RCS bulletin | annals

Developing an artificial intelligence diagnostic tool for paediatric distal radius fractures, a proof of concept study

Journal:	*Annals Journal & Bulletin Journal*
Manuscript ID	RCSJ-2022-0162.R1
Manuscript Type:	Original research – Annals (clinical)
Date Submitted by the Author:	n/a
Complete List of Authors:	Aryasomayajula, Sriharsha; Kingston University, Faculty of Science, Engineering and Computing Hing, Caroline; St George's University Hospitals NHS Foundation Trust, Trauma and Orthopaedics Siebachmeyer, M; St George's University Hospitals NHS Foundation Trust Naeini, FB; Kingston University Ejindu, V; St George's University Hospitals NHS Foundation Trust Leitch, P; St George's Hospital Medical School Gelfer, Y; St George's University Hospitals NHS Foundation Trust Zweiri, Y; Kingston University
Keywords – Go to MeSH to find your keywords.:	Wrist Fracture, Artificial intelligence, X-rays, radiographs

Developing an artificial intelligence diagnostic tool for paediatric distal radius fractures, a proof of concept study

S Aryasomayajula, CB Hing, M Siebachmeyer, FB Naeini, V Ejindu, P Leitch, Y Gelfer, Y Zweiri

Sriharsha Aryasomayajula. Faculty of Science, Engineering and Computing, Kingston University, London, United Kingdom
harshaarya17@outlook.com

Prof. Caroline B Hing*. Department of Orthopaedics, St George's University Hospitals NHS Foundation Trust, London, United Kingdom
caroline.hing@stgeorges.nhs.uk

Martin Siebachmeyer. Department of Radiology, St George's University Hospitals NHS Foundation Trust, London, United Kingdom
martin.siebachmeyer@stgeorges.nhs.uk

Dr. Fariborz Baghaei Naeini. Faculty of Science, Engineering and Computing, Kingston University, London, United Kingdom
f.baghaeinaeini@kingston.ac.uk

Vivian Ejindu. Department of Radiology, St George's University Hospitals NHS Foundation Trust, London, United Kingdom
vivian.ejindu@stgeorges.nhs.uk

Patricia Leitch. Medical student, St George's University London, London, United Kingdom
patricia.leitch@nhs.net

Dr. Yael Gelfer. Department of Orthopaedics, St George's University Hospital, London, United Kingdom
yael.gelfer@ stgeorges.nhs.uk

Dr. Yahya Zweiri. Faculty of Science, Engineering and Computing, Kingston University, London, United Kingdom
y.zweiri@kingston.ac.uk

*Corresponding author

Abstract (250 words)

Introduction: In the UK 1 in 50 children will sustain a fractured bone yearly yet studies have shown that 34% of children sustaining an injury do not have a visible fracture on initial radiographs. Wrist fractures are particularly difficult to identify as the growth plate poses diagnostic challenges when interpreting radiographs.

Materials and methods: We developed convolutional neural network (CNN) image recognition software to detect fractures in radiographs of children. A consecutive dataset of 5000 radiographs of the distal radius in children aged less than 19 years from 2014-2019 were used to train the CNN. Additionally transfer learning from a VGG16 CNN pre-trained on non-radiological images was applied to improve generalization of the network and classification of radiographs. Hypermeter tuning techniques were used to compare the model to the radiology reports that accompanied the original images to determine diagnostic test accuracy.

Results: The training set consisted of 2881 radiographs with a fracture and 1571 without a fracture, 548 radiographs were outliers. With additional augmentation the final dataset consisted of 15,498 images. The dataset was randomly split into three subsets, a training dataset (70%), a validation dataset (10%), and a test dataset (20%). After training for 20 epochs, the diagnostic test accuracy was 85%.

Discussion: A CNN model is feasible in diagnosing paediatric wrist fractures. We demonstrated that this application could be utilized as a tool for improving diagnostic accuracy. Future work would involve developing automated treatment pathways for diagnosis, reducing unnecessary hospital visits and allowing staff redeployment to other areas.

Keywords: Wrist Fracture, Artificial intelligence, Convolutional neural network, Image classification, X-rays, radiographs

Introduction

Wrist and distal forearm fractures are one of the most common types of presentations to the emergency department (ED) in children and adolescent patients [1]. The unique properties of the immature skeleton result in specific fracture patterns in children [2,3]. Furthermore, the growth plate (physis) which is still open in children, poses particular challenges when analyzing radiographs. Not only can the physis mimic the appearance of a fracture for clinicians with less experience, but the fracture can involve the physis itself.

Radiographs of acute injuries are provisionally interpreted by frontline medical staff in the emergency department (ED). The radiographs are then formally reviewed by a radiologist who writes a report, which may not be available in real time, after the radiograph was taken. By this time, the patients may have already been discharged from the ED. When there is a significant discrepancy between the initial interpretation of the radiograph by the frontline staff and the radiologist, the patient may have to be recalled to initiate a different management pathway.

Subtle injuries can be missed altogether at the initial presentation. One study reported that 3.1% of all fractures were not diagnosed at the initial visit to the ED [4]. The same study also observed a diurnal distribution of errors with more errors occurring during night-time, potentially due to less senior support during night shifts. The specific challenge and factors of missing fractures on radiographs of children have been well documented in the literature [1,2,3]. To overcome these challenges, studies have investigated the feasibility of artificial intelligence (AI) based, automated analysis of radiographs in supporting frontline clinicians in fracture detection. To date, most fracture detection models have been developed for the distal radius [5-9]. AI support tools can significantly improve diagnostic accuracy of clinicians' ability to correctly detect fractures [7].

