



City Research Online

City, University of London Institutional Repository

Citation: Ananda, A. (2023). Wrist Fractures Analysis as observed with X-ray imaging. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/31573/>

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Wrist Fractures Analysis as observed with X-ray imaging



Ananda

Supervisors : Dr Constantino Carlos Reyes-Aldasoro,
Professor Eduardo Alonso and Dr Alex Ter-Sarkisov

This thesis is submitted for the degree of
Doctor of Philosophy

School of Science and Technology

Department of Electrical and Electronic Engineering

October 2023

Declaration

I, Ananda confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. This thesis has never been published or submitted elsewhere for obtaining an academic degree or professional qualification. This thesis is my own work and contains nothing that is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

I grant power of discretion to the Librarian at City, University of London to allow the thesis to be copied in whole or in part without further reference. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgment.

Ananda

Abstract

This thesis studies wrist fractures seen on radiographs. Wrist radiographs are analysed by two different approaches; first by traditional image processing to extract geometric measurements, then by deep learning to classify risks as normal or abnormal (i.e. fractures or implants). Two data sets are used. The first data set includes wrist radiographs obtained from the Department of Radiology at the University of Exeter. The second data set corresponds to MURA X-ray images (MUsculoskeletal RADIographs) obtained by the Stanford Machine Learning Team. The MURA data set provides more X-ray images to explore than the first data set.

In the first task, a semi-automated geometric image analysis algorithm is proposed to analyse and compare the radiographs of healthy controls and patients with wrist fractures treated by Manipulation under Anaesthesia (MuA). The first dataset was used in this task. Thirty-two geometric and texture measurements were created. Image texture emerged as a metric of the most distinct geometric features from wrist X-rays associated with fractures.

In the second task, eleven pre-trained convolutional neural network (CNN) architectures were used. CNN classified the MURA data set into normal and abnormal categories. Transfer learning technique applied to all eleven pre-trained CNNs to deal with wrist X-ray datasets. ResNet-50 and Inception-ResNet-V2 were then explored further using data augmentation strategies. Transfer learning techniques and data augmentation strategies greatly enhance CNN's ability to classify wrist X-ray images.

Class activation mapping (CAM) explores the convolutional neural network's activation associated with the abnormality within the wrist X-ray image. It shows that CAM can indicate the abnormality area in the wrist's X-ray image. The graphical heatmap of CAM overlaid on the wrist X-ray image marks the visual point of the area that triggers the CNN's decision.

Contents

Declaration	iii
Abstract	v
Contents	vii
List of Figures	xi
List of Tables	xxiii
Acknowledgements	xxvii
Abbreviations and Acronyms	xxix
Publications & Version control	xxxii
1 Introduction	1
1.1 Overview	1
1.2 Motivation	5
1.3 Research questions	7
1.4 Aim and Objectives	7
1.5 Thesis structure	7
2 Background	11
2.1 Medical background	11
2.1.1 Human wrist	11
2.1.2 Wrist fractures	12
2.1.3 Manipulation under Anaesthesia	17
2.2 Computational background	20
2.2.1 X-ray image	20

2.2.2	X-ray image of Wrist	22
2.3	Literature review	23
2.3.1	Imaging Technology for wrist fracture problems	23
2.3.2	Medical Image Analysis	25
3	Materials	29
3.1	University of Exeter’s DICOM wrist data set	29
3.2	Stanford Machine Learning Group’s MURA data set	34
4	Methods	47
4.1	General Methodologies	47
4.1.1	Geometric Transformations	48
4.1.2	Edge Detection	50
4.1.3	Contrast Limited Adaptive Histogram Equalisation	58
4.1.4	Local Binary Patterns	64
4.1.5	Convolutional Neural Networks (CNN)	67
4.1.6	X-ray data set augmentation	73
4.1.7	Radiograph Transfer Learning	76
4.2	Class Activation Mapping	77
4.3	Performance evaluation	79
5	Wrist Fractures Semi Automatic Analysis	83
5.1	Introduction	83
5.2	Methods	84
5.3	Results	86
5.4	Discussion	89
5.5	Summary	92
6	Wrist Fractures Classification using Deep Learning	97
6.1	Introduction	97
6.2	Methods	99
6.3	Results	104
6.4	Discussion	110
6.5	Summary	111

7	Class Activation Mapping for Wrist Fractures Image Analysis	113
7.1	Introduction	113
7.2	Materials	114
7.3	Methods	115
7.4	Results	118
7.5	Discussion	120
7.6	Summary	124
8	Discussion and Conclusion	129
8.1	Discussion	129
8.2	Contributions	130
8.3	Conclusions	131
8.4	Future work	132
A	The MIUA 2019 Poster	133
B	Samples of CAM on Wrist X-Ray	135
	Bibliography	139

List of Figures

1.1	Medical imaging modalities examples. (a) X-ray [111], (b) MRI [133], (c) CT [92], (d) Ultrasound [4], (e) PET [72], (f) Nuclear medicine [125]. Each image is credited to its respective references.	2
1.2	An example of the open fractures X-ray images. The image taken from [147].	2
1.3	An example of the closed fractures X-ray images. (a) <i>Postero-Anterior</i> view. (b) <i>Lateral</i> view. The images are credited to [49].	3
1.4	Radiographic Parameters. (a) Radius Height (RH). (b) Volar Tilt. The images are credited to [20].	5
1.5	Graphical outline of the thesis.	8
2.1	The drawing of human wrist anatomy. Image credited to [104]. . . .	12
2.2	The drawing of human wrist anatomy completed with ligaments and bones. This image is taken from [104].	13
2.3	Illustration of three different types of wrist fractures. Colles' fracture (a) PA view [84], (d) Lateral view [84]. Barton's fracture (b) PA view [47], (e) Lateral view [47]. Chauffeur's fracture (c) PA view [123], (f) Lateral view [123]. Posterior-Anterior (PA) view and Lateral (LA) are two common positions the radiologist takes to determine conditions of wrist bone fractured visually.	14
2.4	Operation room set-up for distal radius fracture including instruments for surgery [9]	18
2.5	Situation inside the operation room where manipulation under anaesthesia performed [115]	20
2.6	The X-ray image of Wilhelm Conrad Röntgen wife's left hand. He took this image on 22 December 1895 and believed as the first X-ray image. Image is taken from [15].	21

