INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 10, Issue 4 - V10I4-1189)
Available online at: https://www.ijariit.com

Detection and Classification of Bone Fractures in X-Ray Images using Yolov8

Vidya Shree G
vidyashreeg41298@gmail.com
Independent Researcher

ABSTRACT

Hospitals often handle a significant volume of bone fractures, which form a major segment of their medical caseload. X-ray imaging is critical for the accurate detection and classification of these fractures, playing a pivotal role in guiding subsequent treatment strategies. This study investigates the potential of the YOLO (You Only Look Once) deep learning framework to advance the automatic detection and classification of bone fractures in X-ray images. Specifically, it aims to enhance the YOLOv8 model's ability to identify various fracture types by training it on an extensive and diverse set of labelled X-ray images and employing data augmentation techniques to improve performance. The diagnostic tool is expected to support radiologists by providing timely and accurate insights, thereby streamlining the decision-making process. The results show that the YOLOv8 model, enhanced with data augmentation outperforms the standard YOLOv8 model in both accuracy and speed for fracture detection and classification. This approach not only promises to elevate diagnostic standards but also has the potential to reduce the workload on radiologists, leading to more effective and efficient patient care.

Keywords - *You Only Look Once (YOLO), Bone fractures, Convolutional Neural Network (CNN), X-rays, Image analysis, Deep learning, Data Augmentation, YOLOv8*

1. INTRODUCTION

Bone fractures occur when a bone is subjected to excessive stress or force. Fractures can happen in various different ways like a fall, accident, repeated stress from activities, conditions like lack of physical activities, low body weight that weaken the bones like osteoporosis. Fractures often tend to be quite painful since they involve not just the bone but also the surrounding tissue. The pain can be quite sharp and resulting symptoms usually include swelling, bruising and difficulty moving the injured areas. Treating the bone fracture depends on how severe it is. Immediate medical attention is often necessary to ensure proper alignment and to prevent further complications.

X-ray imaging is a key tool for detecting bone fractures by using high-energy radiation to create images of the body's internal structures. Fig.1 shows the X-ray image of a broken bone. The X-rays pass through the body, with dense bones appearing white on the resulting image due to their higher absorption of X-rays, while softer tissues show up darker. Fractures are visible as dark lines or gaps in the bone's white structure. Radiologists analyse these images to assess the typeand severity of the fracture, which informs the appropriate treatment, such as casting, splinting, or surgery.

Fig.1. X-ray image of bone fracture

Manually analysing x-ray images to identify and classify fractures is often slow and prone to human error. To address these issues, there's a growing need for automated systems that can detect and classify bone fractures. Such systems enhance traditional methods by offering quicker diagnosis, minimising the chances of overlooking or misclassifying fractures, and aiding radiologists in making more accurate and informed decisions.

The proposed model leverages YOLOv8 for detecting and classifying bone fractures. Data augmentation is performed to enhance the dataset for optimal results. By employing these methods, the model precisely identifies fractures in X-ray images and highlights them with bounding boxes for classification. Fig.2 illustrates fracture detection and classification using the proposed model. This method achieves both high accuracy and speed in identifying and classifying fractures.

