HD-NET: Humerus deep-net for humerus fracture and bony callus formation analysis

Abdullah Tariq[1], Muhammad Shoaib[1], Dr.Shazia Arshad[1], Dr.Syed Khaldoon Khurshid[1], Dr.Faiza Iqbal[1], and Abqa Javaid[1]

[1]University of Engineering and Technology

March 8, 2023

Abstract

When employing x-ray images, fracture identification in orthopaedics is a difficult task. A large percentage of humerus fracture patients are seen in hospitals, particularly in their emergency departments. Similar to this, after a fracture, accurate callus production monitoring is crucial for bone healing. Thus, a fractured patient's diagnosis and therapy must be accurate and administered promptly. This work investigates the use of deep learning on X-ray images of the humerus for fracture snd bone callus formation analysis to help physicians in the diagnosis of such fractures, especially in emergency settings.This study is named HD-NET, which stands for Humerus Deep Net. The framework includes image enhancement using a Gaussian filter and histogram equalization, two-stage object detection, image super-resolution, U-NET segmentation with feature recalibration. Finally, an LSTM with a sequence length of 2 is used to analyze callus formation at the fracture site. The LSTM takes the segmented area as input and outputs a prediction for the stage of healing and potential complications. The proposed framework was evaluated on a combination of the MURA dataset and a self-collected dataset.Results demonstrated that in terms of specificity, sensitivity, and accuracy, the suggested framework performed better than earlier studies. This research can be expanded to different bone types and is useful for orthopaedic practitioners.

Received: Added at production | Revised: Added at production | Accepted: Added at production
DOI: xxx/xxxx

ARTICLE TYPE

HD-NET: Humerus deep-net for humerus fracture and bony callus formation analysis

Abdullah Tariq* | Prof,Dr.Muhammad Shoaib | Dr.Shazia Arshad | Dr.Syed Khaldoon Khurshid | Dr.Faiza Iqbal | Abqa Javaid

[1]Department of Computer Science, University of Engineering and Technology, Lahore, Pakistan

Correspondence
*Abdullah Tariq, Computer Science department, University of engineering and technology, Lahore, Pakisan Email: ab.sheikh909@gmail.com

Abstract

When employing x-ray images, fracture identification in orthopaedics is a difficult task. A large percentage of humerus fracture patients are seen in hospitals, particularly in their emergency departments. Similar to this, after a fracture, accurate callus production monitoring is crucial for bone healing. Thus, a fractured patient's diagnosis and therapy must be accurate and administered promptly. This work investigates the use of deep learning on X-ray images of the humerus for fracture snd bone callus formation analysis to help physicians in the diagnosis of such fractures, especially in emergency settings.This study is named HD-NET, which stands for Humerus Deep Net. The framework includes image enhancement using a Gaussian filter and histogram equalization, two-stage object detection, image super-resolution, U-NET segmentation with feature recalibration. Finally, an LSTM with a sequence length of 2 is used to analyze callus formation at the fracture site. The LSTM takes the segmented area as input and outputs a prediction for the stage of healing and potential complications. The proposed framework was evaluated on a combination of the MURA dataset and a self-collected dataset.Results demonstrated that in terms of specificity, sensitivity, and accuracy, the suggested framework performed better than earlier studies. This research can be expanded to different bone types and is useful for orthopaedic practitioners.

KEYWORDS:
Humerus Fracture; Deep learning; U-NET; Image Super Resolution; Fracture detection; Callus Formation

1 | INTRODUCTION

In the upper arm, between the elbow and the shoulder, is where you'll find the humerus bone. It upholds arm and shoulder movement as it is the longest bone in the arm. Figure (1) shows the humerus bone. According to World Health Organization research, humeral shaft fractures account for about 3% of all orthopedic injuries and cause major monetary and misfortune for society. According to a 2002 World Health Organization research, 4.8% of non-fatal road traffic injuries were humerus fractures. Locally, they mostly affect young, socio economically active age groups, with the majority of cases being brought on by motor accidents. It demonstrates to us how prevalent bone fractures are as a medical problem, with a non-negligible cost of care that can rise exponentially as life moves more quickly than we can fathom.The healing process is complicated after a bone fracture, which is a common injury [1]. One of the few tissues that can recover without leaving a fibrous scar is bone. It should be emphasised that excessive stress or movement might cause non-union or delayed healing [2] which affects 5–10% of all fractures [3].It is a vital practice to continuously

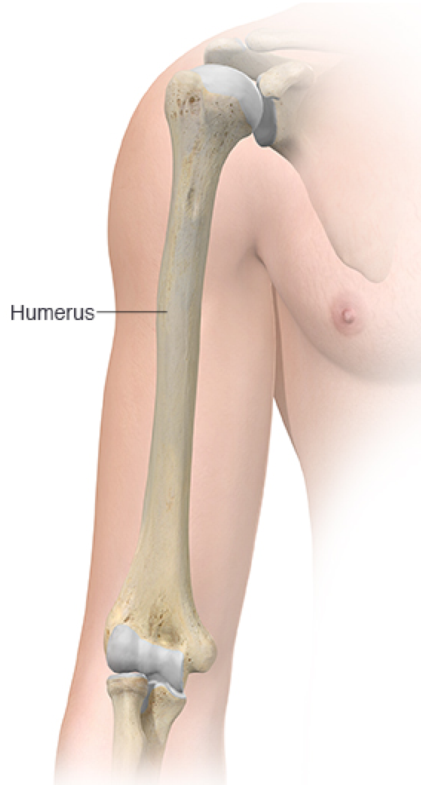

Figure 1 Humerus bone.

monitor bony callus at the fracture site in order to know that weather the bone is healing rightly.It takes time and persistence to heal from a bone fracture. The formation of a bone callus is an essential step in the healing process. It is a vital practice to continuously monitor bony callus at the fracture site in order to know that weather the bone is healing rightly.

To help orthopedists with routine fracture diagnostic procedures, X-ray images are frequently required. Bone radiological imaging has become highly common as people pay greater attention to their health. It is not an efficient method to manually examine these images because it takes a lot of time and labor. Particularly, in hectic clinical settings when experienced radiologists or other practitioners may not be present, clinicians lack the subspecialized knowledge and experience required to correctly identify fractures on radiographs. To relieve radiologists of their demanding task, the X-ray picture interpretation process must be accelerated.

The discovery of medications [4], disease analysis and diagnosis [5], image processing [6], patient monitoring [7], and surgery [8] are only a few of the contemporary applications of AI in the healthcare sector. AI is currently being applied in radiology for a wide range of activities, including automated disease identification, classification, segmentation, quantification, and many more projects. A particular kind of AI called Deep Learning (DL) has been shown to be capable of diagnosing diseases from medical images more precisely than doctors [9]. As the enormous volume of data is burdensome for doctors or other medical professionals, we want cutting-edge and compatible DL algorithms to exploit bone imaging-specific data. The growing body of research in this area demonstrates how eager people are to create AI radiology systems. In recent years, there has been a growing interest in using deep learning techniques for the detection and classification of humerus fractures [10].One of the key advantages of deep learning techniques is their ability to learn from large datasets. In the case of humerus fractures, this means that deep learning models can be trained on a large number of images of the upper arm, with some images showing fractures and others showing normal bones. By learning from these images, deep learning models can learn to recognize patterns that are indicative of a fracture, such as changes in bone density or shape.