2.7	X-ray machine illustration where the patient lay down under the X-ray tube. This image is credited to [16].	22
2.8	Basic X-ray imaging system illustration adapted from [13].	22
2.9	The illustration cartoon of (a) Posterior-Anterior (PA) and (b) Lateral (LA) of the wrist was scanned using an X-ray machine. These two images are taken from [38].	23
2.10	A diagram shows a comparison between certain topics of articles stored in the PubMed database from 2000 to 2022.	25
3.1	These two X-ray images are examples of normal wrist X-rays images from The University of Exeter data set. Image (a) is a type of X-rays position named Lateral. Image (b) is the type of X-ray taken in the Posterior-Anterior(PA) position.	30
3.2	Image (a) is a type of X-rays position taken used to identify abnormalities of the bone in the lateral taken position. Image (b) is the type of X-ray taken in the Posterior-Anterior (PA) position. These are wrist X-rays images with fractures identified from The University of Exeter data set.	31
3.3	These are examples of wrist X-rays images with different values of Photometric Interpretation from The University of Exeter data set. Photometric interpretation is one of the standard attribute metadata of DICOM. Image (a) has a monochrome value of 1, where intended to be displayed on white background. Image (b) has a monochrome value of 2, where intended to be displayed in black background [65] .	31
3.4	These are annotated wrist X-ray images with fractures identified from The University of Exeter data set. A green circle is an abnormality annotated by a radiologist.	32
3.5	Examples of Train positive X-ray images from MURA data set [111]. These are images that are positively labelled and have an abnormality in them. These images are part of the images in the Train folder by MURA. The subfigures show train-positive X-ray images of the Elbow (a), Fingers (b), Forearm (c), Hand (d), Humerus (e), Shoulder (f), and Wrist in PA position (g), and Wrist in Lateral position (h). . . .	37

3.6	Examples of valid positive X-ray images from MURA data set [111] labelled by MURA team experts. Example of valid positive X-ray image of Elbow (a), Fingers (b), Forearm (c), Hand (d), Humerus (e), Shoulder (f), Wrist in PA position (g), and Wrist in lateral position (h).	37
3.7	Examples of Train negative X-ray images from MURA data set. These images are labelled for the training process of normal data. The subfigures show train negative X-ray images of Elbow (a), Fingers (b), Forearm (c), Hand (d), Humerus (e), Shoulder (f), Wrist in PA position (g), and Wrist in Lateral position (h).	38
3.8	Examples of valid negative X-ray images from the MURA data set. The MURA's expert justified that these images in the data set are normal. Images in this label will be used as validity for the negative result of data training. The subfigures show valid negative X-ray images of the Elbow (a), Fingers (b), Forearm (c), Hand (d), Humerus (e), Shoulder (f), and Wrist in PA position (g), and Wrist in Lateral position (h).	38
4.1	The illustration of the coordinate convention used to represent a digital image. This image is adapted from [52].	48
4.2	These are examples of applied Rotation of an image for wrist X-rays images. Image (a) is an original wrist X-ray image in a lateral taken position. Image (b) results from 45° rotations from the original position of the image (a).	51
4.3	These are examples of applied image translation of an image for wrist X-rays images. Image (a) is an original wrist X-ray image in a lateral taken position. Image (b) is the result translation vector of $x = 250.3$ and $y = -300.1$ from the original position of image (a).	51
4.4	These are examples of applied image scaling of an image for wrist X-rays images. Image (a) is an original wrist X-ray image in a lateral taken position. Image (b) is 50 per cent scaled down from the original position of the image (a), which could be observed from the value image coordinates.	52

4.5	These are examples of shear operation to an image of wrist X-rays. Image (a) is an original wrist X-ray image in a lateral taken position. Image (b) is the result of the horizontal Shear of the original position of image (a), and image (c) is the result of the vertical Shear of the original image(a).	52
4.6	These are the ideal representations of edges. From left to right, a step-edge model (a), a ramp-edge model(b), and a roof edge model(c). Below each model is corresponding intensity profiles. These illustrations are adapted from [52].	53
4.7	(a) Contrast stretching function (b) Thresholding function. These graphics are adapted from [52]	54
4.8	(a) An example of the wrist X-ray image, (b) Intensity histogram that can be partitioned by T (threshold).	54
4.9	These are applied edge detection techniques examples to an image for wrist X-rays images. Image (a) is an original wrist X-ray image in a lateral taken position. Image (b) is the result of the Canny edge detection method [21]. Image(c)is the result edge detection using Otsu's method [99].	56
4.10	(a) xy-plane as an illustration of the idea of a connection between two points. (b) An example of parameter space to describe the general equation of a straight line where infinitely lines pass through. These graphics are adapted from [52].	56
4.11	(a) The geometrical interpretation of the parameters ρ and θ . (b) Each sinusoidal curve represents the collection of lines that pass through a particular point (x_k, y_k) in the xy -plane. (c) An example of the accumulator cells, where every non-zero pixel, this sinusoidal curve will be created and counted in these cells. These graphics are adapted from [52]	57
4.12	An example implementation of Hough's transform to align the X-ray wrist image in a certain angle degree. (a) An example of the wrist X-ray image, (b) The result of vote pixels in accumulator cells between ρ and θ on the image proceeds through the transform where the brightest colour represents potential lines in the input image.	58

4.13	An graphic illustration of transformation T grayscale r value of an image to a new grayscale value of pixels s	60
4.14	An example of an image which has size of 4×4 and 16 pixels.	60
4.15	A distribution of pixels within image in Figure 4.14. In assumption, this image has ten levels of the grayscale colour map. This histogram will be the basis of the histogram equalisation process.	60
4.16	These are comparisons between an original image (a) and the result (b) after proceeding through the histogram equalization technique.	61
4.17	A distribution of pixels within image in Figure 4.16(b). In assumption, this image has 10 levels of the grayscale colour map. This histogram is the result of image (a) as the basis of the histogram equalization process.	62
4.18	An example of histogram equalisation applied to a medical image. Image (a) is an original image and (b) is an image resulting from the classic histogram equalization process.	62
4.19	Two histograms applied to medical image in Figure4.18. Image (a) is a histogram of the original image of Figure4.18(a) and (b) is a histogram of Figure4.18(b) resulting from the histogram equalization process.	63
4.20	An example of contrast limited histogram equalisation applied to a medical image. Image (a) is an original image and (b) is an image resulting from CLAHE process.	63
4.21	Two histograms of a wrist X-ray image in Figure 4.20. Image (a) is a histogram of original image of Figure 4.20(a) and (b) is a histogram of Figure 4.20(b) resulted from contrast limited histogram equalisation process.	64
4.22	An illustration of eight pixels neighbourhood. These pixels are ordered in clockwise as the eight elements of Texture Unit. The first element is a as element of E_1 clockwise-way through h as the element E_8	65