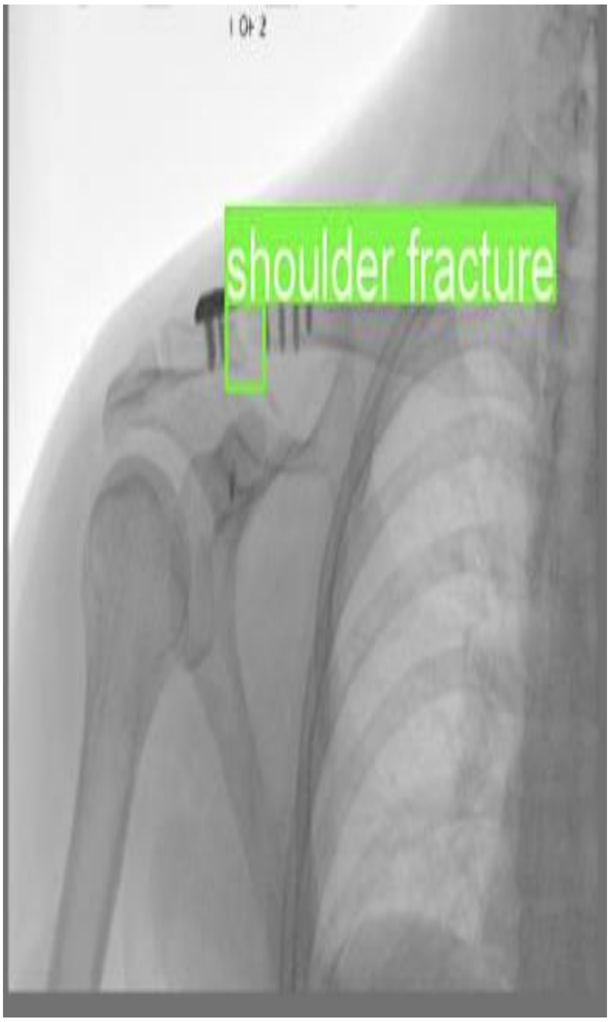

Fig.2. Bone fracture detection and classification

1.1 Problem Statement
Traditionally, bone fractures are identified by examining X-ray images manually. This method depends on human interpretation, which can lead to overlooked fractures and delays in diagnosis and treatment.

To address these issues and minimise the risk of missing critical fractures, an automated, real-time detection system is crucial. This system must be both reliable and highly accurate.

The proposed solution aims to efficiently identify fractures in X-ray images using advanced deep learning algorithms and computer vision techniques. It detects fractures, highlights the affected areas with bounding boxes, and classifies the type of bone fracture. This method not only accelerates the diagnostic process but also enhances precision in fracture detection, thereby improving patient outcomes and reducing the workload for medical professionals.

2. LITERATURE REVIEW

2.1 Radiologist shortage leaves patient care at risk, warns royal college
In this paper [1], the author highlights the significant shortage of radiologists and its risks to patient care. The Royal College has raised alarms about how this shortage could compromise the quality of medical imaging and delay diagnosis. The lack of radiologists may lead to increased workloads, longer wait times for imaging results, and potentially, adverse effects on patient outcomes. The paper stresses the urgent need for solutions to address this shortage and ensure that patient care remains effective and timely despite the growing demand for radiological services.

2.2 Artificial Intelligence Solutions for Analysis of X-ray Images
This paper [2], reviews several models for X-ray images analysis, including convolutional neural networks (CNNs) and deep learning algorithms. These models demonstrated high accuracy in detecting abnormalities, such as fractures and tumours, with performance often surpassing traditional methods. The study emphasises AI's potential to improve radiological practices while acknowledging areas needing further research.

2.3 The Evaluation of Bone Fracture Detection of YOLO Series
In this paper [3], the authors compared several versions of the YOLO object detection model. They evaluated YOLOv3, YOLOv4, and YOLOv5. Each version offers different improvements in terms of speed and accuracy. YOLOv3 was assessed for its balance between detection speed and accuracy, YOLOv4 for its enhancements in both performance and efficiency, and YOLOv5 for its latest updates and optimizations in detection capabilities.

2.4 Abnormal Object Detection In Thoracic X-Ray Using You Only Look Once (YOLO)
The paper [4] employs the YOLO algorithm for abnormal object detection in thoracic X-rays. The authors evaluated YOLO's performance using metrics such as precision, recall, and mean Average Precision (mAP). The results indicate high precision and recall rates, demonstrating YOLO's effectiveness in accurately identifying and localising abnormalities while maintaining fast processing times.

2.5 Conference on Computer Vision and Pattern Recognition
This paper [5], mentions several experiments conducted to validate the effectiveness of Grid R-CNN. They tested the framework on standard object detection benchmarks like COCO and PASCAL VOC. The experiments showed that Grid R-CNN outperformed traditional R-CNN models, achieving higher accuracy in object localization and detection. The authors also evaluated their method's performance across different settings and ablated various components to understand their impact, demonstrating Grid R-CNN's robustness and efficiency.

2.6 Musculoskeletal Images Classification for Detection of Fractures Using Transfer Learning

In this paper [6], the application of transfer learning for classifying musculoskeletal images to identify fractures was explored. The authors review various methods and technologies in medical imaging, emphasising advancements in deep learning techniques that enhance diagnostic accuracy. They highlight the effectiveness of transfer learning in leveraging pre-trained models to improve performance on specific fracture detection tasks, offering a promising approach to automating and refining fracture diagnosis in radiological images.

3. IMPLEMENTATION

3.1 Dataset

The dataset consists of around 10,000 images featuring both normal bone X-rays and images showing various types of bone fractures. The images are categorised into distinct classes based on fracture types, including elbow positive, forearm fracture, shoulder, bow fracture, dislocation, and finger positive. The dataset is divided into three image folders: train (7,531 images), test (838 images), and validate (1,601 images). The model has been established to detect and categorise these bone fractures effectively.