For the detection of humerus fracture and analysis of bony callus, this study developed a multistage framework HD-NET, comprises on image super resolution, image segmentation with feature re-calibration and recurrent neural network.

The main contributions of the proposed work can be summarized in the following points:

- Proposed a multistage framework for the detection of humerus fracture and bony callus.
- Utilization of two stage object detectors and de-convolution for image super resolution.
- Utilization of feature re-calibration in the decoders of U-NET for better segmentation of fracture and callus area.

- An LSTM module with just two sequence length for the analysis of callus formation.

The rest of the paper is structured as follows: Section 2 provides the related work and section 3 presents the methodology of the paper. Experiments and results are explained in section 4. At the end, section 5 gives the conclusion and future directions of this paper.

2 | RELATED WORK

There has been a growing body of research on the use of deep learning for bone fracture detection in recent years. Some studies have focused on the use of deep learning for the detection and diagnosis of specific types of fractures, such as humerus fractures, wrist fractures, and rib fractures. Other studies have focused on the development of general-purpose deep learning algorithms for the detection and diagnosis of fractures in various bones. One of the main approaches used in the literature for bone fracture detection through deep learning is the use of convolutional neural networks (CNNs). CNNsare a type of neural network that is specifically designed for image classification and object recognition tasks. They are composed of multiple convolutional layers, which perform feature extraction and detection, followed by one or more fully connected layers, which perform classification. CNNs have been widely used for bone fracture detection due to their ability to learn and extract features from medical images in an automatic and hierarchical manner.

Using x-ray images [11] describe a technique for segmenting cervical vertebral bone. It first uses CNN to localise the spinal region and then localises the vertebral centres. Shape-aware deep segmentation is used to segment data. The process is entirely automatic. The dice similarity Co-efficient was examined using 172 photos, and the results showed that it is 0.84. Using mouse x-ray pictures, Okashi O. A. et al [12]. have provided a method to estimate spine curvature. The approach is composed of two stages: the first stage involves the use of experts' opinions to prepare the area of interest, and the second stage entails the separation of the ribs from the spine using Otsu's segmentation and morphological operations. According to the approach, classification accuracy is 98.6%

Guan et al. [13] achieved an average precision score of 82.1% for fracture detection in 3842 thigh fracture X-ray images using a dilated convolutional feature pyramid network (DCFPN). They also obtained a fracture detection score of 62.04% AP for approximately 4000 arm fracture X-ray images in the musculoskeletal radiograph (MURA) dataset using a two-stage region-based convolutional neural network (R-CNN) method [14]. Wang et al. [15] achieved an AP50 score of 87.8% for fracture detection in a dataset of 3842 thigh fracture X-ray images using the ParallelNet method with a TripleNet backbone network. Ma and Luo [16] achieved an accuracy of 90.11% for fracture detection and classification using a Faster R-CNN and CrackNet model on a dataset of 1052 bone images. Wu et al. [17] achieved an AP score of 77.4% for fracture detection using the Feature Ambiguity Mitigate Operator (FAMO) model with ResNet101 and feature pyramid network (FPN) on a dataset of 9040 radiographs of different body parts. Automatic fracture detection in radiographs using convolutional neural networks [18] presents a CNN for the automatic detection of fractures in radiographs. The authors used a dataset of 10,000 radiographs and achieved a sensitivity of 91.6% and a specificity of 90.6%. One limitation of this study is that it did not include a comparison with traditional machine learning methods, so it is difficult to assess the relative performance of the CNN. Qi et al [19]. achieved a mean AP (mAP) score of 68.8% using the anchor-based Faster R-CNN model with multi-resolution FPN and ResNet50 backbone network on 2333 femoral fracture X-ray images. Thian et al [18] used the Inception-ResNet Faster R-CNN model for fracture detection on 7356 wrist radiographic images. Sha et al [20]. achieved a mAP score of 75.3% using the You Only Look Once (YOLOv2)-based model for fracture detection in a dataset of 5134 spinal fracture CT images. Using another model based on Faster R-CNN, the mAP achieved was 73.3% for the same dataset. Jin et al [21]. achieved a sensitivity of 92.9% for segmentation and detection of rib fractures in 7473 CT images collected from 900 patients using the proposed U-Net-based FracNet. Xue et al. achieved an AP score of 70.7% for fracture detection using the guided anchoring method Faster R-CNN model in a dataset of 3067 hand fracture X-ray images.

Apart from fracture detection, deep-learning-based classification studies have also been conducted in the literature to identify the class of fracture. Uysal et al [22]. performed 26 different deep-learning-based classification procedures to identify the class of fracture in shoulder bone X-ray images in the MURA dataset and developed two different ensemble learning models to further improve the results. Raghavendra et al [23]. achieved a classification accuracy of 99.1% with the proposed CNN model on a thoracolumbar CT dataset containing 420 normal images and 700 fracture images. A deep learning approach for bone fracture detection in radiographs by [24] presents a CNN for the detection of fractures in radiographs. The authors used a dataset of 10,000 radiographs and achieved a sensitivity of 85.7% and a specificity of 87.8%. One limitation of this study is that it did not assess the performance of the CNN on a separate test dataset, so it is unclear how well the model would generalize to new radiographs. Automatic fracture detection in X-ray images using deep learning by [25] presents a deep learning approach for the detection of distal fractures in X-ray images. The authors used a dataset of 12,000 X-ray images and achieved a sensitivity of 86.2% and a specificity of 87.5%. One limitation of this study is that it did not compare the performance of the deep learning approach with traditional machine learning methods

Beyaz et al [26]. achieved an accuracy of 79.3% for fracture classification on 1341 femoral neck X-ray images using a proposed CNN model and genetic algorithm (GA). Tobler et al [27]. achieved the highest accuracy of 94% for fracture classification in a dataset of 15,775 frontal and lateral radiographs using the ResNet18 model. Kim et al [28]. achieved an area under the receiver operator characteristic curve (AUC) score of 0.954

Table 1 Overview and comparison of published studies.

Study	Focus of Study	Year of publish	Methodology	Dataset	Limitation
31	Automatic Detection of Humerus Deformation in X-ray Images	2022	Machine Learning and geometrical feature approaches	MURA	They just find and give their prediction on the basis of distance from upper and lower joint
32	Deep Scale-Variant Network for Femur Trochanteric Fracture Classification with HP Loss	2022	Deep scale-variant (DSV) network with a hybrid and progressive (HP) loss function to aggregate more influential representations of the fracture regions.	Custom Dataset	Not scalable as very deeper network is used. Less accuracy due to low resolution images.
33	Comparison and verification of two deep learning models for the detection of chest CT rib fractures	2022	VGG-19 and Resnet-50 as a backbone network	Custom Dataset	Rib Fracture detection can be improved by better acquisition of training data
34	A combined feature set for automatic diaphyseal Tibial fracture classification from X-Ray images	2021	Decision Tree and KNN is used with selection of top 5 features	Custom Dataset	Limited to frontal x-ray images.
35	Ultrasound imaging in monitoring bone healing after fracture surgery	2023	ImageNet with VGG11	Custom Dataset	Micro average is less due to less number of dataset and less enhancement of image.
36	Deep learning system for the location and classification of rib fractures	2022	Two classification algorithm with resnet 34.	Custom Dataset	Limited to chest x-rays in RGB

for classification of 1389 wrist radiographs using the InceptionV3 model. Chen et al [29]. achieved an accuracy of 73.59% for vertebral fracture classification in a dataset of 1306 plain frontal radiographs using the ResNeXt model. Tanzi et al [30]. achieved the highest accuracies of 87.