4.23	An illustration to describe Local Binary Patterns (LBP) in a texture. A 3×3 neighbourhood (<i>a</i>) is thresholded by the center pixel's value. The values of each pixel in (<i>c</i>) then multiplied with two-level thresholding results in (<i>b</i>) resulted (<i>d</i>). Texture unit of this neighbourhood is obtained by sum up the eight pixels which got $TU = 180$. Images inspiration from [96].	66
4.24	A diagram showing AlexNet CNN Architecture [70].	68
4.25	A diagram showing VGG-16 CNN Architecture which is adapted from [121].	69
4.26	A diagram showing VGG-19 CNN Architecture which is adapted from [121].	69
4.27	A diagram showing Squeezenet CNN Architecture which is adapted from [64].	70
4.28	A diagram showing ResNet-18 CNN Architecture which is adapted from [59].	70
4.29	A diagram showing GoogLeNet architecture which is adapted from [126].	70
4.30	A diagram showing ResNet-50 CNN Architecture is adapted from [59].	71
4.31	A diagram showing Inception-ResNet-v2 CNN Architecture which was taken from Inception-V3 on Google Cloud as it adapted from [127].	71
4.32	A diagram showing ResNet-101 CNN Architecture is adapted from [59].	72
4.33	A diagram showing DenseNet-201 CNN Architecture is adapted from [61].	72
4.34	A diagram showing Inception-ResNet-v2 CNN Architecture. The diagram is adapted from [127].	72
4.35	An illustration in the tree model of image data augmentation varieties. The graphic is adapted from [120].	75
4.36	An example of the wrist being augmented into several techniques to make a variety of images within the limited size of a data set through image data augmentation strategy.	75
4.37	A generic illustration of the transfer learning process from a pre-trained convolutional neural network. This image has been adapted from MATLAB [®] documentation materials.	77
4.38	Illustration of the Class Activation Mapping. It generates the highlight of the class intention regions within an X-ray image of the wrist.	78

5.1	Graphical illustration of the demographic distribution of the population of the study. (a) The age distribution is shown with boxplots, one per group of the study. (b) Female and Male distribution is shown with bars, one per group of the study.	85
5.2	Six representative radiographs that were collected from previous clinical activity at Royal Devon and Exeter NHS Foundation Trust Emergency Department. The images present considerable variability in the quality, positioning of the arm and presence of lines caused by the collimator also caught during X-ray image acquisition. The images were anonymised and metadata such as age, date of acquisition, gender and clinical outcome was available.	91
5.3	Automatic pre-processing of the radiographs. The six representative cases shown were automatically rotated so that the forearm‘ position aligned vertically. In addition, the artifacts due to the collimator were removed.	92
5.4	Semi-automatic extraction of measurements of the forearm. (a) Original radiograph that presents rotation of the arm and artefactual lines due to the collimator. Three landmarks have been manually located in the base of the lunate (red), <i>Radial Styloid</i> (green), and centre of the middle finger (blue). (b) Automatic pre-processing of the image where the forearm was aligned vertically and the lines removed. (c) Using the lunate landmark as a guide, the boundaries of the forearm were automatically delineated and lines were traced between the boundaries. The distance between the lines is 1 cm and was being used to derive swelling measurements of the wrist.	93

5.5	Semi-automatic extraction of measurements of the finger (a) Region of interest (ROI) of the central finger generated from the landmark, blue dot in Fig 5.4 (a). (b) Identification of regions of cortical bone (shaded in cyan) and trabecular bone (shaded in pink) from which the ratio of cortical to total area was calculated. Notice that the finger was rotated to align vertically as the previous rotation aligned the forearm but the fingers are not necessarily vertical. (c) Intensity profile of the ROI with the following key points: edges of the bone (magenta diamond), the peak of cortical bone (blue asterisk), and centre of the bone (red circle).	93
5.6	Semi-automatic extraction of texture measurements of a region of interest. (a) To analyse the texture of the radius, an ROI is automatically located by traversing a fixed distance from the <i>Radial Styloid</i> landmark. (b) Zoom of the region of interest. (c) Texture coefficients generated by Local Binary Pattern analysis.	93
5.7	Semi-automatic extraction of texture measurements from intensity profiles. (a) Profile lines from the <i>Radial Styloid</i> . Initially, a line (green) is automatically traced between the lunate (red in Figure 5.4 a) the <i>Radial Styloid</i> (green in Figure 5.4 a) landmarks. Two lines are automatically derived from the first, one at 30 degrees (red) and one at 45 degrees (blue) from the <i>Radial Styloid</i> up to the edge of the radius, which is automatically detected. (b) Intensity profiles correspond to the lines traced in (a). Notice the increasing slope. (c) Intensity profiles are adjusted by removing the slope.	94
5.8	Boxplots corresponding to distributions of four representative measurements.	95
6.1	Illustration of different radiographs of the Musculoskeletal Radiographs (MURA) data set corresponding to the training set and negative (no abnormalities) in the top row, and positive (abnormalities) in the bottom row. (a) Elbow, (b) Forearm, (c) The Postero-Anterior view of Wrist, (d) Lateral view of Wrist, (e) Elbow, (f) Fingers, (g) Forearm, (h) Hand.	100

6.2	Block diagram which illustrates the classification of the wrist radiographs with 11 different Convolutional Neural Network (CNN) architectures. 9752 images from Musculoskeletal Radiographs (MURA) Wrist data set were used for training CNN architectures and 659 images were used for validation. Two different metrics, Accuracy (A_c) and Cohen's Kappa (κ) were computed to assess the performance of 11 pre-trained CNNs. Image data augmentation was used during training and a different number of epochs and mini batch sizes were tested.	103
6.3	Illustration of classification results for Lateral (LA) views of wrist radiographs. (a) Corresponds to positive (abnormal) diagnosis image but predicted as negative (normal), (b) Abnormal diagnosis, and abnormal prediction. (c) Normal diagnosis image and normal prediction. (d) Normal diagnosis and abnormal prediction. Notice that the errors in classification may have been biased by artefactual elements on the images.	107
6.4	Illustration of classification results for Postero-Anterior (PA) views of wrist radiographs. (a) Corresponds to positive (abnormal) diagnosis image but predicted as negative (normal), (b) Abnormal diagnosis and abnormal prediction. (c) Normal diagnosis image and normal prediction. (d) Normal diagnosis and abnormal prediction. Notice again that the errors in classification may have been biased by artefactual elements on the images.	108
7.1	Eight examples of radiographs without abnormalities (considered negative) of the M usculoskeletal R adiographs (MURA) data set. (a) Elbow, (b) Forearm, (c) Shoulder, (d) Wrist (lateral view), (d) Lateral view of Wrist, (e) Finger, (f) Hand, (g) Humerus, (h) Wrist. It should be noted the variability of the images in terms of dimensions, quality, contrast, and the large number of labels (i.e., R for right and L for left), which appear in various locations.	115