Most studies published in the recent literature have focused on radiographs of adult patients. Specific paediatric AI models have been developed for the distal tibia and the elbow, but not for the wrist [10,11]. Similar to humans, AI models seem to struggle more with paediatric images. In one study of distal radius fractures, a subgroup analysis found the accuracy of correctly identifying fractures was less with paediatric radiographs compared to adult radiographs with a sensitivity of 92.7% versus 97.5% [12]. A potential explanation is that their AI algorithm was not trained on a dedicated paediatric dataset.

Our study aimed to train an AI model to detect fractures in paediatric wrist radiographs. It is the first study of its kind using a specifically prepared dataset with paediatric wrist radiographs only. Further development of the AI algorithm will enhance radiographic image interpretation in clinical practice, and in particular to support front-line clinicians.

Materials and methods

The objective of this study was to train and optimize an AI-model to detect wrist fractures in paediatric radiographs and to test its accuracy. We trained a Convolutional Neural Network (CNN) using retrospective data from paediatric wrist radiographs and existing text-based reports to categorize the image as 'fracture' or 'no fracture'. Ethics approval was granted by the Health Research Authority (REC 20/PR/0211). The TRIPOD statement checklist was applied for reporting the development of the CNN [13,14].

Data source: A dataset of 5000 retrospectively collected paediatric wrist radiographs was obtained from St George's University Hospitals London. Included were all consecutive wrist radiographs from 2014 to 2018 of patients less than 19 years old until a sufficient quota of 5000 studies was reached. Images of poor quality that were not amenable for interpretation were excluded. Using the radiology text reports that had been produced by consultant musculoskeletal radiologists, the radiographs were labelled as 'fracture' or 'no fracture'. All radiographs with fractures were subcategorized according to the bones involved ***(Figure 1)***.

See figure 1

Data processing: Data collection and de-identification was performed by the direct clinical care team. Radiographs were stored on the PACS (picture archiving and communication system) server of the hospital in DICOM (digital imaging and communications in medicine) format. DICOM is a container format, which contains the imaging data, as well as text metadata. The metadata comprised personal information about the patient, such as name, date of birth etc. The radiographs were downloaded from PACS on a radiology workstation, using specific software tools that allowed mass downloading (RadiAnt DICOM Viewer, https://www.radiantviewer.com or Conquest DICOM, https://ingenium.home.xs4all.nl/dicom.html).

The downloaded radiographs were securely stored on a hospital computer. There was one file for each radiograph. Any text metadata containing personally identifiable information was removed prior to sharing with Kingston University, where the AI model was trained. Information removed from the DICOM images included the patient's name, age, sex, birth date, hospital identity number, NHS number, ethnic group, occupation, referring physician, and institution name and study date. Text metadata was converted to an xlsx file. The xlsx file consisted of entries with accession ID and the presence of fracture. The DICOM files were then split into fracture or no fracture.

There were a few instances where a corresponding ID number was not found in the csv file with the existing DICOM files. These files were separated and gathered for further revision of model. The radiographs and the radiology reports were anonymised before further processing. Further processing included image file conversion, downscaling, and cropping and data augmentation. Augmentation resulted in a larger labelled dataset, intended to make the AI model more robust to real-life data variability.

The dataset was in the format of DICOM which was in an unsuitable format for input into a CNN. The dataset was therefore converted to high resolution png (portable graphics) format using DICOM viewer software. The dataset consisted of images with outliers (features which largely change the model's predictions). Outliers in our dataset were radiographs with intra-osseous metal rods, heavy bandages and alignment tags on the radiographic images. These outliers were removed and the dataset was refined. There were also 239 images which were joined (two radiographs in one image and three radiographs in one image). These images were split manually using image editing software on a windows operating system.

This dataset was then used to create augmented images by applying horizontal and vertical flip as well as random zoom. The augmentations were applied using Python script with a Keras image data generator function [15]. This function was repeated for all the files in fracture and no fracture image directories.

Model training testing and validation: In order to train these neural networks with a high batch size ***(Table 1)***, a considerable amount of computational power was required. This was provided by Google Colab pro account (California, USA). This account was then configured with weights and biases in a third party application which provided an application programming interface (API) reporting the metrics of classification models trained in all experiments.

See table 1

The AI model was based on a Convolutional Neural Network (CNN). CNN is a type of deep learning architecture which performs convolution operation on the image multiple times to extract features from the image. For a given input image and output, the model learns relevant visual features based on convolving the image into the learned filters. Supervised CNN-based models are usually required to be trained on large amounts of labelled data. To reduce the need for an even larger dataset and make the training faster, transfer learning from VGG16 CNN [16] pre-trained on non-radiological images was applied. Our training dataset was used to fine-tune the last hidden layers of the VGG16.

The validation dataset was used to optimize the model's parameters and to perform interval validations. The trained model's ability to detect the presence of fractures in paediatric wrist radiographs was evaluated on the test datasets, previously unseen by the model.

Training: Models were trained to understand the overview of the model's accuracy and bias. The dataset was cleaned and verified for outliers. The outliers were removed and the model was trained. VGG16 with image net weights were used to train an image classification model with two classes. More augmentations were used to increase diversity of datasets in order to improve model generalization. Augmentation of data assisted the network training process to eliminate learning of irrelevant features and noise.

Statistical analysis: The performance of the model was evaluated based on accuracy of the model on the validation set during the training. The loss function of neural networks was defined as binary cross entropy to update the weights and biases of the model with back propagation. As this was the first study performed on the current dataset, an accuracy metric was selected to evaluate the feasibility of this approach to compare fracture detection in paediatric wrist radiographs by AI against the radiologists.