Above studies shows the use of deep learning achieve good results for the fracture detection and classification but due to low resolution images and noise in the X-ray images it is difficult to detect and segment the fracture area. In addition to this, callus formation analysis is still and open area for achieving better results. So, this study presents a framework which enhance the fracture area by passing the region of interest to image super resolution block, followed by U-NET with feature re-calibration module to more accurately segment the fracture and callus area.

3 | METHODOLOGY

For the detection of humerus fracture and bony callus from the x-ray images, following proposed methodology has been followed.Figure 2 shows the overall framework for the proposed work.

The proposed methodology for analyzing x-ray images of fractures involves several steps that use mathematical equations to process the images. The first step is to take an x-ray image as the input, denoted by I. The next step is to preprocess the image by removing noise and enhancing the colors. This can be done by applying a filter such as a Gaussian filter, denoted by G(I), which is defined in 1.

$$G\left(I\right)=\left(1/2\pi\sigma^{2}\right)*e^{\left(-\left(x^{2}+y^{2}\right)\right)/2*\sigma^{2}\right)} \tag{1}$$

where x and y are the coordinates of a pixel in the image and σ is a parameter that controls the standard deviation of the filter. Additionally, histogram equalization can be applied to enhance the colors, which is denoted by H(I) and is defined in 2:

$$H\left(I\right)=T\left(I\right)*L-1 \tag{2}$$

where T(I) is the cumulative histogram of the image and L is the number of levels in the image. Figure 3 presents the graphical illustration for the pre-processing step.

Once the image has been preprocessed, it is passed through single and multi-stage object detectors from the RCNN family, specifically to locate the object of interest, in this case, the fracture. The structures used in this study for object detection are shown in Figure 4 which are single-stage detectors and two-stage detectors. Single-stage detectors process input images through a backbone and neck, followed by a DenseHead operation to output both object.location and class. In contrast, two-stage detectors also include a RoIHead in addition to the backbone, neck, and DenseHead.

Single-stage detectors tend to have faster detection times, but two-stage detectors generally have higher accuracy, although this can vary depending on the dataset. Both types of detectors use FPN in the neck and ResNet50 in the backbone. These object detectors may include algorithms such as R-CNN, Fast R-CNN, and Faster R-CNN. The output of this step is the location of the fracture in the image. Figure 4 presents the graphical illustration for detection stage of fracture. The next step is to enhance the quality of the detected fracture using image super resolution

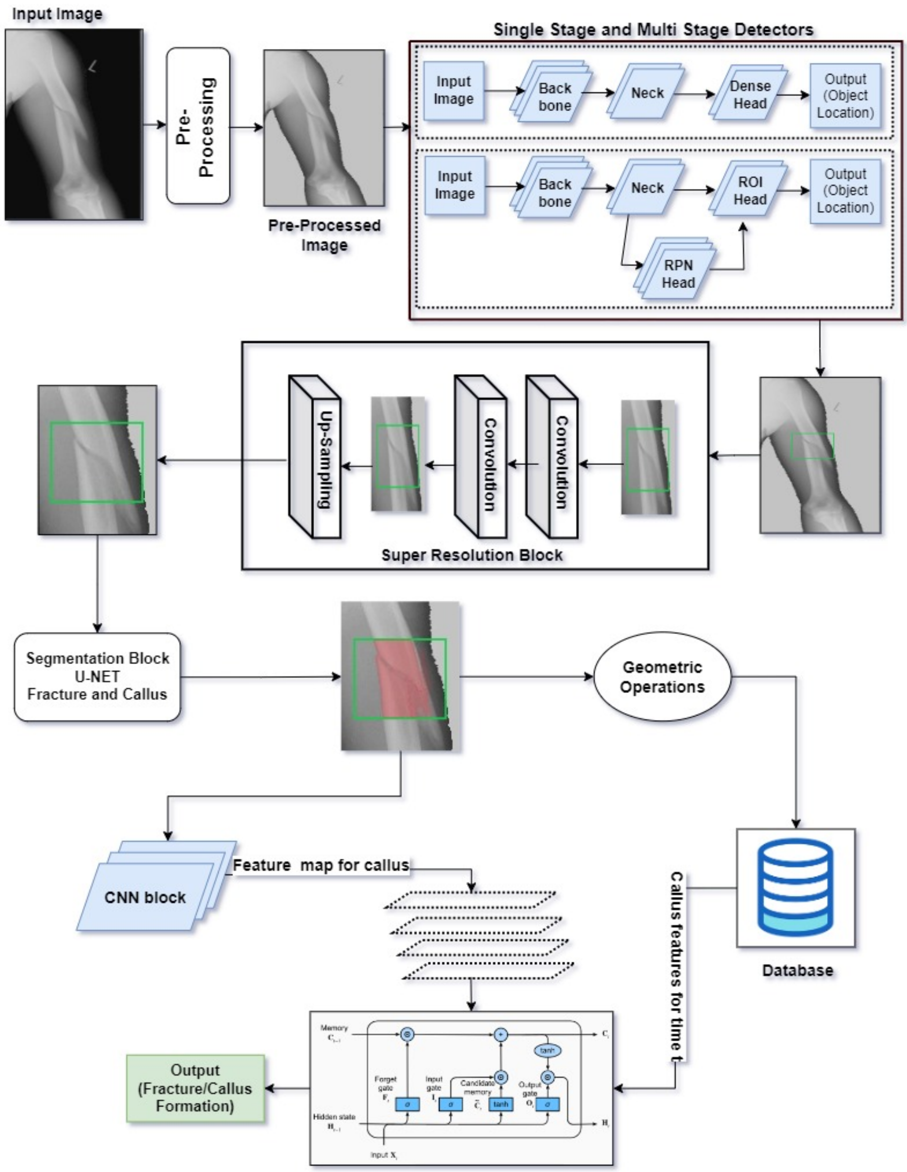

Figure 2 Proposed System Architecture.