7.2	Eight examples of radiographs with abnormalities (considered positive) of the M usculoskeletal R adiographs (MURA) data set. (a) Elbow, (b) Forearm, (c) Shoulder, (d) Wrist (lateral view), (e) Finger, (f) Hand, (g) Humerus, (h) Wrist (PA view). As for the cases without abnormalities, it should be noted the variability of the images and in addition the abnormalities themselves. There are cases of metallic implants some of which are smaller (a) than others (b), as well as fractures.	116
7.3	Illustration of Class Activation Mapping which generates the highlight of the class intention regions.	116
7.4	Illustration of the Class Activation Mapping overlaid on the four classification results for (a,b) Postero-Anterior views shown in Figure 6.4.	119
7.5	Illustration of the Class Activation Mapping (CAM) overlaid on the four classification results for (a,b) Lateral views shown in Figure 6.3	119
7.6	Illustration of activation maps overlaid over the eight radiographs with abnormalities of Figure 7.2 to indicate the regions of the image that activated a ResNet 50 architecture. (a) Elbow, (b) Forearm, (c) Shoulder, (d) Wrist (lateral view), (e) Finger, (f) Hand, (g) Humerus, (h) Wrist. The activation maps illustrate the location of the abnormalities, e.g., (a,e), but appears spread in other cases (b,g) where the abnormality is detected together with a neighbouring region. In other cases (c), the abnormality is not detected.	120
7.7	Illustration of activation maps overlaid over the eight radiographs without abnormalities of Figure 7.1 to indicate the regions of the image that activated an Inception-ResNet-V2 architecture. (a) Elbow, (b) Forearm, (c) Shoulder, (d) Wrist (lateral view), (e) Finger, (f) Hand, (g) Humerus, (h) Wrist. It should be noted that the activation regions are more localised than those of the ResNet-50.	121

7.8	Illustration of activation maps overlaid over the eight radiographs with abnormalities of Figure 7.2 to indicate the regions of the image that activated an Inception-ResNet-v2 architecture. As for the cases without abnormalities, the activation regions are more located, e.g., ((c) Shoulder,(d) Lateral view of Wrist, and (h) Posterior-Anterior view of Wrist)) and in addition, the abnormalities are better located, e.g., ((a) Elbow, (b) Forearm, (e) Finger, (f) Hand, and (g) Humerus).	122
7.9	Illustration of class activation mapping overlaid over the eight radiographs without abnormalities of Figure 7.1 to indicate the regions of the image that activated a ResNet-50 architecture. (a) Elbow, (b) Forearm, (c) Shoulder, (d) Wrist (lateral view), (d) Wrist (lateral view), (e) Finger, (f) Hand, (g) Humerus, (h) Wrist. As these cases are positive (no abnormality), the regions of activation are not as critical as those with abnormalities.	123
7.10	Illustration of the effect of the number of layers of architectures against the two metrics used in this work accuracy and Cohen’s Kappa.	126
A.1	Poster presented at the MIUA 2019 conference	133
B.1	Thirty sample images of Wrist X-ray overlaid with CAM based on ResNet-50 architecture.	136
B.2	Thirty sample images of Wrist X-ray overlaid with CAM based on Inception-ResNet-V2 architecture.	138

List of Tables

2.1	This table is an example of metadata attributes from the DICOM image. These attributes are standard for communicating and managing medical imaging information and related data [90].	27
3.1	This table shows the exact number of data from Exeter’s data set, which has a value of the PatientSex metadata attribute.	33
3.2	This table shows the exact number of data from Exeter’s data set with a value of SeriesDescription, ViewPosition metadata attribute related to Lateral and Posterior-Anterior which are focus positions of this work.	33
3.3	This table shows the exact number of data from Exeter’s data set that has a value of the PhotometricInterpretation metadata attribute.	34
3.4	Distribution of cases of the Stanford MURA (musculoskeletal radiographs) data set [111] for studies of the upper body.	34
3.5	Distribution of images in MURA (musculoskeletal radiographs) data set for training and validation.	35
3.6	Distribution of images in the Stanford MURA (MUsculoskeletal RA-diographs) data set into Abnormal and Normal groups. This work concentrated on the wrist radiographs.	36
3.7	MURA’s Wrist train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	39
3.8	MURA’s Wrist valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	40
3.9	MURA’s Elbow train image data set breakdown based on the number of studies taken for the patient and condition detected on the image	40
3.10	MURA’s Elbow valid image data set details the number of studies taken for the patient and condition detected on the image.	41

3.11	MURA’s Forearm train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	41
3.12	MURA’s Forearm valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	42
3.13	MURA’s Finger train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	42
3.14	MURA’s Finger valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	42
3.15	MURA’s Hand train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	44
3.16	MURA’s Hand valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	44
3.17	MURA’s Humerus train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	44
3.18	MURA’s Humerus valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	45
3.19	MURA’s Shoulder train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	45
3.20	MURA’s Shoulder valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.	45
4.1	These are types of geometric transformation usually applied to an image in the image processing field [52].	50
4.2	Values of r_k if $L = 8$	59
4.3	Detail calculation of histogram equalisation of image Figure 4.14. Abbreviations: Frequency (freq), Cumulative Distribution (Cumltv Dist).	61
4.4	Details of convolutional neural networks (CNNs) that were used in this work.	67