Results

The final training dataset consisted of 2881 real radiographs with a fracture and 1571 without a fracture, 548 radiographs were outliers **(table 1)**. The training model tended to start over fitting after reaching 70% training accuracy. To reduce over fitting additional images were added to the dataset by augmentation techniques. These techniques ensured the models generalized key features of the dataset instead of memorizing all features. This yielded an accuracy of 65% and also started over fitting at 70% training accuracy. Following data augmentation to produce a further 6457 images with fracture and 4589 with no fracture, a total of 15,498 images were available for analysis **(table 1)** to assist the network training process to eliminate learning of irrelevant features and noise, the model generalization improved and overfitting of the model on the validation set reduced. Consequently, the validation accuracy of the model was greatly improved to 85%.

Deep learning models have a large number of parameters which can lead to overfitting with small datasets. Therefore, it was necessary to split the dataset into (i) Training set: to train the model and update weights; (2) Validation set: to select the best model during the training process; (3) Test set: to evaluate the selected model outcome and report the results. In this study, the dataset was randomly split into three subsets, a training dataset (70%), a validation dataset (10%), and a test dataset (20%).

The model was trained for 20 epochs for each model and training accuracy and validation accuracy of the last model is shown in **figure 2,3**. The model was based on vgg16 and a few more layers added after that. Flatten layer was added just before the final activation layer. The final layer had a dense layer with SoftMax activation function. Softmax was used to avoid a vanishing gradient problem. The model had pretrained weights from Imagenet. Input was fed through model and these weights were recalculated corresponding to the training data and an inference made with these weights for any image passed through input layer of neural network.

See figure 2

The training loss and validation loss displayed in **figure 2,3** respectively. The model was evaluated using accuracy. The accuracy was measured based on unseen data i.e., test data, with accuracy defined as the number of correct predictions with the total number of predictions. With that as a metric of measurement, the test data accuracy was 85%.

See figure 3

The CNN has a feature extractor layer and a classification layer. The classification layer as shown in **figure 4** arranges the final output of convolution block into a column. In the case of the algorithm implemented here it has 25088 neurons that are then multiplied to the weight and summed. This sum is then presented to the activation function. The activation function then draws probability of the class of the image and fires the respective neuron and a decision is omitted. The dropout layer ensures it does not fire all the neurons in the network which reduces the overfitting of the network.

Grad-CAM a third party library was used to visualize important features used to make a prediction. A heat map is generated of all the key features represented in red with decreasing intensity of color to blue as shown in the **figures 5, 6,7**. This gives us a better understanding of what the model abstraction sees while making a prediction to classify as fracture or no fracture where the colour red indicates regions with greater significance and blue with lower significance. The regions include the background and boundaries of the image. If a heat map colour is not superimposed on an image it signifies that those features of the image were not considered in making the decision.

Discussion

Distal radius fractures account for around 25% of fractures in the paediatric population [17]. The incidence is increasing with an ensuing increased burden on ED services for diagnosis and treatment [17]. The cost implications are not insignificant with the cost of treating paediatric forearm fractures quoted as $2 billion per year in the United States (US) [18].

Additionally taking a child to the ED often necessitates time off work for parents for initial diagnosis and further follow up [1,2,3,4]. Improving pathways and automating systems can therefore have a positive impact both in terms of hospital resources and reducing unnecessary outpatient attendances for the child and carers.

Studies using AI utilizing CNNs to detect fractures of the distal radius in adults are well reported in the literature [5,8,9,19]. With accuracy, sensitivity, specificity and Youden Index all showing that CNNs can perform better than a group of radiologists in diagnosing adult distal radius fractures [5,8]. Current commercially available CE-marked applications in paediatric musculoskeletal radiology using AI models have mainly concentrated on bone age, bone health, fractures around the elbow and diagnosing child abuse from inflicted fractures [20].

Few models exist to detect distal radial fractures with a recent systematic review identifying only two studies using AI for the distal radius [21]. Of these, one study by Zhang et al used ultrasound and the other radiographs to detect fractures in the entire paediatric appendicular skeleton (Dupuis et al) [22,23]. The systematic review noted that the studies assessed using ultrasound of the distal radius were subject to bias due to the ultrasound being performed by medical students on a 'convenience sample' of suspected wrist fractures in children [22]. The review noted that studies had strict exclusion criteria (healing bones, certain types of fracture, treatment with cast) reducing the applicability of the models in clinical practice. The diagnostic accuracy rates were 92% test accuracy in ultrasound diagnosis of the distal radius but the study design was weakened by selection bias [22].

In our study a large consecutive dataset was used with augmentation to a final dataset of 15,498 images. The AI model classified images into "fracture" and "non-fracture" groups to assist radiologists for detection of bone fracture. For evaluation of the model, all the reported metrics were based on the test dataset which had not been seen by the AI model in the training process. Therefore, the accuracy of model represents the model performance without over-fitting on the dataset.

For future work, the accuracy of the AI model would be calculated on a per-radiograph basis and a per-study basis. Per-radiograph true-positive determination requires a fracture diagnosis which corresponds to a fracture diagnosis in the radiology report. Per-study true-positive determination requires at least one true-positive for either or both of the radiographs of a typical radiographic examination (anteroposterior and lateral view of the same wrist). The sensitivity, specificity, positive and negative predictive values and area under the ROC curves would be estimated with 95% confidence intervals. The model's false-positive and false-negative predictions would undergo further evaluation by a second reading of the original radiographic images.

