# Automated localization and classification of bone fractures in radiology

Mohamed Afsal

BSc Computer Science specialization in Cyber Security

School of Computer Science And Application

8186afsal@gmail.com

Mohammed Ali Athar

suhaib.athar@gmail.com

# INTRODUCTION

A Radiologist plays a critical role in patient care. Their ability to rapidly assess and detect issues such as brain haemorrhage, pulmonary embolism, cancer, and fracture is all that lies between a patient's full recovery or death and disability. Through their swift detection, they can minimize further patient care and reduce the burden on the healthcare system. In addition to that, there is an alarming global shortage of radiologists currently and this shortage is reportedly projected to grow further. The introduction of artificial intelligence (AI) to radiology has formed a sense of urgency among radiologists to learn about this new technology and its applications. Works show that medical specialists have recognized the major transformation that can happen to radiology.

Why is there a shortage ?

"The demand for imaging is outpacing what we're doing on the training side," said Dr. Yaghmai, professor and chair of radiological sciences at the University of California, Irvine. "The number of radiologists in the workforce is not growing as fast as the population and the demand for imaging." According to WHO (World Health Organization) the population of older ages is increasing and said to be 22% by 2050 of the global population compared to 11% on 2015. This means that the older population requires more imaging. Europe has 13 radiologists per 100,000 population but in the U.K., the rate is only 8.5 per 100,000 , This is an very low rate.

How its affecting the medical world ?

Newshub , a New Zealand news service reports "**Staff shortages at radiology departments across NZ could cost lives**" , This shortage of Radiologists is causing backlog of patients waiting for CT scans is growing quickly , The effect of the shortage is not only limited to New Zealand but rather its effects are being felt across the entire world , so this problem needs to be addressed before it causes more harm

How will it solve real-world problems?

The famous Quote "Time is priceless" is a good one but it stands out especially when Someone's life is on the line, Emergency X-ray scans often require the Radiologist's Quick attention where a scan has to be done and reported within a few minutes to hours but this also means that Radiologist won't be present for the Normal level scans which are not of a higher priority thus they take 1 to 2 days to produce results. This proposed idea will detect the fractures of both normal scans and High priority (Emergency) scans and produce the result which will be later determined by the Radiologist, this not only reduces the time taken to produce results but also give Radiologist more time to focus on emergency scans which is time vital as the patient's life is on the line, this idea will highly increase the efficiency of the production of reports

With all these facts stated, Artificial Intelligence (AI) will become more mainstream in clinical care over the next few years, and it will become an essential part of the diagnostic care process. This proposed idea where An AI based system is built to detect bone fractures from an X-ray machine scan will solve a lot of problems in the medical world. A substantial fraction of radiology data are still collected manually and often in an unstructured format, including patient demographics and radiology reports.

# METHODS

## Datasets

The dataset used for the study is assembled from the website of Stanford ML group. Musculoskeletal Radiographs (MURA) is released by Stanford ML group. This dataset contains X-ray images of different bones including elbow Xrays, hand X-rays, shoulder X-rays and finger X-rays, etc. 250 elbow X-rays are used for this study from the dataset including 190 normal and 60 abnormal X-rays.

Another dataset on Kaggle contains mixture of Chest, Pelvic And C-spine Fractures of a total of 530 X-ray Images it can be found at url:

https://www.kaggle.com/datasets/pardonndlovu/chestpelviscspinescans

## Algorithms and Models for Object Dectection

### Traditional Methods

The traditional approach of object detection usually has three stages:

1) informative region selection
2) feature extraction
3) classification of the object

1) In the first stage, we try to find the object's location. Objects have a different size and aspect ratio and may appear at different locations of an image. Due to this, we have to scan the whole image using a multiscale sliding window. But this method is computationally expensive and it produces many irrelevant candidates.
2) In the second step, we do the feature extraction stage by using techniques like SIFT, HOG to extract the visual feature for recognizing the object. These visual features provide a semantic and robust representation. However, due to different illuminations conditions, viewpoint differences, and different backgrounds it is very difficult to manually design a robust feature descriptor to perfectly describe all types of objects.
3) In the third classification stage we use Support Vector Machine(SVM) or Adaboost for the classification of target objects from all the other categories and to make the representations more hierarchical, semantic, and informative for visual recognition.