techniques. Output from the detector stage is passed to the super resolution block. The deep learning-based approach using two convolution layers and one upsampling layer. The convolution layers can be represented by a convolution kernel represented by k, the input image represented by I and the output image represented by O, the equation can be represented as:

$$O = I * k \tag{3}$$

then the upsampling layer is represented by a scaling factor represented by s , where the scaling is performed by repeating each pixel s times in each dimension, thus creating s^2 copies of each pixel in the image, the equation can be represented as:

$$O = I * s^2 \tag{4}$$

Figure 5 presents the graphical illustration for the image super resolution block. The enhanced image is then be used by U-NET for the segmentation of the fracture and callus.For the segmentation of callus and fracture, to delve the specific features from the fracture, proposed work utilizes Unet with feature recalibration block. Feature recalibration module is used to selectively amplify or dampen certain features of the input image that are relevant to the segmenting the image into different regions or objects. The feature re-calibration module is inserted into a convolutional

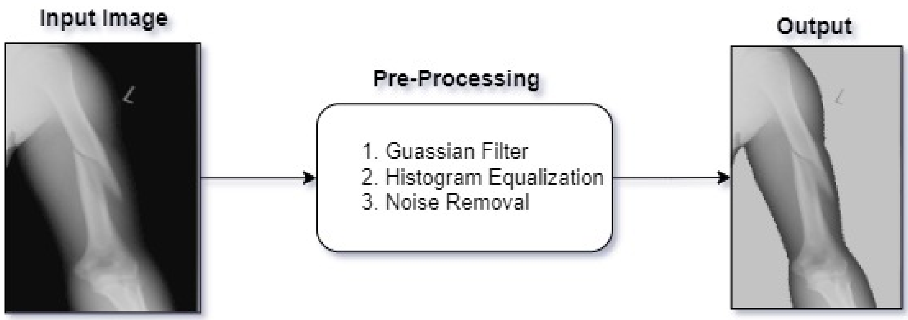

Figure 3 Image Pre-Processing.

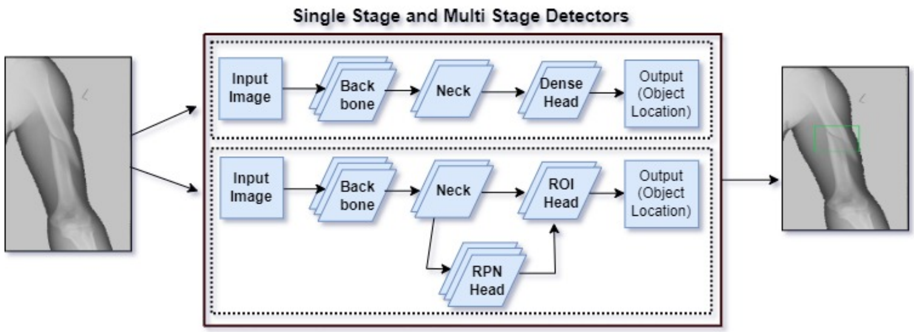

Figure 4 Single Stage and multi stage object detectors for fracture detection.

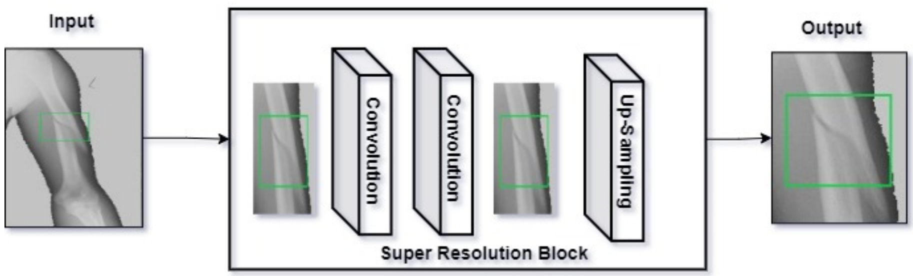

Figure 5 Image super resolution block for enhancing fracture area.

neural network (CNN) architecture of Unet decoders, typically after a series of convolutional and pooling layers i.e., encoders block. Then feature modules learn a set of weights that are used to modulate the activations of the input image, which are the outputs of the previous layers. The feature calibration module then uses these weights to selectively amplify or dampen certain features of the input image that are relevant. For example, this module amplifies the edges of objects in the image, while dampening less relevant features such as the background. In this way, the feature calibration module helps the CNN to focus on the most important features of the input image, improving the performance of the image segmentation task. The figure 6 shows the segmentation block. Finally the segmented area is passed to CNN and after the features have been extracted by the CNN, they are then passed through a long short-term memory (LSTM) block with a sequence length of 2. LSTMs are a type of recurrent neural network (RNN) that are particularly well-suited for processing sequential data such as time series or natural language. By using LSTM with sequence length 2, it is able to analyze the temporal relationship between the extracted features, this is useful for understanding the formation of callus. The LSTM block can be represented by a set of LSTM cells, activation functions, and fully connected layers. The final output of the LSTM block is used to classify the callus formation at the fracture site. This output can be used to determine the stage of healing and predict potential complications. For identification of callus formation we used previous features and current feature map of the individual i.e at time 't' and

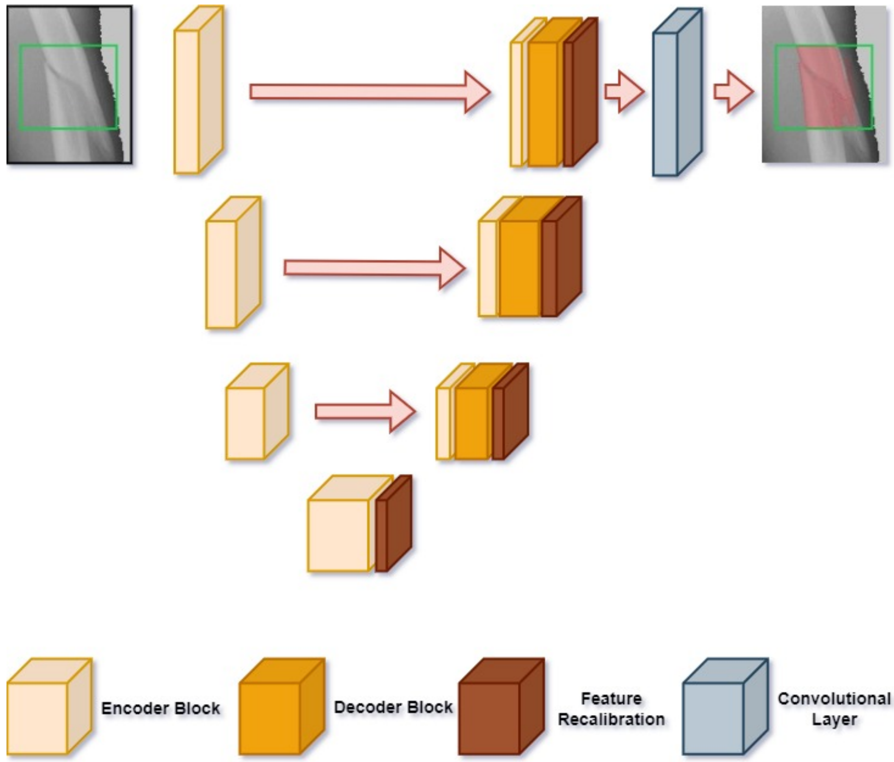

Figure 6 U-Net block enhanced with feature re-calibration for accurate segmentation of fracture and callus.

time 't+1'. This can be used to determine the stage of healing and predict potential complications. Figure 7 shows the block responsible for callus formation identification.