5.1	Measurements extracted from the radiographs. The third column corresponds to the landmark used to calculate the measurement. Columns 4 – 10 show the p-values result of paired t-tests between different groups. Values lower than 0.05 are highlighted in bold. Abbreviations: Ratio of width line 1 / width line 4 (W1/W4), Local Binary Pattern (LBP), Standard Deviation (Std), Measurement (Msrmt), Landmark (Ldrk), Control (Ctrl), Patient (Pat) Unsuccessful (Unsucc), Successful (Succ), Pre-Successful (Pre-Succ), Post-Successful (Post-Succ), Pre-Unsuccessful (Pre-Unsucc), Post-Unsuccessful (Post-Unsucc)	89
6.1	MURA’s wrist radiographs are subdivided into four studies.	99
6.2	Summary of convolutional neural networks (CNNs) hyperparameters.	101
6.3	Results of accuracy for eleven architecture of Convolutional Neural Networks used to classify the wrist images in MURA data set experiments. The best results for each row are highlighted in <i>italics</i> and the overall best results are highlighted in bold . Abbreviations: Epoch (Ep).	105
6.4	Cohen’s Kappa results from eleven architecture of Convolutional Neural Networks used to classify the wrist images in MURA data set experiments. The best results for each row are highlighted in <i>italics</i> and the overall best results are highlighted in bold . Abbreviations: Epoch (Ep).	106
6.5	The effect of mixed domain knowledge of anatomy data set to the area under curve and times of computation (ResNet-50 case study). Abbreviations: Times of computation (Time of comp), Area Under the Curve (AUC).	106
6.6	Results of accuracy for eleven Convolutional Neural Networks used to classify the wrist images in the MURA data set. The best results for each row are highlighted in <i>italics</i> and the overall best results are highlighted in bold . Abbreviation: augmentation (aug).	109

6.7	Cohen’s Kappa results from eleven Convolutional Neural Networks architectures used to classify the wrist images in the MURA data set. The best results for each row are highlighted in <i>italics</i> and the overall best results are highlighted in bold . Abbreviation: augmentation (aug).	109
7.1	Summary of CNNs Activation Size and Parameters	125
7.2	Summary of accuracy within chosen activation layer of CNN.	126

Acknowledgements

I would like to thank my first supervisor, Dr Constantino Carlos Reyes-Aldasoro, for his unlimited guidance and support from the beginning of accepting me as his doctoral student and starting my PhD study at the City, University of London. Thank you City, University of London for the PhD studentship to fund my study. I would also like to thank Professor Eduardo Alonso, my second supervisor, who gave me a lot of advice and supports in this stage of my PhD study. I would like to thank Dr Alex Ter-Sarkisov, my third supervisor, who gave me valuable suggestions on my doctoral research's both in technical and non-technical aspects. Huge appreciation and thanks to Dr Karen Knapp from the University of Exeter for providing me with valuable clinical knowledge, a precious wrist X-ray data set, and explaining this classic yet challenging clinical problem. Thank you Stanford University Machine Learning Group for MURA data set.

I spent my doctoral years in two research centres: the Research Centre for Biomedical Engineering (RCBE) and the giCentre. These two research centres are places where I belonged during my PhD years. I would like to acknowledge Professor Panayiotis Kyriacou and all members of the Research Centre of Biomedical Engineering and also to Professor Jo Wood and all members of giCentre for accepting, sharing and welcoming me into your research groups. I would also like to shout my big regards to Cefa Karabağ, Kwun Ho Ngan, José Alonso Solís-Lemus, Zaibaa Patel, Tomas Ysehak Abay, Nystha Baishya, Riad Ibadulla, Youssef Arafat, Mirela Reljan-Delaney, Kevin Allain, SMCSE staff Nathalie Chatelain, Gabrielle To, Asif Nawaz, City's canteen crews and my employer in Indonesia, Politeknik Caltex Riau.

My special thank you to my parents, my parents-in-law, whole family from both sides. My ultimate special thank you to my lovely beautiful wife **Fadilla Rizki** and my beloved son **Mahdi Rizki Ananda** for enormous support during this journey. Thank you England, thank you London. Thank you.

Abbreviations and Acronyms

<i>Ac</i>	Accuracy
<i>Adam</i>	Adaptive moment estimation
<i>AHE</i>	Adaptive Histogram Equalization
<i>AI</i>	Artificial Intelligence
<i>AIA</i>	<i>Anterior Interosseous Artery</i>
<i>AUC</i>	Area Under Curve
<i>Aug</i>	Augmentation
<i>BAA</i>	Bone Age Assessments
<i>CAM</i>	Class Activation Mapping
<i>CLAHE</i>	Contrast-Limited Adaptive Histogram Equalization
<i>CNN</i>	Convolutional Neural Networks
<i>CT</i>	Computed Tomography
<i>Cumltv Dist</i>	Cumulative Distribution
<i>DICOM</i>	Digital Imaging and Communications in Medicine
<i>Ep</i>	Epoch
<i>FOOSH</i>	Fall onto an outstretched hand
<i>GP</i>	Greulich and Pyle
<i>LA</i>	Lateral
<i>LBP</i>	Local Binary Pattern
<i>MRI</i>	Magnetic Resonance Imaging
<i>MuA</i>	Manipulation under <i>Anaesthesia</i>
<i>MURA</i>	MUsculoskeletal RAdiographs
<i>ORIF</i>	Open Reduction and Internal Fixation
<i>PNG</i>	Portable Network Graphics
<i>PA</i>	Postero-Anterior
<i>ReLU</i>	Rectified Linear Unit
<i>RH</i>	Radius Height
<i>SDI</i>	Socio-demographic Index
<i>SGDM</i>	Stochastic Gradient Descent with Momentum
<i>RMSPprop</i>	Root Mean Square Propagation
<i>Time of comp</i>	Times of computation
<i>TU</i>	Texture Unit
<i>TW2</i>	Tanner Whitehouse
<i>TWA</i>	Total Wrist <i>Arthroplasty</i>
<i>VT</i>	<i>Volar</i> Tilt