The problem with the traditional approach is that the generation of candidate bounding boxes using the sliding window technique is computationally expensive and also the hand-engineered features are not always sufficient to perfectly describe all types of objects. The feature extraction was done through various mathematical techniques such as HOG(histogram of oriented gradients) or SIFT (Scale Invariant Feature Transforms) and were not learnt by a machine learning algorithm therefore were not that robust.

## Main Algorithm - Convolutional Neural Network based Object Detection

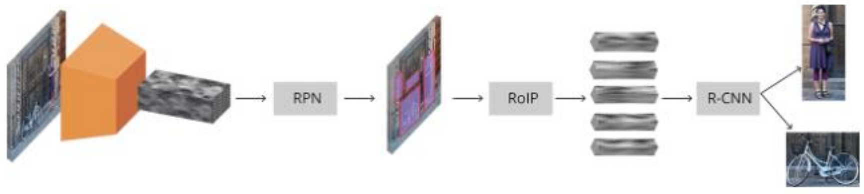
Modern techniques have been brought forward by the emergence and popularity of deep convolutional neural networks. Using these methods not only brings improvements to the computational complexity but also improves the performance of the models and the inference times. CNN based object detection methods utilize convolutional layers and conv-nets along with techniques such as regression to solve the problem of object detection, rather than utilizing traditional methods of feature extraction, CNNs can be trained to automatically extract useful features from an image, moreover, with the prevalence of pretrained networks, CNNs only need to be fine tuned to fit to a specific task. Even object detection techniques using CNNs are of two types; 1) dual-stage detectors, 2) single-stage detectors.

## Sub-Algorithm 1 - Dual-Stage Object Detectors

As the name suggests, dual-stage object detection models consist of two stages in the object detection pipeline; 1) region extraction, 2)region classification. Region extraction is the process of extracting viable regions from a given image where an object exists, rather than using something like a sliding window technique (which returns all possible locations with an image) this process is carried out using more efficient techniques such as “selective search” or more modern techniques such as a region proposal network, or feature pyramids.

## Model 1 for Dual-Stage : Faster R-CNN

This model is the third in line of a serious of region based CNN methods. This model consists of two parts; 1) region extraction, 2) predicting the label and probability of each region. The first R-CNN network used the traditional technique of selective seach for region extraction, while this was superior to sliding window techniques, it was still slow since selective search was a CPU intensive process and still required the model to train on ~2000 regions per image. Faster R-CNN improved this deficit by introducing the region proposal network. The region proposal network takes the features extracted from the CNN backbone and outputs a user-defined number of anchors with different scales. These anchors reference the original image. The R-CNN prediction module then predicts the probability of an actual object being present for an anchor, and the bounding box coordinates tied to that anchor. If a bounding box is predicted to be of a background class, it is discarded. The bounding boxes containing actual objects are returned along with their probabilities. Furthermore, a bounding box regressor is used to ensure the fit of the bounding box is more accurate.

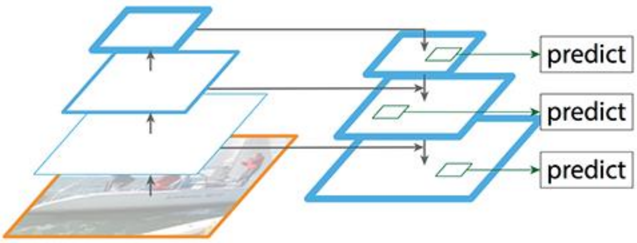



Complete Faster R-CNN architecture

Faster R-CNN is one of the state of the models in the object detection task.

### Model 2 for Dual-Stage : FPN (Feature Pyramid Network)

FPN (feature pyramid network) is a model used for the extraction of features from an image, rather than simply utilizing a convolution backbone. The feature pyramid network provides the advantage of extracting features from different scales, this allows small objects present in images to also be captured. The features extracted using the feature pyramid are present as multiple feature map layers. These layers are then passed onto a region proposal network and bounding box predictor to obtain the object detection prediction.




Add skip connections (Source)

### Sub Algorithm 2 – Single-Stage Detectors

The Dual-stage detectors provide accurate results due to the use of a region proposal method which outputs all likely regions within an image, despite the accuracy they lack in speed. Since the regions have to be produce first, then the prediction, this makes these networks unsuitable for real-time environments. A one-stage detector, on the other hand, requires only a single pass through the neural network and predicts all the bounding boxes in one go. Single stage detectors utilize bounding box regression and non-maxima suppression to detect objects in images in one pass. These models are trained end-end.