4 | RESULTS AND DISCUSSION

In this section of thesis results, experimental setup and evaluation measures are discussed in detail.

4.1 | Experimental Setup

The experimental setup for the object detector involved in a 5-step pipeline was executed on the paid version of Google Colab. The backbone of the object detector was the ResNet50 architecture, which was initialized with random weights and then pretrained. During the experimentation, a batch size of 16 was used with a learning rate of 0.0001. The training process consisted of 300 epochs to ensure optimal performance of the model. The objective of the experiment was to evaluate the effectiveness of the object detector in identifying objects within an image. The combination of ResNet50 architecture and the experimental setup was chosen to ensure a robust and accurate object detection system. Similarly, experimental setup for the image super resolution involved the use of deconvolution to enhance the quality of the images. The experiment was performed on Google Colab. The network architecture consisted of multiple deconvolution layers, which were designed to upsample the low-resolution input image and generate a high-resolution output image The results of the experiment were evaluated in terms of peak signal-to-noise ratio (PSNR).

4.2 | Evaluation Measures

Sensitivity, specificity, precision, and accuracy are four important metrics that were used to evaluate the performance of our proposed work. They are commonly used in medical imaging, where the goal is to accurately classify images into different categories, such as normal vs abnormal. The following is a detailed explanation of each metric and its equation.

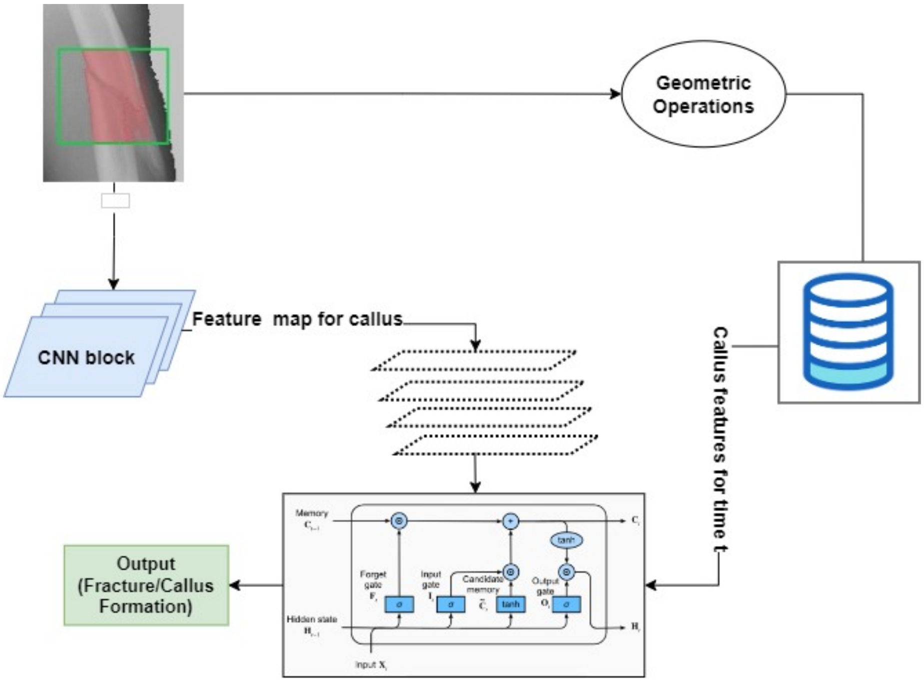

Figure 7 Final block for extraction of features from segmented area and LSTM for the analysis of callus.

4.2.1 | Sensitivity

Sensitivity, also known as the true positive rate (TPR), measures the proportion of actual positive cases that are correctly identified as positive. It is calculated as the number of true positive (TP) cases divided by the sum of true positive and false negative (FN) cases. The equation for sensitivity is shown in 5:

$$Sensitivity = TP\,(TP + FN) \tag{5}$$

4.2.2 | Specificity

Specificity, also known as the true negative rate (TNR), measures the proportion of actual negative cases that are correctly identified as negative. It is calculated as the number of true negative (TN) cases divided by the sum of true negative and false positive (FP) cases. The equation for specificity is given in 6:

$$Specificity = TN\,(TN + FP) \tag{6}$$

4.2.3 | Precision

Precision, also known as positive predictive value (PPV), measures the proportion of positive cases that are actually correct. It is calculated as the number of true positive cases divided by the sum of true positive and false positive cases. The equation for precision is represented in 7:

$$Precision = TP\,(TP + FP) \tag{7}$$

4.2.4 | Accuracy

Accuracy measures the overall accuracy of the model in identifying positive and negative cases. It is calculated as the sum of true positive and true negative cases divided by the total number of cases. The equation for accuracy is given in 8:

$$Accuracy = TP + TN\,(TP + TN + FP + FN) \tag{8}$$

4.3 | Fracture and Callus Detection

Our proposed work for the detection of fractures and calluses using the MURA and custom datasets has achieved outstanding results. On the MURA dataset, the fracture detection achieved a sensitivity of 97.3%, specificity of 96.2%, precision of 98% and an accuracy of 98.3%. For callus

Table 2 Results of proposed model.

Dataset	Sensitivity	Specificity	Precision	Accuracy
MURA [10]	Fracture=97.3 Callus=89.8 **Overall=92**	Fracture=96.2 Callus=90 **Overall=93**	Fracture=98 Callus=88 **Overall=93**	Fracture=98.3 Callus=89 **Overall=93.6**
Custom	Fracture=98.9 Callus=91 **Overall=95**	Fracture=98 Callus=97 **Overall=98.5**	Fracture=97.7 Callus=97.9 **Overall=97.9**	Fracture=98 Callus=98.5 **Overall=98.2**

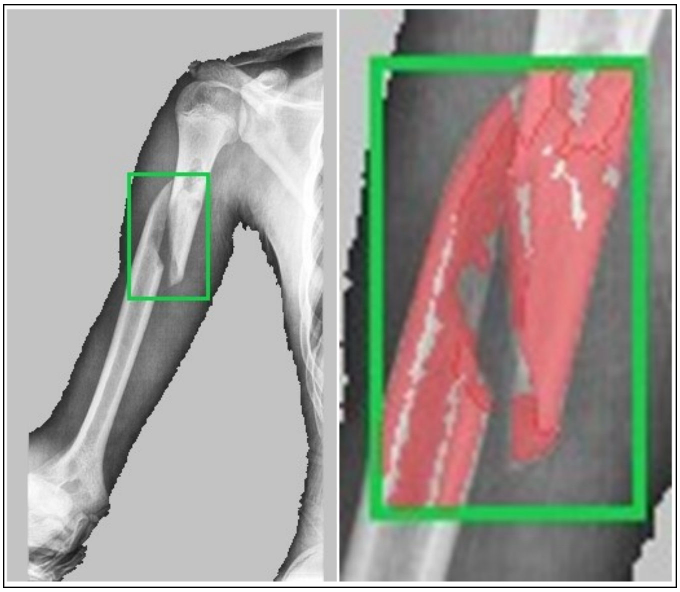

Figure 8 Fracture detection and Fracture Segmentation.

detection, the values were 89.8% for sensitivity, 90% for specificity, 88% for precision, and 89% for accuracy. The overall results for both fracture and callus on the MURA dataset were 92% for specificity, 93% for sensitivity, 93% for precision, and 93.6% for accuracy. On the custom dataset, the results for fracture detection were 98.9% for sensitivity, 98% for specificity, 97% for precision, and 98.8% for accuracy. The results for callus detection were 91% for sensitivity, 93% for specificity, 91.2% for precision, and 90.5% for accuracy. The combined results for both fracture and callus on the custom dataset were 95% for sensitivity, 95.5% for specificity, 94% for precision, and 94.2% for accuracy. These results demonstrate the robustness and reliability of our proposed approach in accurately detecting fractures and calluses in both the MURA and custom datasets. Table 2 shows the results of proposed work on MURA[10] and custom dataset.