3.2 Image annotation

Image annotation involves tagging an input image with predefined labels to provide the model with information about the content of the image. The labels highlight key features that will be used to train the model through supervised machine learning techniques. During the annotation process, metadata is also incorporated into the dataset, which is pivotal for enhancing the model's accuracy and precision.

Every image in the dataset is labelled with bounding boxes or pixel-level segmentation masks, highlighting the position and scope of detected fractures. The annotation files are present in a separate directory called labels. The annotation file name is the same as the image name in the images directory. The format used here is YOLO.txt for the annotation file.

3.3 YOLO - You Only Look Once

YOLO represents a series of deep learning models crafted for rapid, real-time object detection. It was developed by Joseph Redmon using a custom framework called Darknet. Darknet is a versatile framework for research written in low-level languages and is known for producing highly effective real-time object detectors.

YOLO was groundbreaking because it integrated both bounding box prediction and class label identification into a single, end-to-end network, unlike traditional models which handle these tasks separately in two stages. While models like Faster RCNN use a two-stage approach, YOLO processes object detection in one go, which typically makes it faster and more efficient.

YOLOv8 is one of the latest advanced models in the YOLO series, designed for tasks such as object detection, image classification, and instance segmentation. Created by Ultralytics, the same team behind the widely recognized YOLOv5, YOLOv8 brings several architectural upgrades and enhancements to developer experience compared to its predecessor. Fig.3 shows the latest object detections using YOLOv8.

Fig.3. Object detection in real-time using YOLOv8

Our object detection implementation utilises YOLOv8, and its architecture will be detailed in the following section.

3.4 YOLOv8 Model Architecture

The model consists of 3 essential blocks - head, neck and backbone. This is depicted in Fig.4. Everything that happens in the model happens in these three blocks and the functionalities of the three blocks are discussed below.

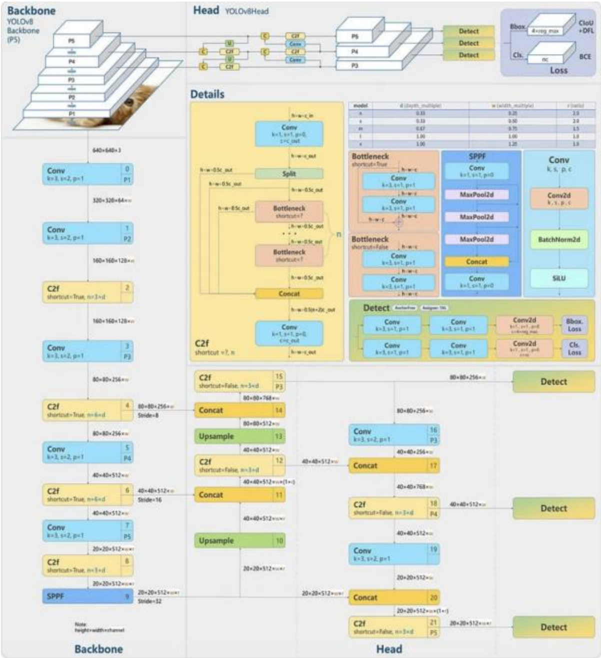

Fig.4 YOLOv8 architecture

3.4.1 Backbone

Backbone is the CNN (Convolutional Neural Network). This is responsible for the meaningful feature extraction from the input images and classification. CNN is made up of a multitude of layers; the input layer, hidden layers, and the final output layer. In the network, nodes are linked to each other and each has a weight and a threshold. A node becomes active and forwards data to the next layer only if its output exceeds the threshold value. If the output is below the threshold, no data is transmitted further.

CNN working is as shown in Fig.5. A feature map is generated by using various filters and then applying a Rectified Linear Unit (ReLU) function to introduce non-linearity.

The ReLU activation function is defined as:

$$f(x) = \max(0, x)$$

where:

- x is the input to the activation function.
- $f(x)$ is the output after applying ReLU.

Pooling is performed on these feature maps, and the resulting pooled images are flattened into a single long vector. This vector is then fed into a fully connected neural network for further processing. The network is trained using forward and backward propagation over multiple epochs, and the final output is produced by the fully connected layer. YOLO object detection utilises this CNN-based approach.