Publications & Version control

1. A. Ananda, K. M. Knapp, A. Appelboam, E. Alonso and C. C. Reyes-Aldasoro, 'Digital Image Analysis of Distal Radius Fractures X-Ray for Manipulation under Anaesthesia', Poster session, in the 23rd Medical Image Understanding and Analysis (MIUA) conference, Jul.2019.
2. A. Ananda, C. Karabağ, A. Ter-Sarkisov, E. Alonso, and C. C. Reyes-Aldasoro, 'Radiography Classification: A Comparison between Eleven Convolutional Neural Networks', in 2020 Fourth International Conference on Multimedia Computing, Networking and Applications (MCNA), Oct. 2020, pp. 119–125. doi: 10.1109/MCNA50957.2020.9264285.
3. C. C. Reyes-Aldasoro, K. H. Ngan, A. Ananda, A. d'Avila Garcez, A. Appelboam, and K. M. Knapp, 'Geometric semi-automatic analysis of radiographs of Colles' fractures', PLOS ONE, vol. 15, no. 9, p. e0238926, Sep. 2020, doi: 10.1371/journal.pone.0238926.
4. A. Ananda.; Ngan, K.H.; Karabağ, C.; Ter-Sarkisov, A.; Alonso, E.; Reyes-Aldasoro, C.C. Classification and Visualisation of Normal and Abnormal Radiographs; A Comparison between Eleven Convolutional Neural Network Architectures. Sensors 2021, 21, 5381. <https://doi.org/10.3390/s21165381>
5. PhD's project GitHub address : <https://github.com/amno/WristMatlab>

Chapter 1

Introduction

1.1 Overview

Image analysis is the process to extract information from an image or collection of images. Digital image processing as stated by [52] ‘encompasses processes whose inputs and outputs are images and, in addition, includes processes that extract attributes from images up to, and including, the recognition of individual objects’. Digital image processing covers techniques such as edge detection, image filtering, thresholding, affine transformations, and morphological operations [52].

Digital image processing techniques can be applied to problems such as finding shapes, removing noise, segmentation, classification, feature extraction, and statistic calculations for texture analysis [52].

Encyclopedia Britannica defines medical imaging as ‘the use of electromagnetic radiation and certain other technologies to produce images of internal structures of the body for the purpose of accurate diagnosis’ [41].

Medical imaging includes many modalities such as X-ray [13], magnetic resonance imaging (MRI) [133], computed tomography (CT) [92], ultrasound [4], positron emission tomography (PET) [72], and nuclear medicine [125]. Examples of these medical imaging modalities show in Figure 1.1.

Medical image analysis provides tools to aid the diagnosis of clinical problems as well as follow the effects of certain treatments [76]. The origin of medical image analysis is Computer Vision, which has become a sub-discipline of Computer Science. This field emerged in the early 1990s when a group of researchers and others with backgrounds in Computer Vision began applying mathematical methods to analyse

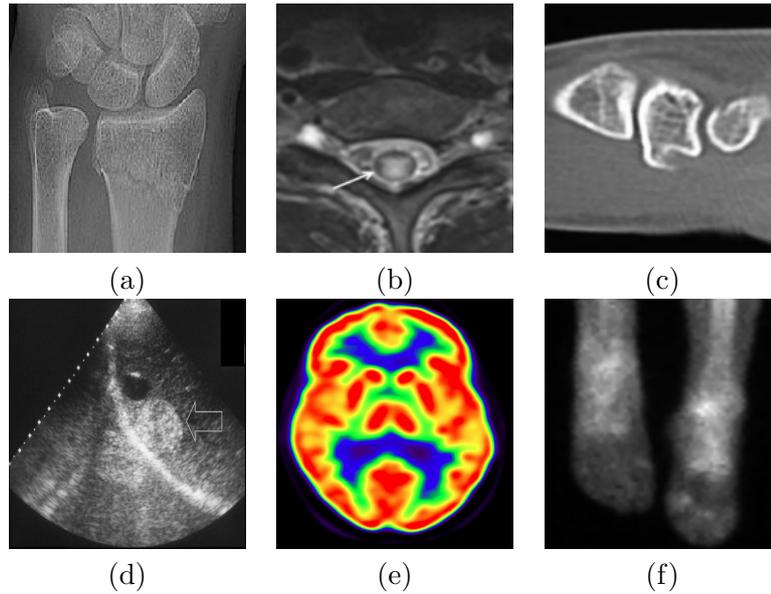


Figure 1.1: Medical imaging modalities examples. (a) X-ray [111], (b) MRI [133], (c) CT [92], (d) Ultrasound [4], (e) PET [72], (f) Nuclear medicine [125]. Each image is credited to its respective references.

problems related to medical images [141].

Medical image analysis has helped doctors analyse medical conditions like fractures [79, 80], tumours [48, 51], diseases like *osteoporosis* [69, 89], *tuberculosis* [78, 100] and cancer [23, 119].



Figure 1.2: An example of the open fractures X-ray images. The image taken from [147].

This thesis is focused on wrist bone fractures, which can be classified into open or closed fractures. An *open fracture* is a fracture that has pierced the skin. Open fractures are typically caused by high-energy injuries such as car accidents, a gunshot,

sports injuries, or falls. *Open fractures* require immediate surgery to avoid infection because of the break in the skin. Figure 1.2 shows an example of radiographs of an open fracture of the *distal radius* with dislocation of distal *radioulnar* joint.

A *closed fracture* is a fracture that has not punctured the skin and has no open wound. The injured body part may be swollen or bruised. The X-ray images are used to check how severe the condition is [109]. Figure 1.3 shows an example of a *closed fracture* of the distal radius [33] as resulting from a *fall onto an outstretched hand* (FOOSH) [67].

X-ray images of the wrist are normally acquired in two projection views. *Postero-Anterior* (PA) and the *Lateral* (LA). The PA view of the wrist (Figure 1.3 (a)) is obtained with the hand pronated or supinated, the hand resting flat and in direct contact with the surface of the image receptor. The LA view (Figure 1.3 (b)) is obtained by maintaining the *Lateral* position of the wrist and on the same horizontal plane as the surface of the image receptor. The LA view is taken from the side part of the wrist. The overlapping bones also can be seen from the LA view [129].

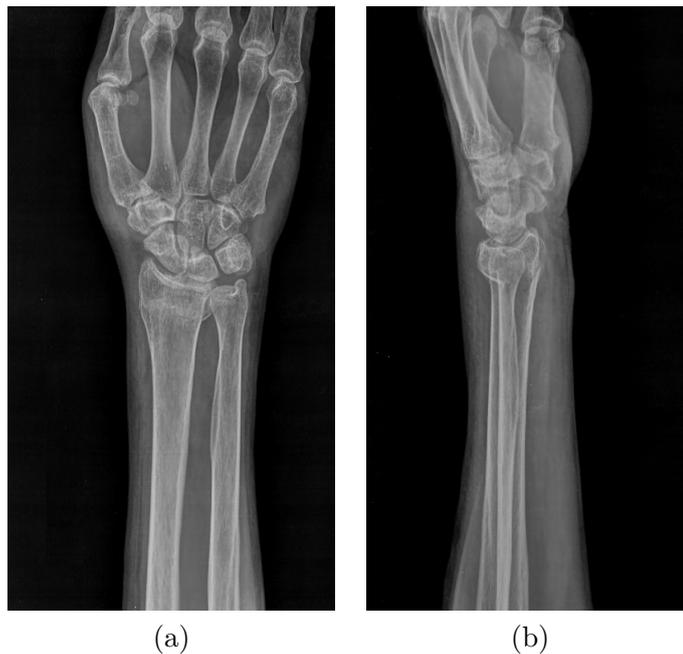


Figure 1.3: An example of the closed fractures X-ray images. (a) *Postero-Anterior* view. (b) *Lateral* view. The images are credited to [49].