## Model – YOLOv4

YOLO is a one-stage detector. The One-stage method is one of the two main state-of-the-art methods used for the task of Object Detection, which prioritizes on the inference speeds. In one-stage detector models ROI (Region of Interest) is not selected, the classes and the bounding boxes for the complete image is predicted. Thus, this makes them faster than two-stage detectors. It divides the object-detection task into regression task followed by a classification task. Regression predicts classes and bounding boxes for the whole image in single run and helps to identify the object position. Classification determines the object's class. YOLOv4 is a **SOTA (state-of-the-art)** real-time **Object Detection** model. It was published in April 2020 by Alexey Bochkovsky; it is the 4th installment to YOLO. It achieved SOTA performance on the COCO dataset which consists of 80 different object classes.

The architecture consists of various parts, broadly they are - The input which comes first and it is basically what we've as our set of training images which will be fed to the network - they are processed in batches in parallel by the GPU. Next are the Backbone and the Neck which do the feature extraction and aggregation. The Detection Neck and Detection Head together can be called as the Object Detector.

## RetinaNet

**RetinaNet** is a one-stage object detection model that utilizes a focal loss function to address class imbalance during training. Focal loss applies a modulating term to the cross entropy loss in order to focus learning on hard negative examples. RetinaNet is a single, unified network composed of a *backbone* network and two task-specific *subnetworks*. The backbone is responsible for computing a convolutional feature map over an entire input image and is an off-the-self convolutional network. The first subnet performs convolutional object classification on the backbone's output; the second subnet performs convolutional bounding box regression. The two subnetworks feature a simple design that the authors propose specifically for one-stage, dense detection.