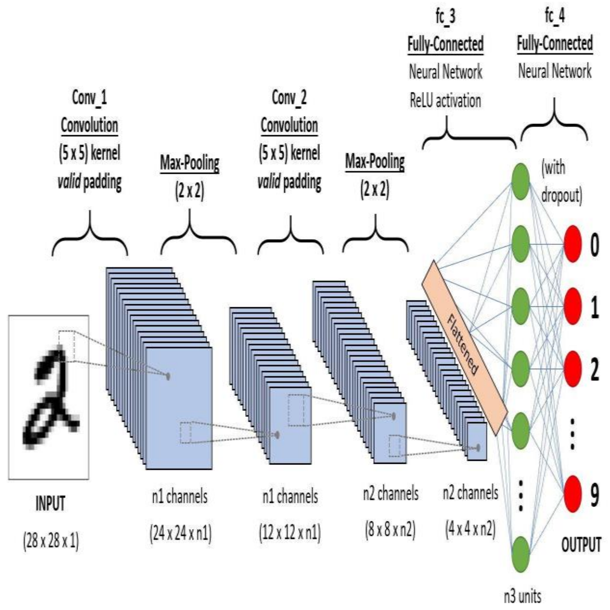

Fig.5 Working of CNN

The backbone, or feature extractor, is crucial for identifying and extracting significant features from input data. It starts by detecting basic patterns like edges and textures in its early layers. As you progress through the layers, it captures features at various levels of detail, offering a comprehensive hierarchical representation. YOLOv8 utilises a specialised backbone called CSPDarknet53, which incorporates Cross Stage Partial(CSP) connections to enhance information flow and increase accuracy. The core idea behind CSP connections is to split the feature map into two parts, process them separately through different stages of the network, and then merge them back.

3.4.2 Neck

The neck in a neural network connects the backbone to the head, playing a crucial role in merging and processing features. It gathers feature maps from various stages of the backbone and combines them to handle objects of different sizes effectively. This process involves fusing features at different scales and integrating contextual information to enhance detection accuracy. Additionally, the neck reduces the spatial resolution and dimensionality of the features, which speeds up computation but can impact model quality. In YOLOv8, the neck uses a C2f module instead of the traditional Feature Pyramid Network (FPN), which helps improve detection accuracy by better integrating high-level semantic features with low-level spatial details, particularly benefiting the detection of small objects.

The core concept of C2F is to improve the integration of features from different levels of a network. In conventional architectures, there can be a disconnect between earlier layers and the final prediction stage, leading to missed opportunities for utilising important features. C2F addresses this issue by promoting interactions between features from various stages, thereby boosting the model's capacity to harness relevant information throughout the network.

3.4.3 Head

The head is the final component of the network, tasked with producing the output for object detection tasks. It generates bounding boxes around potential objects in an image, assigns confidence scores to these boxes indicating the likelihood of an object being present, and categorises the objects. In YOLOv8, the head uses several detection modules to predict bounding boxes, objectness scores, and class probabilities for each grid cell in the feature map. These predictions are then combined to deliver the final detection results. Additionally, the head now handles classification and regression separately, as shown in Fig.4, which enhances the model's performance.

3.5 PROPOSED MODEL

The proposed model aims to detect and classify bone-fractures using the x-ray images . If any fracture is detected in the image, a bounding box will be drawn around the image with the corresponding type.

The proposed model is built upon the YOLOv8 model. The pre-trained model significantly sped up convergence and improved performance. The number of epochs and patience has been varied multiple times to get the best results. The model performed best at 100 epochs and patience of 25.

Data augmentation for X-ray images, such as by applying colour jittering, plays a vital role in machine learning and medical imaging. This process involves modifying the images through different techniques to expand and diversify the dataset artificially. Colour jittering has been added to enhance the visibility of the bone anomaly. Due to the varying contrast levels(high, normal and low), and to ensure the augmentation techniques do not alter the clinical significance of the anomalies or introduce artefacts or distortions in the x-ray images affecting the alignment of anatomical structures and introducing unrealistic features, Random Contrast Adjustment technique with a range of 0.65 to 1.35 has been selected to be applied after experimentations in various other ranges. The contrast values can fall between 0.0 to 3.0.

Contrast adjustment changes the intensity range of the image. The augmented image I_{aug} after applying colour jittering and contrast adjustment can be represented as

$$I_{contrast} = Clip(\alpha \cdot Jitter(I,c1,c2,c3)+\beta)$$

where:

- $Jitter(I,c1,c2,c3)$ applies colour jittering to the original image.
- α is the contrast factor within [0.65, 1.35].
- β is typically 0, unless a brightness adjustment is also needed.