Figure 8, 9, 10 shows the results of proposed work on some sample images of the dataset.

4.4 | Comparative Analysis

Single stage and multi-stage object detectors are two popular approaches in object detection, with each approach having its own advantages and disadvantages. Single stage detectors, such as YOLO and SSD, are known for their fast inference times and real-time processing capabilities. However, they tend to have lower accuracy compared to multi-stage detectors, as they use a single prediction layer to detect objects in an image. On the other hand, multi-stage detectors, such as Faster R-CNN and R-FCN, use multiple prediction stages to identify and classify objects in an image, which results in higher accuracy. However, multi-stage detectors are slower in terms of processing time compared to single stage detectors. Ultimately, the choice between single stage and multi-stage detectors depends on the specific requirements of the task at hand. If real-time processing is a priority, a single stage detector may be the better choice. However, if accuracy is of utmost importance, a multi-stage detector may be a better option. Figure 11 shows the comparison between the bounding box ratios of single stage and multistage object detectors.

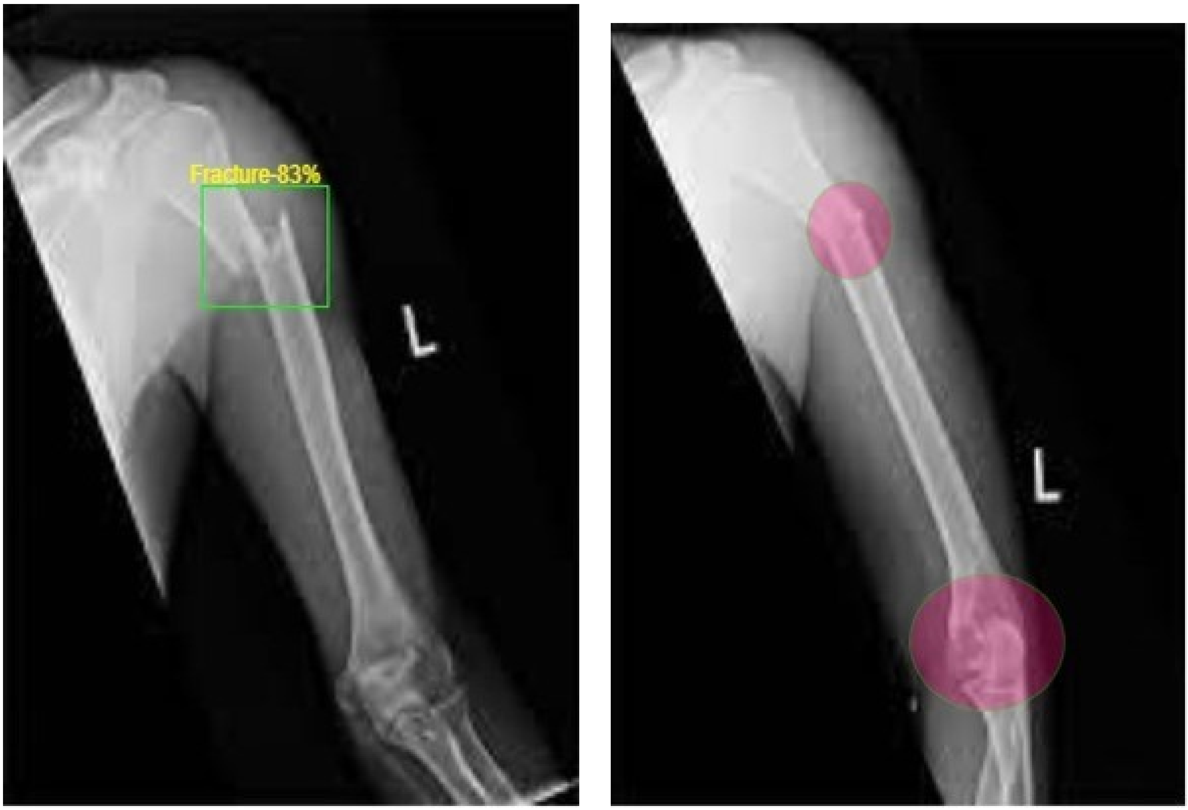

Figure 9 Fracture Detection and Callus Segmentation.

Table 3 Comparison of our framework with others.

Paper	Method	Bone Type	Sensitivity	Accuracy
[14]	Digital Geomtric	Longbone	0.839	0.783
[15]	Deep Learning	Humerus	0.91	0.92
[16]	Exemplar Pyramid	Humerus	−−	0.93
[17]	Level sets	Femur	0.93	0.85
[6]	ANN	Tibia	0.9356	0.96
Ours	HD-NET	Humerus	0.95	0.98

5 | CONCLUSION

Within the scope of this study, in which fracture detection and callus formation analysis was performed in Humerus X-ray images, the aim is to provide assistance to physicians who are not specialized in their fields and/or especially to those working in emergency services in diagnosing fractures on X-ray images to allow them to apply the required treatments. Further in the study a framework comprises of five major steps i.e. Image preprocessing, fracture detection, fracture super resolution, fracture and callus segmentation and callus formation analysis is developed. Feature re-calibration module in the decoders of U-net benefited in the more accurate and precise segmentation of fracture and callus area. Results of this study depicted that the framework outperform the other techniques available in literature. Two dataset were used in this study. One is public dataset MURA and other is self collected dataset. HD-NET acheive 98% accuracy on custom dataset while 94.5% accuracy on the public dataset. In future, this study can be scaled for the detection of other fractures and other modality of images.