DESIGN / PROTOTYPE

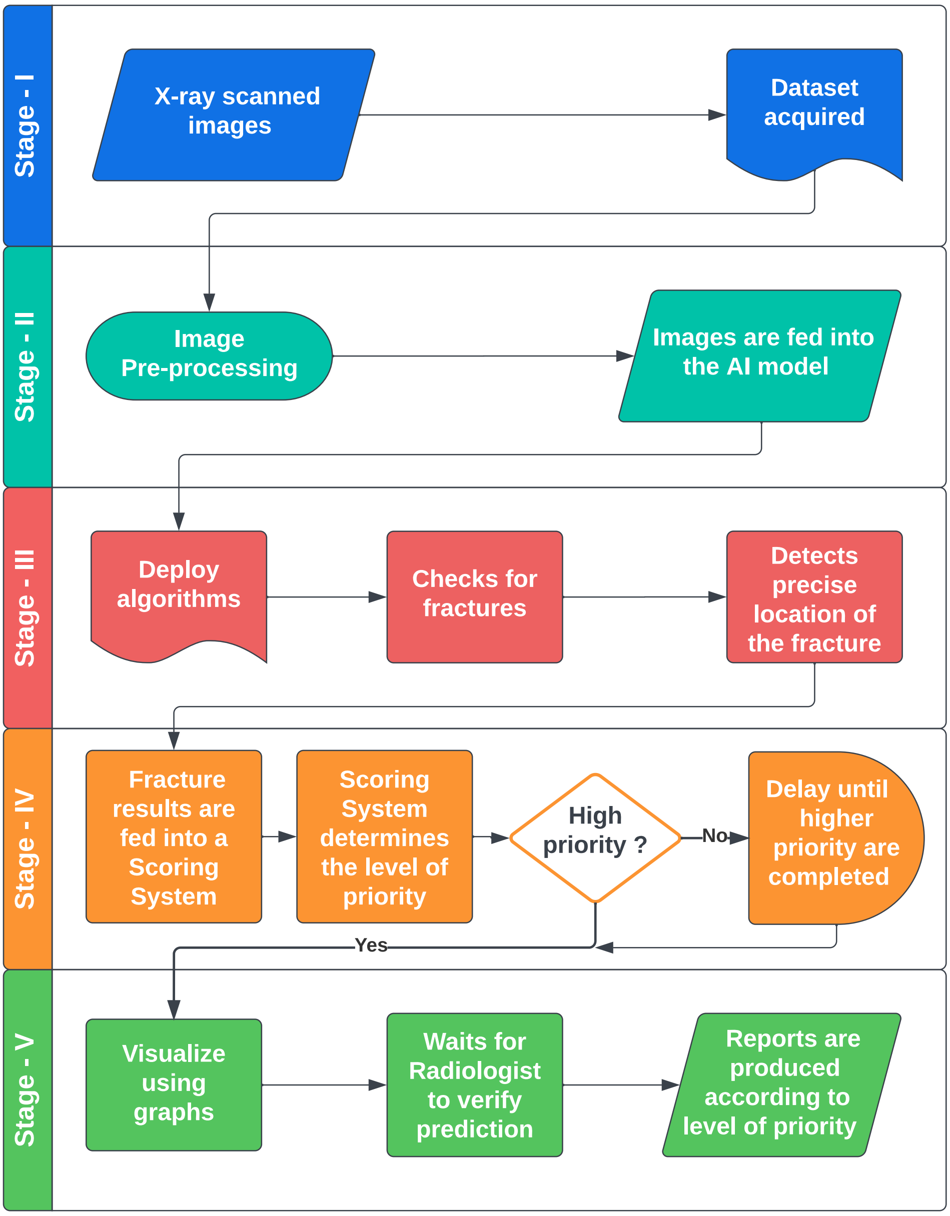



Figure 1 : Block Diagram for Model design

The above Diagram "Figure 1" gives An overview of the process flow of the AI Model

**Stage – I**

This is the first stage CT-Scanners are used to produce X-ray scanned images which will act as an Input to the AI model <Data set accquired>

**Stage – II**

In this stage Image / Data pre-processing takes place to make the images optimal for the AI model , then these pre-processed images are fed into the AI model / Training AI model

**Stage – III**

In the 3rd stage methods / algorithms are deployed which will aid the AI model to check for fractures detecting its precise location using bounding boxes

**Stage – IV**

The results of these fractures are fed into a scoring system , the scoring system is designed to determine the level of priority of the fracture according to its severity level on a priority scale of 1 to 10 , 1 being the most critical or prioritized score. The model will conditionally check if other X-ray scans of Higher priority than the current Scan is present , if there is no higher priority scan present than this current Scan , then the current Scan proceeds to the next stage , If other X-ray scans of Higher priority than the current Scan is present , then the current scan priority scan proceeds to the next stage

**Stage – V**

In the Final stage the AI model produces graphical representation of its result , giving its detection of precise location of the fracture and accuracy level of the detected fracture. These results are verified and compared by the Radiologists and then approved by the Radiologists , then the X-ray reports are produced