This approach notably improved the performance of the model and gave the best results for maintaining the clinical relevance and visibility of anomalies. Fig.6 shows the before(left) and after(right) of an x-ray image after data augmentation with random contrast adjustment.

Fig.6 The original and augmented image after applying Random Contrast Adjustment

3.6 METHODOLOGY
The flow diagram of the proposed model is as shown in Fig.7

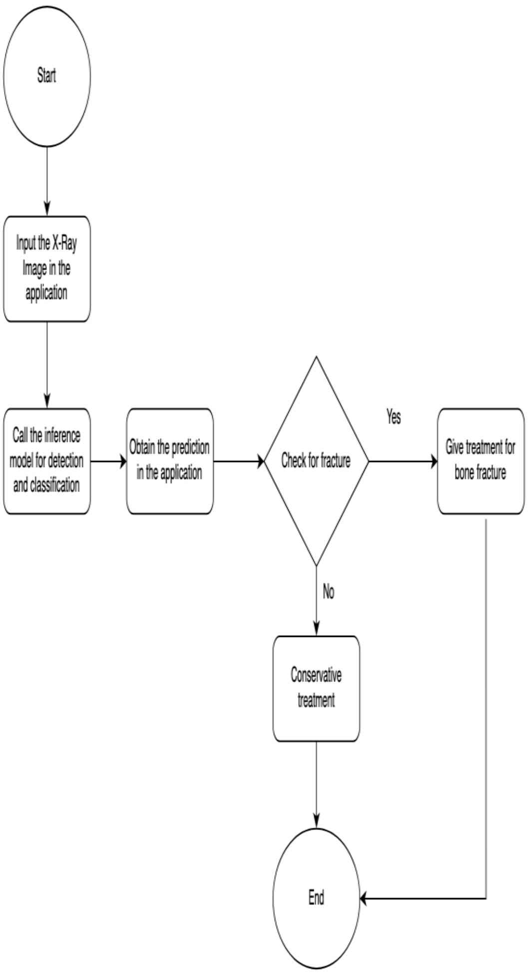

Fig.7 Proposed model flow chart

The steps followed for detecting and classifying the bone-fractures in the x-ray images are as follows:

i. Image collecting and preprocessing
ii. YOLOv8 model configuration for custom object detection.
iii. Apply data augmentation on the training dataset
iv. Training the YOLOv8 model
v. Validating and testing the trained model
vi. Evaluating the accuracy
vii. Saving the inference model
viii. Deploying the model using frameworks like react and react native

4. EXPERIMENTAL RESULTS
Fig.8 shows the training results of the proposed model. The evaluation metrics includes the mAP, mean average precision which assesses their effectiveness by summarising precision and recall across various classes and detection thresholds. The custom YOLOv8 model achieved a mAP of 0.68 compared to the standard YOLOv8 model that achieved a mAP of 0.634. This increase in the mAP for the custom model indicates a significant improvement in the performance and overall accuracy.

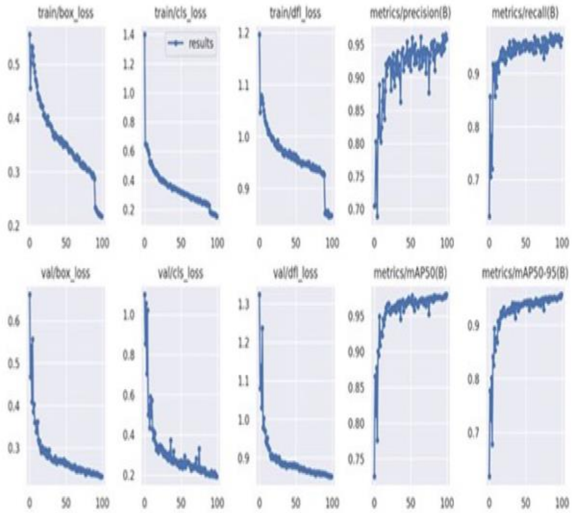

Fig.8 Results of training the custom YOLOv8 model

Once the image is input to the inference model in the application, it checks for fractures. If there are any, a box is drawn around the area and the type is specified. Fig.9 shows the finger positive bone fracture.

Fig.9 Fingers Positive bone fracture detection

The inference model accurately predicts the bone fractures and classifies them. Once the type of fracture is received in the application, further treatment can be provided based on the severity. Fig.10 shows the results of forearm bone fracture classification. It can identify the injuries in multiple spots.