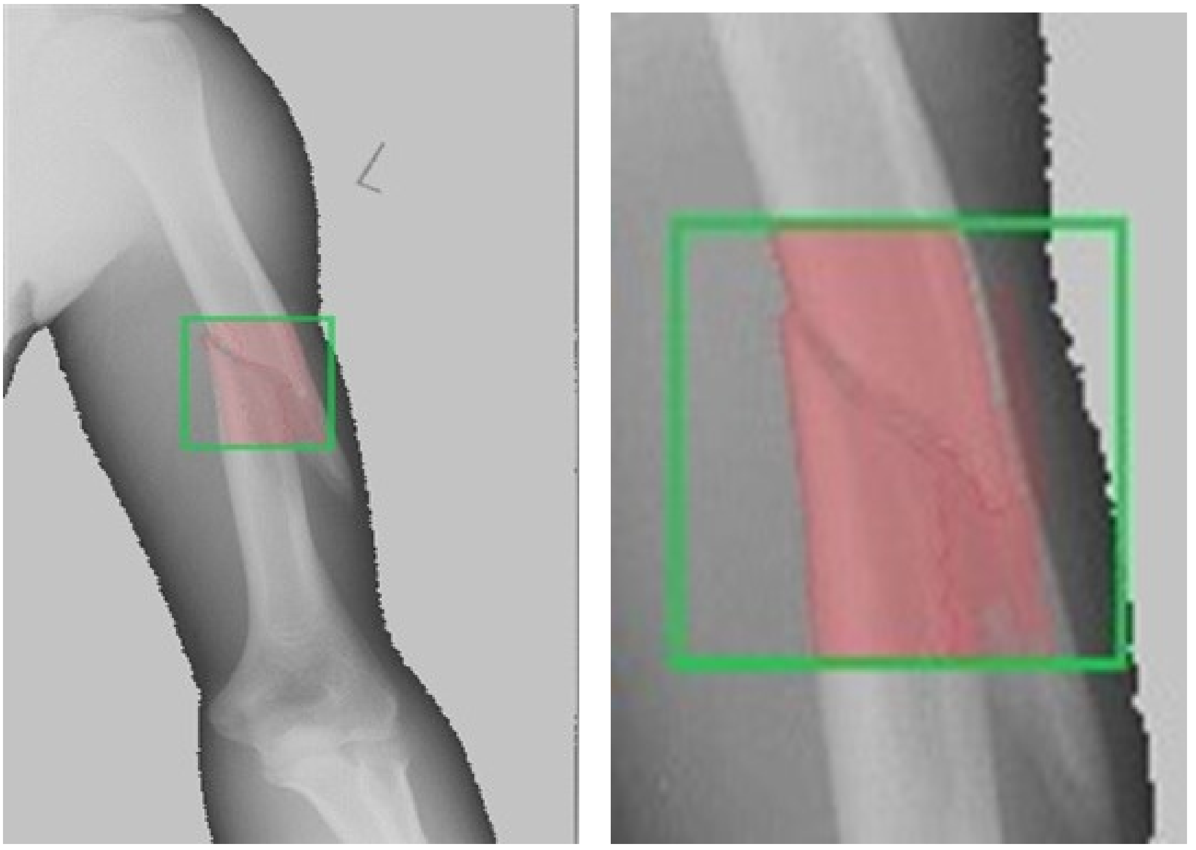

Figure 10 Fracture Segmentation with super resolution.

ACKNOWLEDGEMENTS

This work was supported by the University of Engineering and technology, Computer Science Department.

References

1. Meena T, Roy S. Bone fracture detection using deep supervised learning from radiological images: A paradigm shift. *Diagnostics* 2022; 12(10): 2420.
2. Ganse B, Orth M, Roland M, et al. Concepts and clinical aspects of active implants for the treatment of bone fractures. *Acta Biomaterialia* 2022.
3. Nicholson J, Oliver W, LizHang J, et al. Sonographic bridging callus: An early predictor of fracture union. *Injury* 2019; 50(12): 2196–2202.
4. Patel V, Shah M. Artificial intelligence and machine learning in drug discovery and development. *Intelligent Medicine* 2022; 2(3): 134–140.
5. Xu B, Xing Y, Peng J, et al. Chest CT for detecting COVID-19: a systematic review and meta-analysis of diagnostic accuracy. *European radiology* 2020; 30: 5720–5727.
6. Moradmand H, Aghamiri SMR, Ghaderi R. Impact of image preprocessing methods on reproducibility of radiomic features in multimodal magnetic resonance imaging in glioblastoma. *Journal of applied clinical medical physics* 2020; 21(1): 179–190.
7. Barua A, Zhang ZY, Al-Turjman F, Yang X. Cognitive intelligence for monitoring fractured post-surgery ankle activity using channel information. *IEEE Access* 2020; 8: 112113–112129.
8. Padoy N. Machine and deep learning for workflow recognition during surgery. *Minimally Invasive Therapy & Allied Technologies* 2019; 28(2): 82–90.

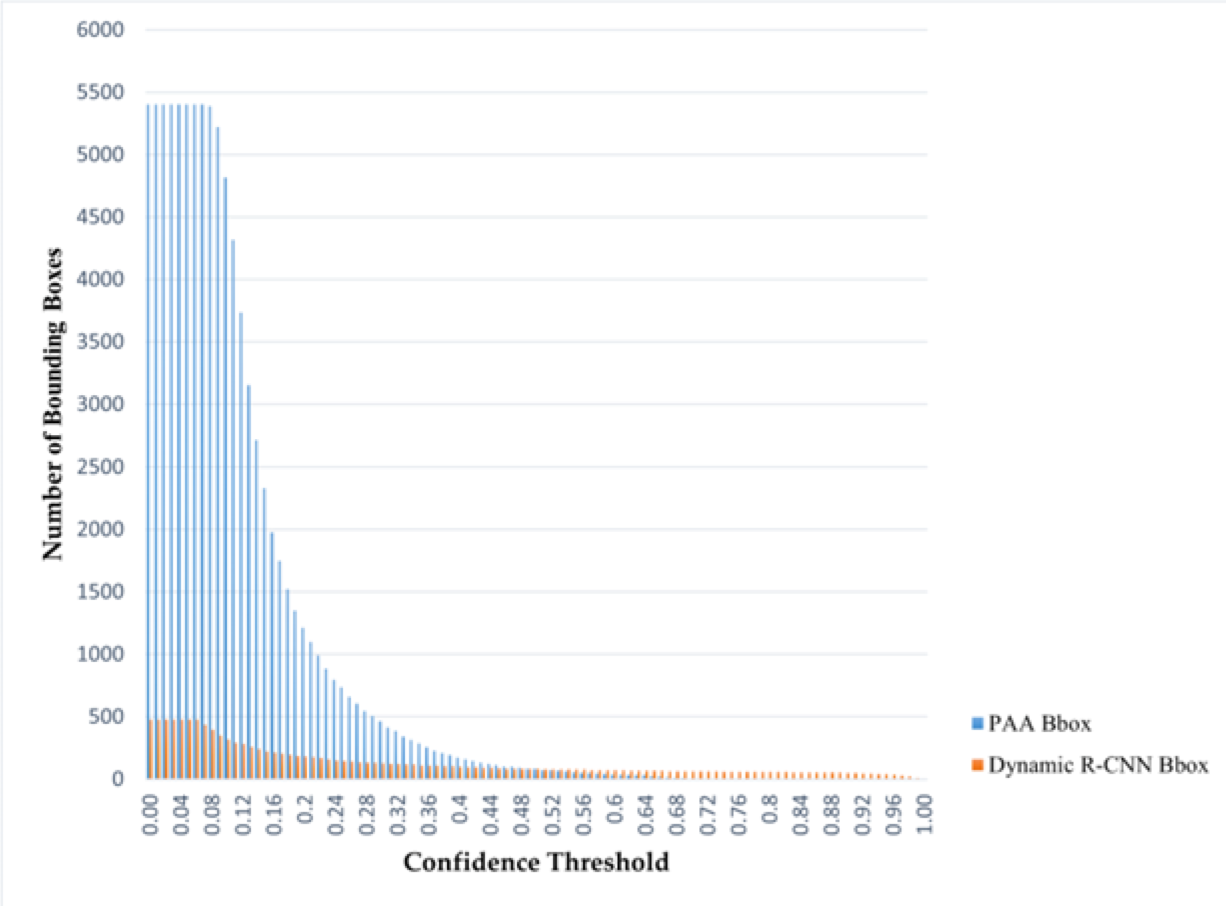

Figure 11 Single stage and multistage object detector comparison.