Diagnosis, as defined by the Cambridge dictionary, is a judgment about what a particular illness or problem is, made after examining it [108]. The diagnosis generally includes a physical observation of the affected body part and medical imaging. The radiograph is one of the medical imaging technologies to support a

diagnosis. Radiology is a branch of medicine using radiation for the diagnosis and treatment of disease [42].

The diagnosis of medical images entirely depends on the radiologist. The interpretation of findings in medical images by a radiologist could be different from another radiologist [135]. The radiograph interpretation skill takes years of education, training, experience, and expertise to identify the abnormality in the image observed [140]. Radiologists principally rely on visual inspection that makes errors and discrepancies in radiology practice inevitable [19, 135].

An X-ray image of bone injury is diagnosed by looking at the anatomical location, type of fracture, the direction of fracture lines, fracture positions, and the joint between fragments [109]. Radiographic parameters such as *Radius Height* (RH) and *Volar Tilt* (VT) shown in Figure 1.4 are usually used when evaluating wrist X-ray images.

Radius Height (RH) also known as *Radial Height* is a radiograph parameter that measures the difference in length between the *ulnar* head and the tip of the *radial styloid* on the PA view [102]. *Volar Tilt* (VT) is a radiographic parameter that measures from an angle at the articular surface of the radius, a tangent line is drawn from *dorsal-to-volar*, followed by a line perpendicular to the long axis of the *radius* [102].

Radius Height (RH) and *Volar Tilt* (VT) are then compared with the clinical outcomes of the patients whilst still requiring general agreement to treat young and old patients [20]. A systematic approach to diagnosing trauma like wrist fracture has been proposed by [73]. After a diagnosis is concluded, the treatment process could be started.

The treatment of wrist fracture can be through a surgical fixation, also known as Open Reduction and Internal Fixation (ORIF) [11, 94] or through manual therapy such as Manipulation under Anaesthesia (MuA) [54]. Surgical fixation is needed for immediate action to avoid infection in open wounds and could involve many treatment options [11, 94]. The Manipulation under Anaesthesia (MuA) which includes closed reduction and casting, applied when there is no need for a surgical procedure based on radiologists' treatment decision. This non-surgical procedure is often the primary option undertaken in Emergency Departments for displaced fractures in an attempt to correct the deformity and represents a significant proportion of the department workload [12]



Figure 1.4: Radiographic Parameters. (a) Radius Height (RH). (b) Volar Tilt. The images are credited to [20].

In this thesis, the main focus is the analysis of wrist fractures observed with radiographs. The contributions of this thesis are: (i) using geometric attributes to analyse fracture patterns, (ii) abnormality detection on the wrist X-ray image through deep learning techniques (iii) visualising suspected abnormalities within the wrist X-ray through training of convolutional neural networks.

1.2 Motivation

The human wrist is a complex joint between the five metacarpal bones of the hand and the radius and ulna bones of the forearm [43]. The wrist together with the hand has interactions that depend on the integrity and function of the ligaments, tendons, muscles, joints, and bones [104]. Problems in any of these can affect upper extremity function, causing disruptions of daily activities and negatively impacting the quality of life [134]. An example of a medical problem related to the wrist is bone fractures.

In the United Kingdom, wrist fractures have become an intensive study, for example, the association between frailty with incidence and mortality of fractures in people aged over 80 years [113] as well as the epidemiology of fractures which analysed in varieties of demographic factors such as age, sex, geographic location, ethnicity, and socioeconomic status which may inform public health policy in the

UK and elsewhere [29, 85].

Globally, the trend of hand and wrist trauma has also been studied. According to [28], the incidence of hand trauma has modestly decreased but low-middle and middle Socio-Demographic Index (SDI) nations have demonstrated increased rates of fracture and amputation in almost three decades as well as the limited access to better surgical hand care is still an issue [28]. PubMed (<https://pubmed.ncbi.nlm.nih.gov>), the on-line tool of the National Library of Medicine of the United States, reported 51,343 entries of ‘Wrist’, 12,212 entries of ‘Wrist Anatomy’, 10,411 entries related to ‘Wrist Fractures’, and 8,638 entries indexed as ‘Distal Radius Fracture’ (last accessed on the 11th of April 2022).

The decision to treat a wrist injury is still based on visual observations both physically and through medical imaging such as X-ray images where errors and discrepancies occur [19]. The other important factor whilst diagnosing a patient is the agreement between the doctors which can be related to personal expertise [135]. The decision to treat a patient with Manipulation under Anaesthesia (MuA) is primarily of lower-level of evidence without widely accepted guidelines [54].

A comprehensive review of radiographs allows the doctor to understand the injury better and then provide appropriate treatment [18]. The presence of bone-related diseases such as *osteoporosis* also needs to be highlighted during the imaging review which can provide an appropriate fracture diagnosis [69].

Studies related to detecting abnormality such as fracture within X-ray image of the wrist has become a strategic issue to be explored [69]. An appropriately performed radiographic examination helped to improve clarity and certainty interpretation of the potential area of bone injury [50].

A medical image’s quality is characterised by contrast, noise, and spatial resolution. A medical image’s quality depends on the imaging system hardware and how the radiologist obtained the images. Image defects such as blurry images, images with too low or high brightness and sharpness levels, and images with an improperly positioned wrist scan can be a visual barrier to the radiologist. Therefore, it is important to obtain good quality images [62].

Nowadays, the medical image analysis field has evolved by leveraging existing medical image collections. Exploration opportunities are carried out by utilising computational capabilities such as machine learning to analyse the patterns of images in the image collection like X-ray images [31]. Cases such as wrist fractures pattern

could be analysed through available data sets.