As the model is trained on a limited sample of images, further data would need to be collected in a diverse manner to be robust against all different types of 'noise' in the image. For example, radiographic imagery devices may have different 'noise' levels which are not even captured by human eye while the noise can affect the model performance. Even a static electro-magnetic noise in the imaging room can introduce a bias in the dataset.

In addition, the current classifier model provides a single outcome to assist radiologists which can be improve in future by the development of image "*segmentation"* models and "*explainable AI*" to highlight the exact region of fracture. The segmentation method will require radiologists to highlight the fracture region in order to train a new AI-based segmentation model to learn the regional features.

Conclusion

In conclusion the diagnosis of paediatric wrist fractures with CNN is feasible and could help radiologists reduce the time they take to diagnose a child's fracture. The CNN trained here was VGG16 with imagenet weights and augmentations to dataset were applied to reduce the over fitting of the model. The model's accuracy improved significantly after multiple tests and runs to 85% on test data.

Word count 2554

Acknowledgements

This work has been funded by a grant from AO UK

References

1. SM Naranje, RA Erali, WC Warner, JR Sawyer, DM Kelly. Epidemiology of pediatric fractures presenting to emergency departments in the United States. Journal of Pediatric Orthopaedics 2016;36(4):e45-48. Doi:10.1097/BPO:0000000000000595
2. MP George, S Brixby. Frequently missed fractures in pediatric trauma. A pictorial review of plain film radiography. Radiologic Clinics of North America 2019; 57(4):843-855. Doi:10.1016/j.rcl.2019.02.009
3. LS Segal, MW Shrader. Missed fractures in pediatric trauma patients. Acta Orthopedica Belgica 2013. 79(6):608-15
4. P Hallas, T Ellingsen. Errors in fracture diagnosis in the emergency department – characteristics of patients and diurnal variation. BMC Emergency Medicine 2006;6:1:4. Doi:10.1186/1471-227X-6-4
5. K Gan, D Xu, Y Lin, Y Shen, T Zhang, K Hu, K Zhou, M Bi, L Pan, W Wu, Y Liu. Artificial intelligence detection of distal radius fractures: a comparison between convoluted neural networks and professional assessments. Acta Orthopaedica 2019;90:4:394-400. Doi:10.1080/17453674.2019.1600125
6. DH Kim, T MacKinnon. Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks. Clinical Radiology 2018; 73:5:439–445. Doi:10.1016/j.crad.2017.11.015.
7. R Lindsey, A Daluiski, S Chopra et al. Deep neural network improves fracture detection by clinicians. Procedures National Academy of Sciences USA 2018;115:45:11591-6. Doi:10.1073/pnas.1806905115
8. YL Thian, Y Li, P Jagmohan, D Sia, VEY Chan, RT Tan. Convolutional neural networks for automated fracture detection and localization on wrist radiographs. Radiology Artificial Intelligence 2019;1:1:e180001. Doi.org/10.1148/ryai.2019180001
9. C Blüthgen, AS Becker, I Vittoria de Martini, A Meier, K Martini, T Frauenfelder. Detection and localization of distal radius fractures: Deep learning system versus radiologists. European Journal of Radiology 2020;126:108925. Doi:10.1016/j.ejrad.2020.108925
10. ZA Starosolski, H Kan, and AV Annapragada. CNN-based radiographic acute tibial fracture detection in the setting of open growth plates. bioRxiv 2019; 506154. Doi:https://doi.org/10.1101/506154
11. JC Rayan, N Reddy, JH Kan, W Zhang, A Annapragada. Binomial Classification of Pediatric Elbow Fractures Using a Deep Learning Multiview Approach Emulating Radiologist Decision Making. Radiology Artificial Intelligence 2019;1:1:e180015. Doi:https://doi.org/10.1148/ryai.2019180015
12. C-T Cheng, T-Y Ho, T-Y Lee et al. Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. European Radiology 2019;29:10:5469–5477. Doi:10.1007/s00330-019-06167-y
13. GS Collins, KGM Moons. Reporting artificial intelligence prediction models. Doi: https://doi.org/10.1016/S0140-6736(19)30037-6
14. GS Collins, JB Reitsma, DG Altman, KGM Moons. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the Tripod statement. BMC Medicine 2015;13:1. Doi:10.1186/s12916-014-0241-z
15. K. Team. Keras documentation: Image data preprocessing. Keras.io, 2021. [Online]. Available: https://keras.io/api/preprocessing/image/.
16. K Simonyan, A Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. Published as a conference paper at ICLR 2015. https:aexiv.org/abs/1409.1556
17. KW Nellans, E Kowalski, KC Chung. The epidemiology of distal radius fractures. Hand Clinics 2012;28(2):113-125. Doi:10.1016/j.hcl.2012.02.001
18. LM Ryan, SJ Teach, K Searcy, SA Singer, R Wood, JL wright, JM Chamberlain. Epidemiology of pediatric forearm fractures in Washington DC. Journal of Trauma and Acute Care Surgery 2010;69:S200-S205. Doi:10.1097/TA.0b013e3181fle837
19. K Oka, R Shiode, Y Yoshii, H Tanaka, T Iwahashi, T Murase. Artificial intelligence to diagnosis distal radius fracture using biplane X-rays. Journal of orthopaedic surgery and research 2021;16:694. Doi:10.1186/s13018-021-02845-0
20. AC Offiah. Current and emerging artificial intelligence applications for pediatric musculoskeletal radiology. Pediatric Radiology 2021;doi.org/10.1007/s00247-021-05130-8