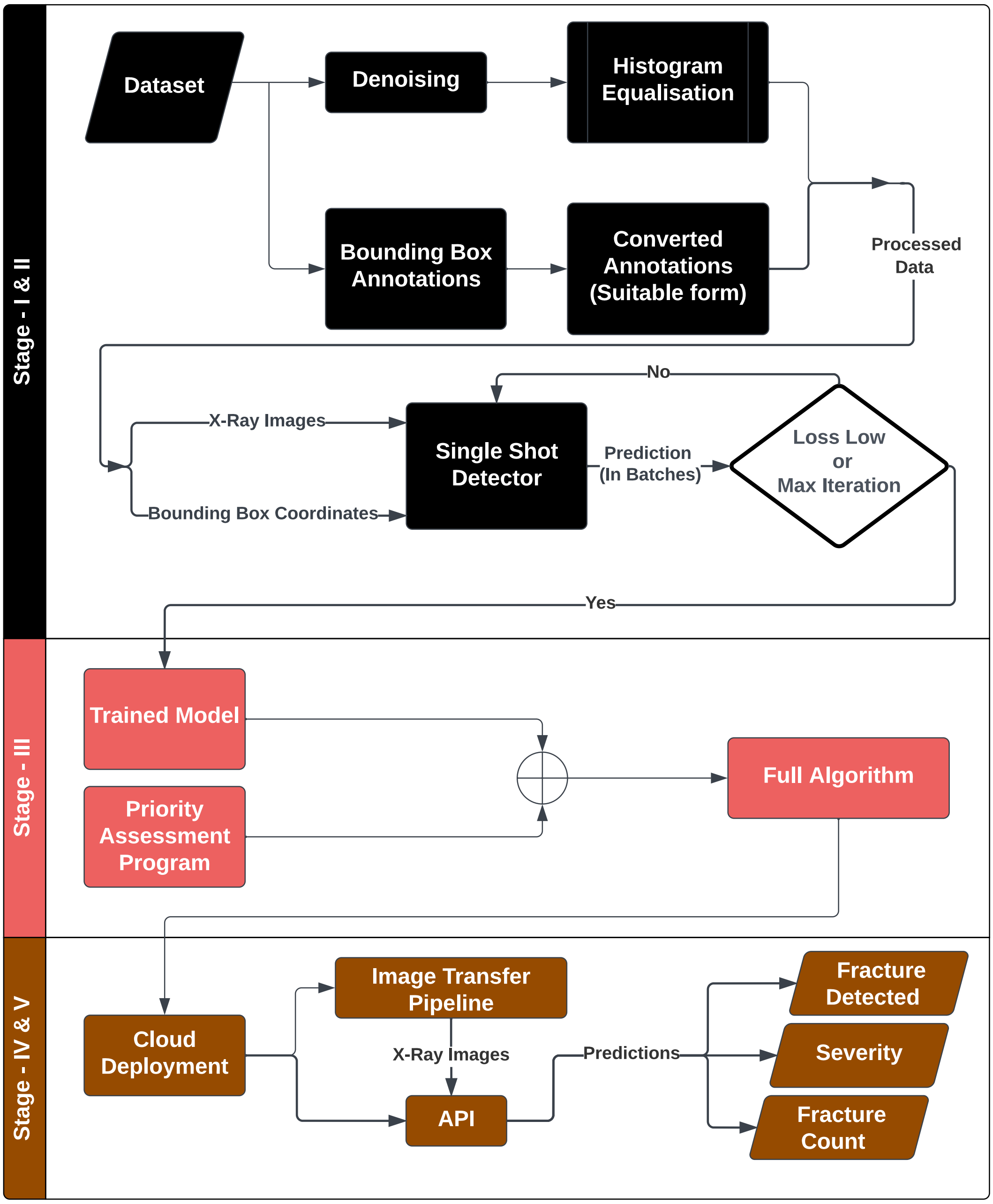



Figure 2 : Proposed Architecture of the Model

The above Diagram “Figure 2” gives a deeper look of the architecture of the “Figure 1” Prototype / Design

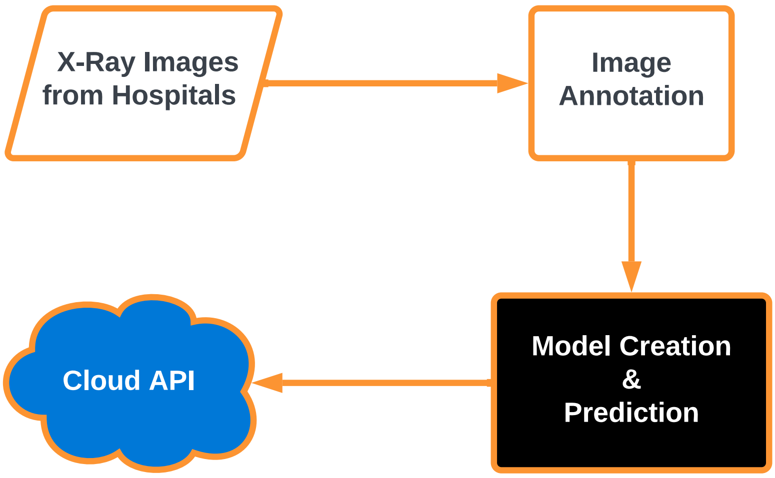



Figure 3 : Summary Architecture of the Model

The above Diagram “Figure 3” is a summarized version of the “Figure 2” architecture of the Model

In “Figure 3” , First X-Ray images are sourced from hospital or publicly sourced from the internet and the X-rays which have fractures are required to be marked by bounding boxes which is annotated by experts , then comes model creating and training : “The Model Creation & Prediction” Box of “Figure 3” has been expanded in detail in “Figure 2”, Once the trained model is ready then its uploaded to the cloud API , the API will contain the methods / functions which will allow predictions to be obtained when they are sent through the API

Clouds such as Google Cloud Platform (GCP) can be used where the Object detection models are going to be created using Tensor flow frameworks which are provided by Google Cloud Platform (GCP) , It helps to deploy models more efficiently and easily due to Tensor flow integration in the Google Cloud Platform (GCP)