Fig.10 Multiple fractures detection using the model

5. CONCLUSION

YOLOv8 is utilised for detecting and classifying bone fractures in X-ray images. It processes these images in real-time at approximately 45 frames per second, offering both speed and precision. This model excels in providing high accuracy while maintaining quick performance. Once a fracture is identified, treatment can begin promptly. This approach not only shortens the time between X-ray analysis by radiologists and the initiation of treatment but also reduces the likelihood of human error.

The model's performance improves with a larger dataset, allowing it to identify a broader range of fracture types and classifications. Customising the algorithm can enhance its efficiency and lead to the development of more advanced architectures. By expanding the dataset, we can achieve better results, minimise overfitting, and improve the handling of rare cases. The success of the model is largely influenced by the quality and variety of the data, with a more extensive dataset contributing to the creation of more robust deep learning models.

REFERENCES

[1] Rimmer A. Radiologist shortage leaves patient care at risk, warns royal college. BMJ. 2017 Oct 11;359:j4683. doi: 10.1136/bmj.j4683. PMID: 29021184.

[2] Adams SJ, Henderson RDE, Yi X, Babyn P. Artificial Intelligence Solutions for Analysis of X-ray Images. Can Assoc Radiol J. 2021 Feb;72(1):60-72. doi: 10.1177/0846537120941671. Epub 2020 Aug 6. PMID: 32757950.

[3] P. Samothai, P. Sanguansat, A. Kheaksong, K. Srisomboon and W. Lee, "The Evaluation of Bone Fracture Detection of YOLO Series," 2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), Phuket, Thailand, 2022, pp. 1054-1057, doi: 10.1109/ITC-CSCC55581.2022.9895016.

[4] W. A. K. Adji, A. Amalia, H. Herriyance and E. Elizar, "Abnormal Object Detection In Thoracic X-Ray Using You Only Look Once (YOLO)," 2021 International Conference on Computer System, Information Technology, and Electrical Engineering (COSITE), Banda Aceh, Indonesia, 2021, pp. 118-123, doi: 10.1109/COSITE52651.2021.9649500.

[5] Lu, X., Li, B., Yue, Y., Li, Q. & Yan, J. Grid r-cnn. Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 7363–7372 (2019).

[6] Kandel I, Castelli M, Popovič A. Musculoskeletal Images Classification for Detection of Fractures Using Transfer Learning. J Imaging. 2020 Nov 23;6(11):127. doi: 10.3390/jimaging6110127. PMID: 34460571; PMCID: PMC8321195.

[7] Mohamed Khalifa, Mona Albadawy, AI in diagnostic imaging: Revolutionising accuracy and efficiency, Computer Methods and Programs in Biomedicine Update, Volume 5, 2024, 100146, ISSN 2666-9900, https://doi.org/10.1016/j.cmpbup.2024.100146.

[8] https://tagvault.org/blog/types-of-bone-fractures/

[9] K. C. Santos, C. A. Fernandes and J. R. Costa, "Feasibility of Bone Fracture Detection Using Microwave Imaging," in IEEE Open Journal of Antennas and Propagation, vol. 3, pp. 836-847, 2022, doi:10.1109/OJAP.2022.3194217.

[10] https://doi.org/10.48550/arXiv.2406.15958

[11] Meena T, Roy S. Bone Fracture Detection Using Deep Supervised Learning from Radiological Images: A Paradigm Shift. Diagnostics (Basel). 2022 Oct 7;12(10):2420. doi: 10.3390/diagnostics12102420. PMID: 36292109; PMCID: PMC9600559.

[12] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788

[13] Wang, Chien-Yao, Hong-Yuan Mark Liao, Yueh-Hua Wu, Ping-Yang Chen, Jun-Wei Hsieh, and I-Hau Yeh. "CSPNet: A new backbone that can enhance learning capability of CNN." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp. 390-391. 2020.

[14]YOLOv8Architecture.https://blog.roboflow.com/content/images/size/w1000/2023/01/image-16.png

[15] He, J.C., Leow, W.K., Howe, T.S. (2007). Hierarchical Classifiers for Detection of Fractures in X-Ray Images. In: Kropatsch, W.G., Kampel, M., Hanbury, A. (eds) Computer Analysis of Images and Patterns. CAIP 2007. Lecture Notes in Computer Science, vol 4673. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-74272-2_119

[16] https://viso.ai/deep-learning/yolov8-guide/