21. S Shelmerdine, H Liu, OJ Arthurs, NJ Sebire. Artificial intelligence for radiological paediatric fracture assessment: a systematic review. Research Square 2022. Doi.org/10.21203/rs.3.rs-1415235/v1
22. J Zhang, N Boora, S Melendez, A Rakkunedeth Hareendranathan, J Jaremko. Diagnostic accuracy of 3D ultrasound and artificial intelligence for detection of pediatric wrist injuries. Children 2021;8(6): 431 doi:10.3390/children8060431
23. J Olczak, N Fahlberg, A Maki et al. Artificial intelligence for analysing orthopaedic trauma radiographs. Acta Orthopedica 2017;88:581-586. Doi:10.1080/17453674.2017.1344459

List of figure and table legends

Figure 1. Schematic of image labelling, training and validation prior to testing.

Figure 2. Graph of accuracy with respect to epochs showing that training accuracy reached its peak by the end of training whereas validation accuracy peaked at 865% for 20 epochs.

Figure 3. Graph illustrating that training loss decreased with each epoch whereas the validation loss flattened after 10~~increases for each~~ epochs.

Figure 4. Decision making process of the CNN classifier layer illustrating 'flatten', 'activation' and the final dense layer with two neurons. The output of the softmax function is the probability of the class of the image being 'fracture' or 'no fracture'.

Figure 5. Heat map of the features learnt by the model at 'block 2' convolution on lateral view (red greater significance, blue lower significance).

Figure 6. Heat map of the features learnt by model at block2_convolution on AP view (red greater significance, blue lower significance).

Figure 7. Heat map of the features learnt by model at block5_pooling on AP view (red greater significance, blue lower significance).

Table 1. Table of the dataset used to train the model split into train, test and validation set.

List of figures and tables

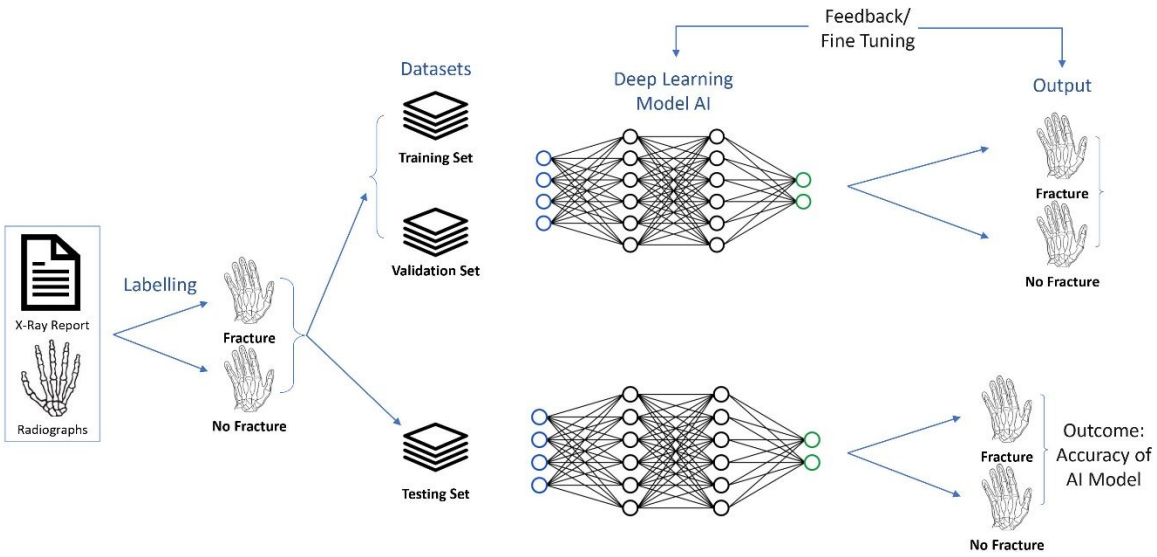

Figure 1. Schematic of image labelling, training and validation prior to testing.

