Song, X., Tan, Y., Wang, C., Jin, C., Chen, M., Wang, H., & Zhang, M., "Rib fracture detection system based on deep learning," *Scientific Reports*, vol. 11(23513), 2021.

- [11] Ahmed, K.D., & Hawezi, R. "Detection of bone fracture based on machine learning techniques," *Measurement: Sensors*, vol. 27, 100723, 2023.
- [12] Qi, Y., Zhao, J., Shi, Y., Zuo, G., Zhang, H., Long, Y., Wang, F., & Wang, W., "Ground Truth Annotated Femoral X-Ray Image Dataset and Object Detection Based Method for Fracture Types Classification," *IEEE Access*, vol. 8, 189436-189444, 2020.
- [13] Kim, H.E., Cosa-Linan, A., Santhanam, N., et al., "Transfer learning for medical image classification: a literature review," *BMC Med Imaging*, vol. 22, 69, 2022.
- [14] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 770-778, 2016.
- [15] Zhao, Zhong-Qiu, et al. "Object detection with deep learning: A review," *IEEE transactions on neural networks and learning systems*, vol. 30(11), 3212-3232, 2019.
- [16] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 39, no. 06, pp. 1137-1149, 2017.
- [17] Dwyer, B., Nelson, J., Solawetz, J., et. al. Roboflow (Version 1.0) [Software], Available from <https://roboflow.com>, 2022.
- [18] Ju, RY., Cai, W., "Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm," *Sci Rep*, vol. 13, 20077, 2023.
- [19] Guan, B., Yao, J., Zhang, G. & Wang, X., "Thigh fracture detection using deep learning method based on new dilated convolutional feature pyramid network," *Pattern Recogn. Lett.*, vol. 125, 521–526, 2019.
- [20] Wang, M. et al., "Parallelnet: Multiple backbone network for detection tasks on thigh bone fracture," *Multimed. Syst.*, vol. 27, pp. 1091–1100, 2021.
- [21] Guan, B., Zhang, G., Yao, J., Wang, X. & Wang, M. Arm fracture detection in x-rays based on improved deep convolutional neural network. *Comput. Electric. Eng.* 81, 106530 (2020).
- [22] Wu, H.-Z. et al, "The feature ambiguity mitigate operator model helps improve bone fracture detection on x-ray radiograph," *Sci. Rep.*, vol. 11, pp. 1–10, 2021.
- [23] Ma, Y. & Luo, Y., "Bone fracture detection through the two-stage system of crack-sensitive convolutional neural network," *Inf. Med. Unlocked*, vol. 22, 100452, 2021.
- [24] Xue, L. et al., "Detection and localization of hand fractures based on ga\_faster r-cnn," *Alex. Eng. J.*, vol. 60, pp. 4555–4562, 2021.
- [25] Sha, G., Wu, J. & Yu, B., "Detection of spinal fracture lesions based on improved yolov2," *Proc. 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, pp. 235–238, IEEE, 2020.
- [26] Sha, G., Wu, J. & Yu, B., "Detection of spinal fracture lesions based on improved faster-rcnn," *2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIS)*, pp. 29–32, IEEE, 2020.
- [27] Tanzi, L., Vezzetti, E., Moreno, R., Aprato, A., Audisio, A., Massè, A., "Hierarchical fracture classification of proximal femur X-Ray images using a multistage Deep Learning approach," *European Journal of Radiology*, vol. 133, 109373, 2020.