# RESULTS AND DISCUSSION

The results produced by the AI model helped us to understand which model training methods to use in terms of their efficiency and accuracy , YOLO algorithm seems to be best fit for model training for this project

**Datasets are split into 70% for training, 15% for validation and 15% for testing**

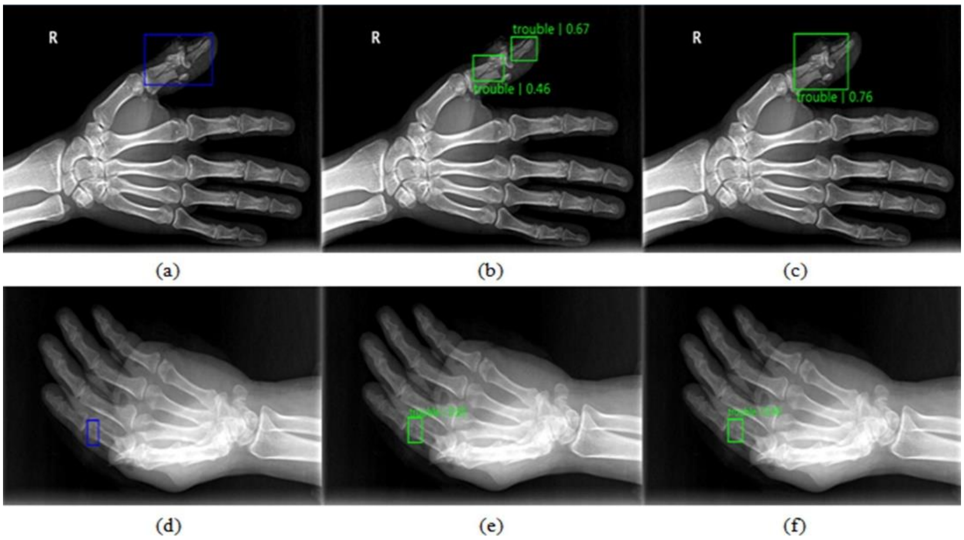


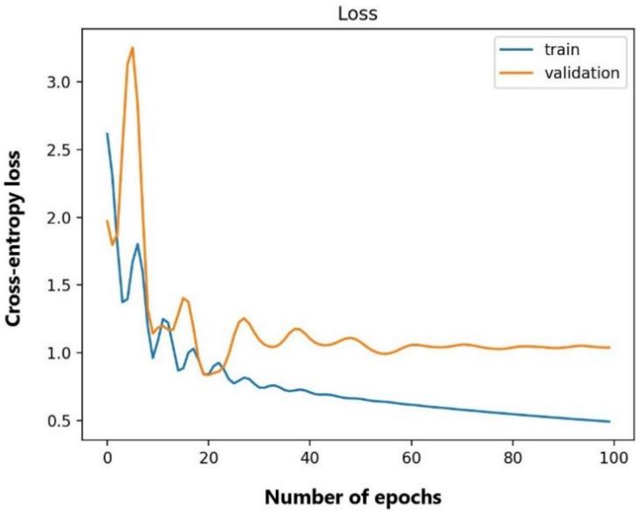

Figure 4 : Close-up views of examples of hand fracture detection results

The above "Figure 4" shows No fractures in BLUE bounding boxes and it also shows Fraction probability in GREEN bounding boxes

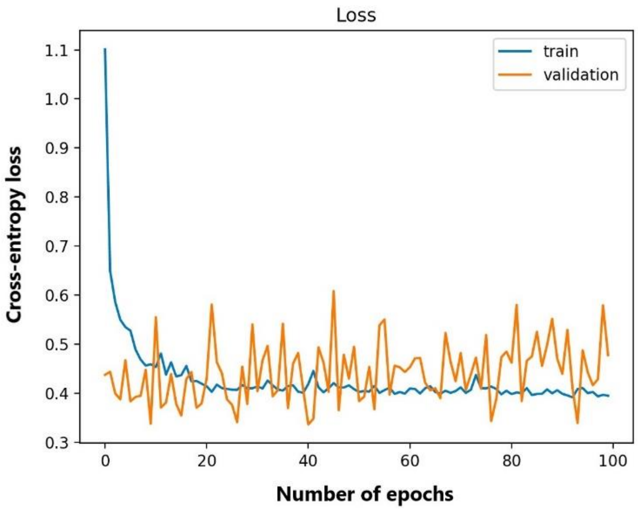
| Fractures | Models | | | |
|---|---|---|---|---|
| | RetinaNet | YoloV4 | Faster R-CNN | FPN R-CNN |
| Finger | 0.8634 | 0.9167 | 0.9027 | 0.901 |
| Hand | 0.8725 | 0.9341 | 0.8929 | 0.8934 |
| Shoulder | 0.8432 | 0.9001 | 0.9102 | 0.9023 |
| Elbow | 0.8345 | 0.8643 | 0.8592 | 0.8447 |
| Pelvis | 0.8238 | 0.8592 | 0.8513 | 0.8492 |
| Chest | 0.8323 | 0.8542 | 0.8482 | 0.8312 |