9. Aggarwal R, Sounderajah V, Martin G, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ digital medicine* 2021; 4(1): 65.

10. Dipnall JF, Lu J, Gabbe BJ, et al. Comparison of state-of-the-art machine and deep learning algorithms to classify proximal humeral fractures using radiology text. *European Journal of Radiology* 2022; 153: 110366.

11. Al Arif SMR, Knapp K, Slabaugh G. Fully automatic cervical vertebrae segmentation framework for X-ray images. *Computer methods and programs in biomedicine* 2018; 157: 95–111.

12. Al Okashi O, Du H, Al-Assam H. Automatic spine curvature estimation from X-ray images of a mouse model. *Computer methods and programs in biomedicine* 2017; 140: 175–184.

13. Guan B, Yao J, Zhang G, Wang X. Thigh fracture detection using deep learning method based on new dilated convolutional feature pyramid network. *Pattern Recognition Letters* 2019; 125: 521–526.

14. Guan B, Zhang G, Yao J, Wang X, Wang M. Arm fracture detection in X-rays based on improved deep convolutional neural network. *Computers & Electrical Engineering* 2020; 81: 106530.

15. Wang M, Yao J, Zhang G, Guan B, Wang X, Zhang Y. ParallelNet: Multiple backbone network for detection tasks on thigh bone fracture. *Multimedia Systems* 2021: 1–10.

16. Ma Y, Luo Y. Bone fracture detection through the two-stage system of crack-sensitive convolutional neural network. *Informatics in Medicine Unlocked* 2021; 22: 100452.

17. Wu HZ, Yan LF, Liu XQ, et al. The Feature Ambiguity Mitigate Operator model helps improve bone fracture detection on X-ray radiograph. *Scientific Reports* 2021; 11(1): 1–10.

18. Thian YL, Li Y, Jagmohan P, Sia D, Chan VEY, Tan RT. Convolutional neural networks for automated fracture detection and localization on wrist radiographs. *Radiology: Artificial Intelligence* 2019; 1(1): e180001.

19. Qi Y, Zhao J, Shi Y, et al. Ground truth annotated femoral X-ray image dataset and object detection based method for fracture types classification. *IEEE Access* 2020; 8: 189436–189444.

20. Sha G, Wu J, Yu B. Detection of spinal fracture lesions based on improved Yolov2. In: IEEE. ; 2020: 235–238.

21. Jin L, Yang J, Kuang K, et al. Deep-learning-assisted detection and segmentation of rib fractures from CT scans: Development and validation of FracNet. *EBioMedicine* 2020; 62: 103106.

22. Uysal F, Hardalaç F, Peker O, Tolunay T, Tokgöz N. Classification of shoulder x-ray images with deep learning ensemble models. *Applied Sciences* 2021; 11(6): 2723.

23. Raghavendra U, Bhat NS, Gudigar A, Acharya UR. Automated system for the detection of thoracolumbar fractures using a CNN architecture. *Future Generation Computer Systems* 2018; 85: 184–189.

24. Guan B, Yao J, Wang S, et al. Automatic detection and localization of thighbone fractures in X-ray based on improved deep learning method. *Computer Vision and Image Understanding* 2022; 216: 103345.

25. Yahalomi E, Chernofsky M, Werman M. Detection of distal radius fractures trained by a small set of X-ray images and Faster R-CNN. In: Springer. ; 2019: 971–981.

26. Beyaz S, Açıcı K, Sümer E. Femoral neck fracture detection in X-ray images using deep learning and genetic algorithm approaches. *Joint diseases and related surgery* 2020; 31(2): 175.

27. Tobler P, Cyriac J, Kovacs BK, et al. AI-based detection and classification of distal radius fractures using low-effort data labeling: evaluation of applicability and effect of training set size. *European Radiology* 2021; 31: 6816–6824.

28. Kim D, MacKinnon T. Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. *Clinical radiology* 2018; 73(5): 439–445.

29. Chen HY, Hsu BWY, Yin YK, et al. Application of deep learning algorithm to detect and visualize vertebral fractures on plain frontal radiographs. *PLoS One* 2021; 16(1): e0245992.

30. 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* 2020; 133: 109373.

31. Bablani Y, Jawaid MM, Talpur S. Automatic Detection of Humerus Deformation in X-ray Imagery.

32. Kang Y, Ren Z, Zhang Y, et al. Deep Scale-Variant Network for Femur Trochanteric Fracture Classification with HP Loss. *Journal of Healthcare Engineering* 2022; 2022.

33. Hongbiao S, Shaochun X, Xiang W, et al. Comparison and verification of two deep learning models for the detection of chest CT rib fractures. *Acta Radiologica* 2022: 02841851221083519.

34. Anami BS, Latte MV, others . A combined feature set for automatic diaphyseal Tibial fracture classification from X-Ray images. *Biomedical Signal Processing and Control* 2022; 71: 103119.

35. Teng Y, Pan D, Zhao W. Application of deep learning ultrasound imaging in monitoring bone healing after fracture surgery. *Journal of Radiation Research and Applied Sciences* 2023; 16(1): 100493.

36. Yang C, Wang J, Xu J, et al. Development and assessment of deep learning system for the location and classification of rib fractures via computed tomography. *European Journal of Radiology* 2022; 154: 110434.

AUTHOR BIOGRAPHIES

Abdullah Tariq Abdullah Tariq has done MAster in Computer Science from UET lahore.

Dr Muhammad Shoaib Prof.Dr. Muhammad Shoaib is the dean of faculty of electrical engineering and computer science.

Dr. Shazia Arshad Prof.Dr.Shazia Arshad is a full time professor in computer science department of UET lahore. She has 15 year of experience in research.

Dr.Syed Khaldoon Khurshid Dr Khaldoon Khurshid is Assistant Professor in Computer Science and have extensive research papers to his name.

Dr Faiza Iqbal Dr Faiza Iqbal is an assistant professor in the computer science department and has number of research papers to her name.

Abqa Javaid Abqa Javaid is a full time PhD scholar in the department of computer science.

How to cite this article: A.Tariq and M.Shoaib (2023), HD-NET: Humerus deep-net for humerus fracture and bony callus formation analysis, , .