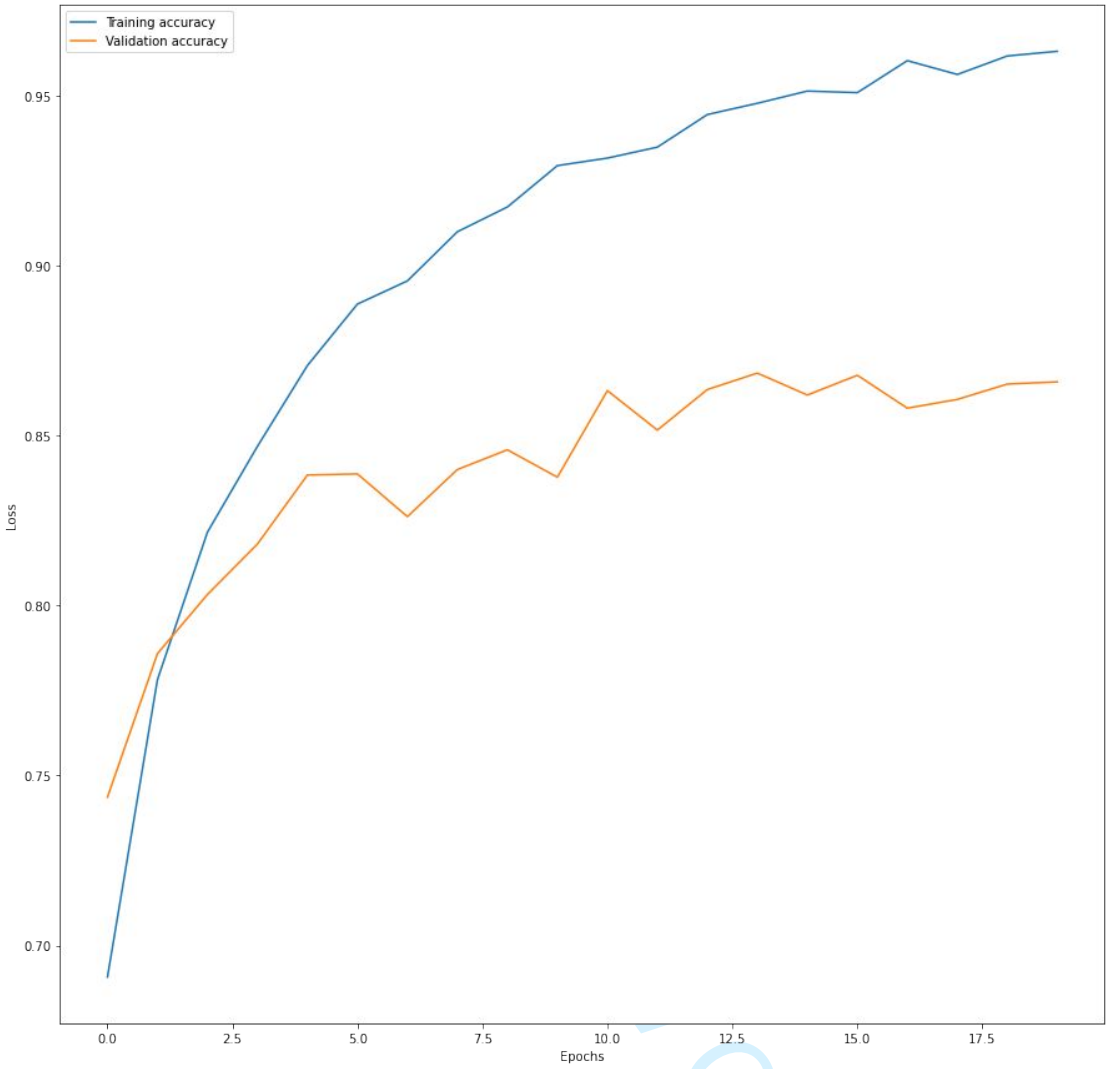

Figure 2. Graph of accuracy with respect to epochs showing that training accuracy reached its peak by the end of training whereas validation accuracy peaked at 86% for 20 epochs.