This is a Confidence score table for the different fractures and different models and each value is the mean confidence score , YOLO has consistently better performance than the other models

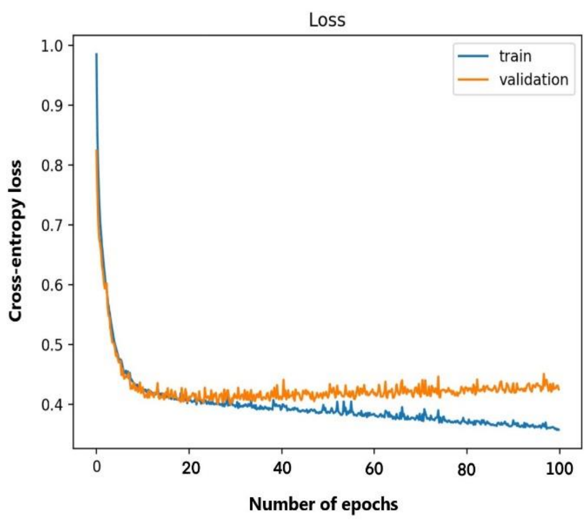
**The graphs below are training-validation loss for model training**

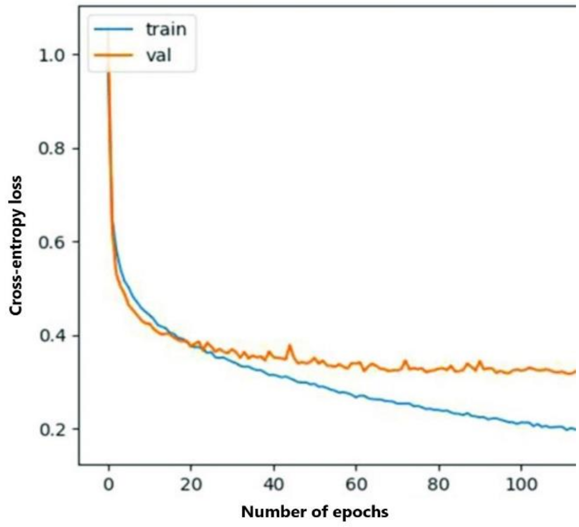



RetinaNet : This is the loss progression for the training and validation sets on the RetinaNet model over 100 epochs. . At approximately 20 epochs, the model gives the lowest loss for both the validation and training dataset, after 20 epochs the model can be seen to overfit.



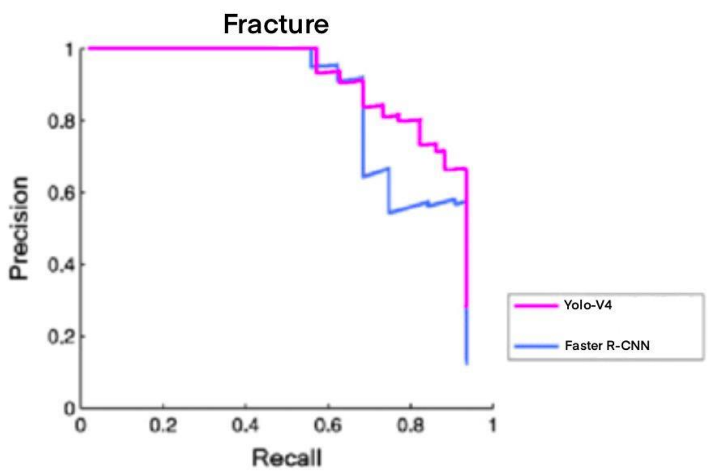

FPN : This is the loss progression for the training and validation sets on the FPN model over 100 epochs. The progression of the loss during training is quite volatile compared to the other models.