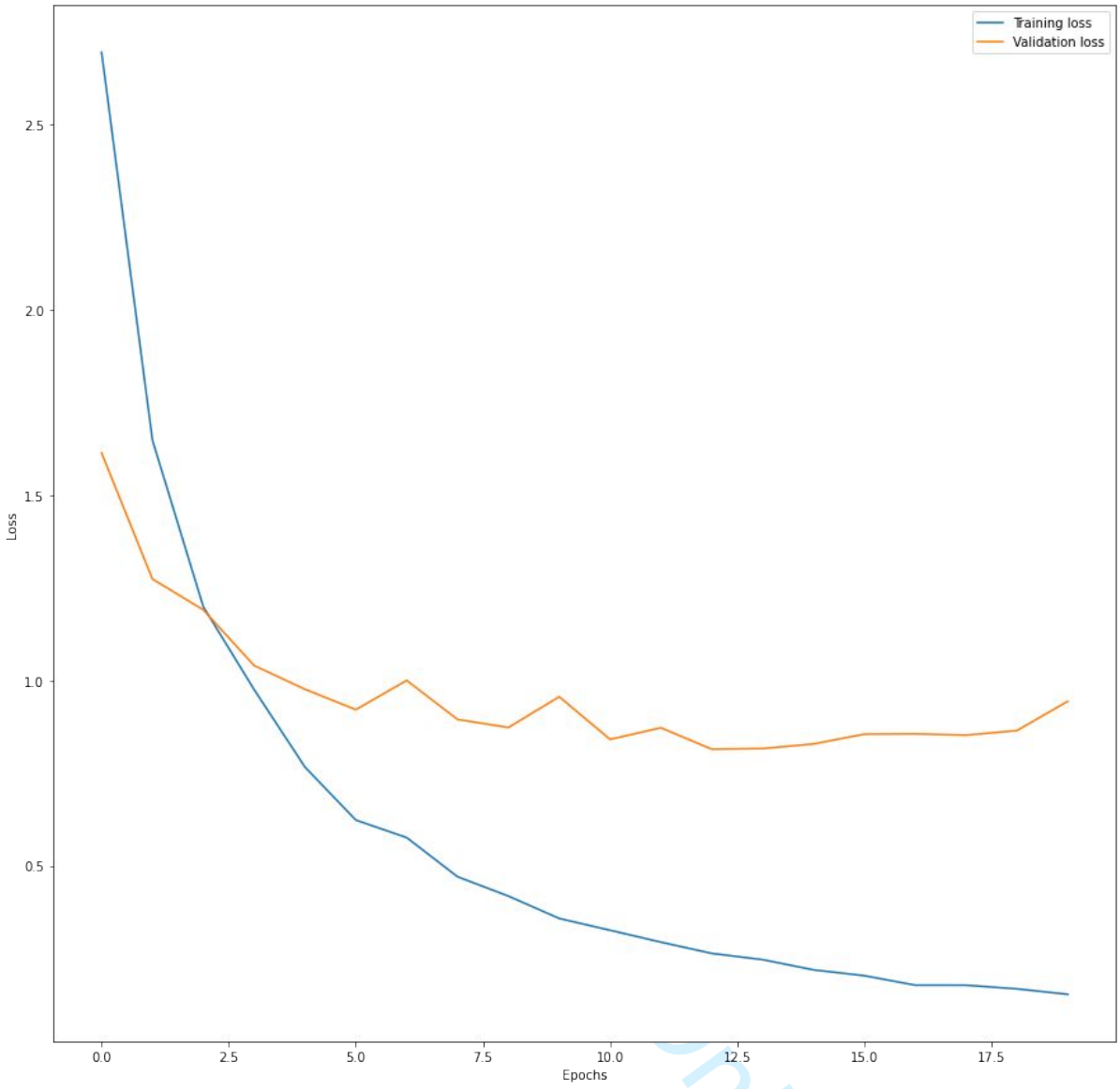

Figure 3. Graph illustrating that training loss decreased with each epoch whereas the validation loss flattened after 10 epochs.

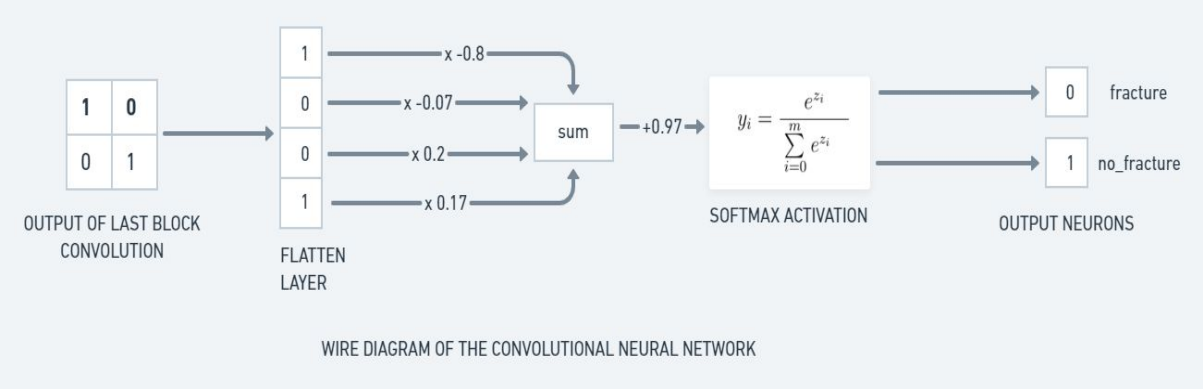

Figure 4. Decision making process of the CNN classifier layer illustrating 'flatten', 'activation'and the final dense layer with two neurons. The output of the softmax function is the probability of the class of the image being 'fracture' or 'no fracture'.

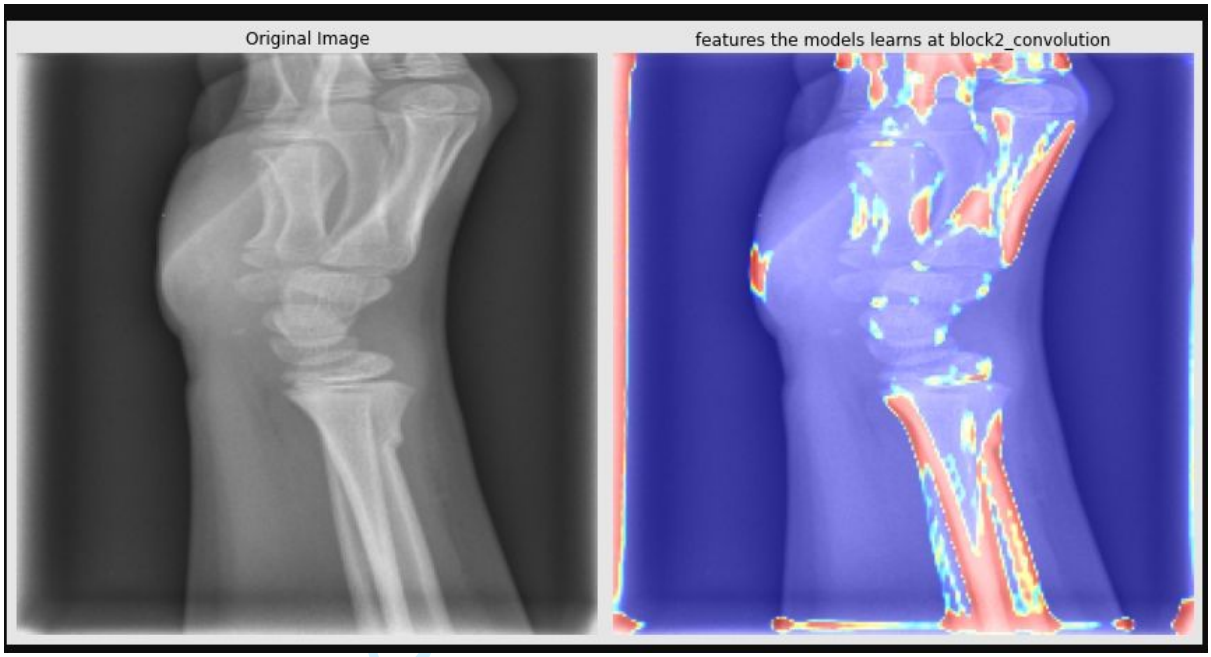

Figure 5. Heat map of the features learnt by the model at 'block 2' convolution on lateral view (red greater significance, blue lower significance).