YOLO (On the left) : This is the loss progression for the training and validation sets on the YOLO model over 100 epochs. The progression of the loss during training is stable for the yolo model. At approximately 10 epochs, the model gives the lowest loss for both the validation and training dataset, after 20 epochs the model can be seen to overfit.

Faster R-CNN (On the right ) : This is the loss progression for the training and validation sets on the Faster R-CNN model over 100 epochs. The progression of the loss during training is stable for this model. At approximately 20 epochs, the model gives the lowest loss for both the validation and training dataset, after 20 epochs the model can be seen to overfit.




This is a precision-recall curve for two models, the yolo and the faster R-CNN. YOLO is outperforming the faster R-CNN as it's able to keep a better balance between precision and recall

# CONCLUSION

Provide possible outcomes for a new idea This idea has not yet been implemented but has been under development for years, this idea could help high-end hospitals to clinics, and Radiology centres, this model will work as a second pair of hands for the Radiologist. This idea will aid the Radiologist by decreasing the time taken to produce reports from days to minutes thus increasing the overall efficiency of the process of producing Radiology reports. Additionally, Radiologists will have better Priority Management so that they can devote their time and resources to higher priority (emergency) cases.

The AI model further helps in detecting and organizing the scanned X-rays according to their priority so the Radiologists don't need to check for the priority level of the scans themselves and once the AI gives the scanned X-ray results according to their severity, the Radiologist can produce reports for the higher priority cases first so the patients of higher risk are treated first, this further improves efficiency thus reducing the time taken to produce scan reports.

Currently there is no huge publicly available datasets present on X-ray scans of bone fractures as its private but if the AI model is trained with huge datasets In the future , it is bound to produce more accurate prediction.

Further improvements to this AI model can be made by adding AI aided report generation where instead of the Radiologists producing the report after results obtained from the AI model , the Model itself will input the X-ray Images and its findings into a X-ray report and the Radiologists can fill in the rest of the information , this will further save time in the productions of reports.

# REFERENCES


1. **Jha, S.** (2020). Artificial Intelligence in Radiology: The Future is Here. *American Journal of Roentgenology*, 215(1), 145-153. https://doi.org/10.2214/AJR.20.23282
2. **Kirk, S. H., & Flemming, D. J.** (2019). Understanding the Role of Artificial Intelligence in Radiology. *Radiology Management*, 41(2), 91-96.
3. **Yaghmai, V., & Heller, S.** (2019). A Growing Crisis: The Radiology Workforce. *American Journal of Roentgenology*, 213(1), 1-2. https://doi.org/10.2214/AJR.19.20956
4. **Lakhani, P., & Thoden, J.** (2018). Deep Learning for Radiology: A Review. *The British Journal of Radiology*, 91(1087), 20180015. https://doi.org/10.1259/bjr.20180015
5. **Khalilzadeh, O., et al.** (2020). Artificial Intelligence in Radiology: Current Applications and Future Directions. *Radiology: Artificial Intelligence*, 2(4), e200015. https://doi.org/10.1148/ryai.2020200015
6. **Rajpurkar, P., et al.** (2017). CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. *Proceedings of the National Academy of Sciences*, 115(46), 11591-11596. https://doi.org/10.1073/pnas.1711239115
7. **Bai, W., et al.** (2020). Automated Assessment of Bone Fractures using Deep Learning. *Nature Biomedical Engineering*, 4(6), 1120-1130. https://doi.org/10.1038/s41551-020-0572-5
8. **Li, L., et al.** (2021). A Comprehensive Review on Deep Learning in Medical Imaging: Applications and Challenges. *Journal of Medical Imaging*, 8(3), 030901. https://doi.org/10.1117/1.JMI.8.3.030901
9. **Chen, M., et al.** (2019). Deep Learning for Medical Image Analysis: A Comprehensive Overview. *Medical Image Analysis*, 58, 101547. https://doi.org/10.1016/j.media.2019.101547
10. **Chung, Y. C., et al.** (2018). A Review of Convolutional Neural Networks for Medical Image Processing. *Journal of Healthcare Engineering*, 2018, 1-14. https://doi.org/10.1155/2018/1026853