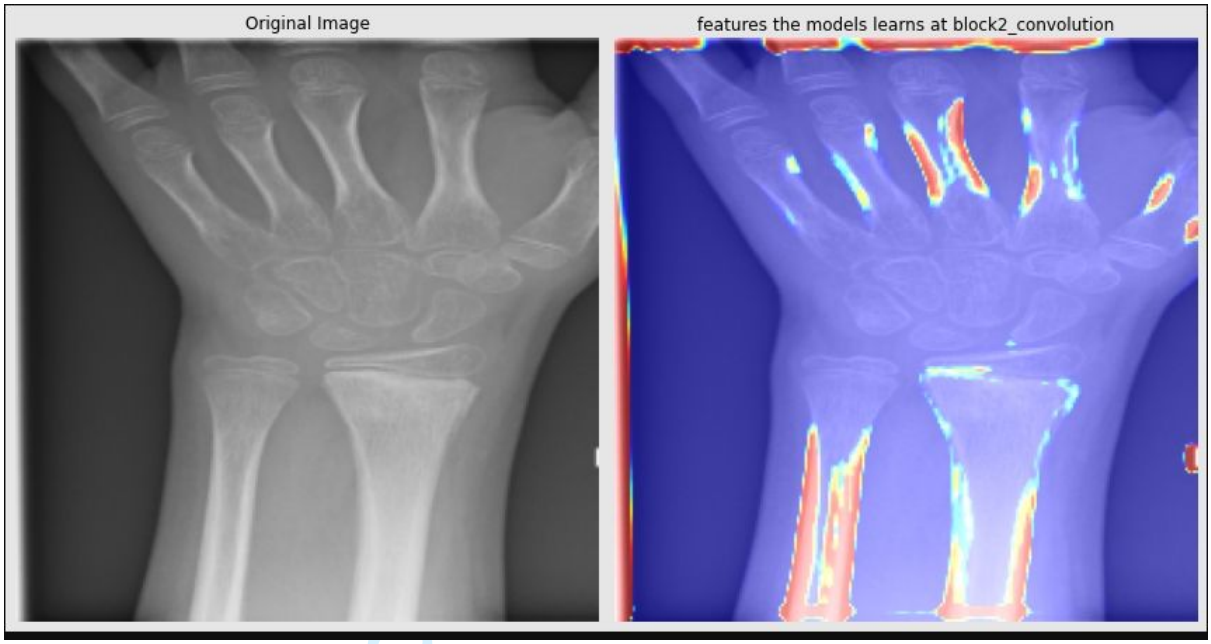

Figure 6. Heat map of the features learnt by model at block2_convolution on AP view (red greater significance, blue lower significance).

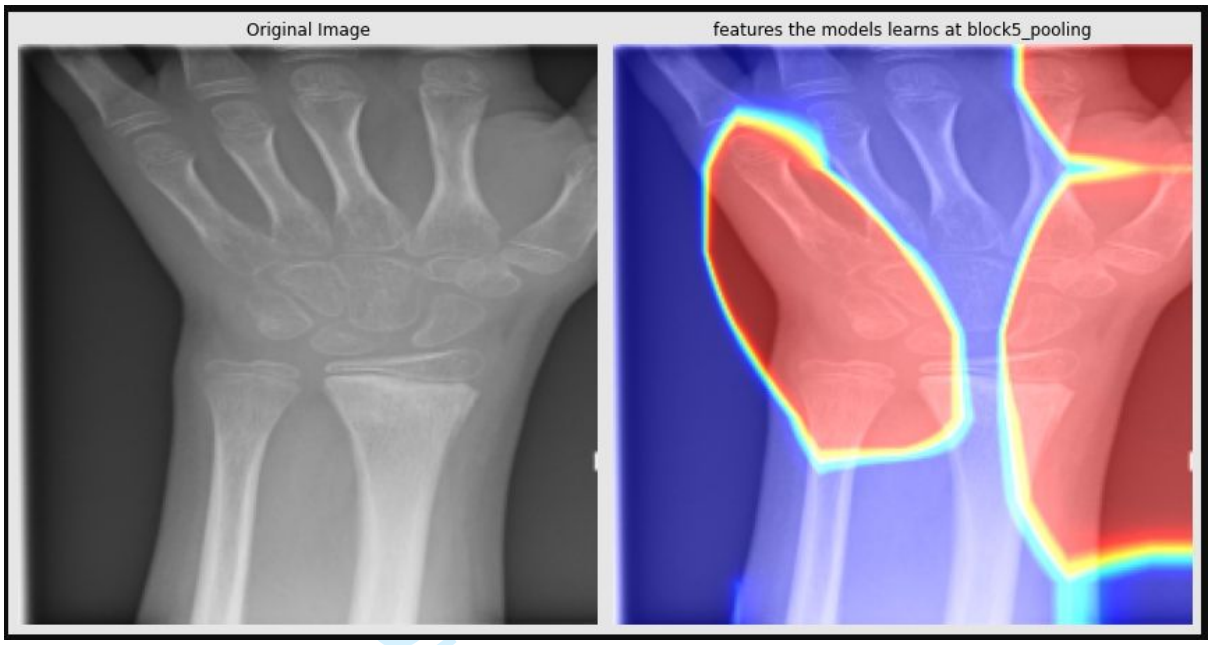

Figure 7. Heat map of the features learnt by model at block5_pooling on AP view (red greater significance, blue lower significance).

TABLES

DATASET SIZE SPECIFICATIONS

	Fracture	No fracture
Augmented	6457	4589
Real	2881	1571
Total	9338	6160

Table 1. Table of the dataset used to train the model split into train, test and validation sets.