1.3 Research questions

The wrist X-ray images present challenges in detecting abnormalities like fractures. In this thesis, several research questions with the wrist X-ray image are considered:

1. How can the geometric features of a wrist X-ray be used to analyse the wrist in relation to fractures?
2. How can Deep Learning be used to classify wrist X-ray images into normal and abnormal X-ray images?
3. How do we know which part of the wrist X-ray image might trigger the decision to be abnormal by a deep learning model?

1.4 Aim and Objectives

The aim of this thesis is to analyse wrist fractures as observed with X-ray imaging through traditional image processing techniques and deep learning models which are expected to be able to provide a better technological understanding of patient treatment. The following objectives are presented:

1. To analyse the wrist fractures through the geometric features of a wrist X-ray image.
2. To investigate the performance of deep learning models in classifying wrist X-ray images into normal and abnormal categories.
3. To investigate the visualisation of the suspected areas of abnormality within a wrist X-ray image from a trained Convolutional Neural Network through the Class Activation Mapping technique.

1.5 Thesis structure

An overview summary of each chapter in this thesis is shown in Figure 1.5. It provides a graphic description of this thesis and the relation of each chapter in this thesis.

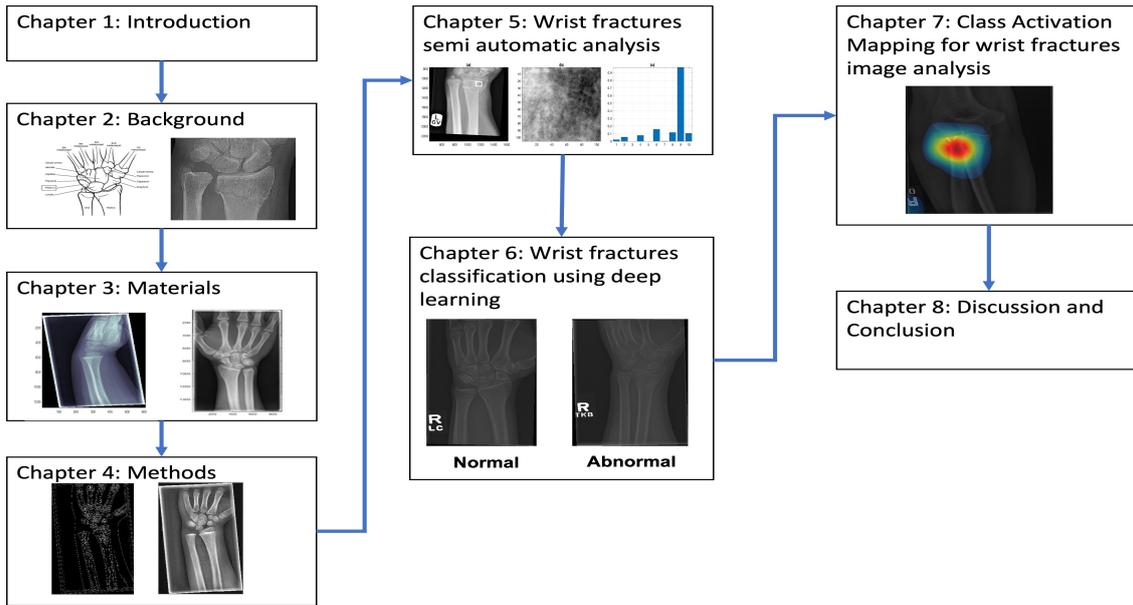


Figure 1.5: Graphical outline of the thesis.

Chapter 1. Introduction describes the motivation of the research, problem definition, aim and objectives of the research and report structure of this thesis. An illustration outline to give a brief thesis summary.

Chapter 2. Background This chapter is divided into medical background, computational background and literature review.

Chapter 3. Materials presents the data for this thesis. There are two X-ray data sets explored. First, the data set that experts from the University of Exeter acquired. The second data set came from the Stanford Machine Learning Group named Musculoskeletal Radiographs (MURA).

Chapter 4. Methods describes the most common medical image analysis techniques. First, general image processing methodologies such as geometric transformations, edge detection, and contrast-limited adaptive histogram equalisation. Second, an image texture-based technique named Local Binary Patterns (LBP) was used to analyse texture-based measurements extraction from wrist X-ray images. Third, the theory of class activation mapping to visualise attention areas within wrist X-ray images leads to a decision from the convolutional neural networks of an input image to classify.

Chapter 5. Wrist fractures semi automatic analysis investigates how to identify features within the wrist X-ray image related to dorsally displaced wrist fractures. This chapter explains how semi-automatic analysis consists of a manually identified finger, lunate, and radial styloid as landmarks and automatic processing

to generate geometric features and texture measurements.

Chapter 6. Wrist fractures classicification using deep learning contains an explanation of identification of fractures within wrist X-ray images. The wrist images were categorised as abnormalities identified using eleven convolutional neural network architectures.

Chapter 7. Class Activation Mapping for wrist fractures image analysis explores the benefit of class activation mapping as visualisation prediction by deep learning. Class activation mapping has the benefit of identifying networks' intentions when deciding on a classified input image.

Chapter 8. Discussion and Conclusion offers a discussion and concludes the thesis.

Chapter 2

Background

This chapter describes the background of the thesis. The background is divided into medical background and computational background. The medical background begins with the anatomy of the human wrist, wrist-related fractures, and the diagnosis and treatment of wrist fractures. A non-surgical treatment called *Manipulation under Anesthesia* (MuA) is also described. The computational background includes imaging techniques for diagnosing wrist fractures.

2.1 Medical background

2.1.1 Human wrist

The human wrist contains 15 unique bone parts: eight *carpal* bones, one *ulna* bone, one *radius* bone, and five *metacarpals*. The bones are *Trapezium*, *Trapezoid*, *Capitate*, *Scaphoid*, *Radial* bones such as *Radial Styloid Process* and *Radial Articular Surface*, *Ulna Styloid Process*, *Lunate*, *Triquetrum*, *Pisiform*, and *Hamate* [109] (Figure 2.1). The *Radius* and the *Ulna* bones are two bones usually called the forearm. *Lunate*, *Pisiform*, *Triquetral*, *Capitate*, *Hamate*, *Trapezoid*, *Trapezium* and *Scaphoid* are collectively named the palm. Metacarpals bone from the 1st to the 5th are called the fingers.

A ligament is a fibrous connective tissue that attaches bone to bone (Figure 2.2). An adult human wrist consists of 16 ligaments connected to each bone in the hand anatomy. The ligaments are uniquely coded to make it easier to identify which bone it is connected to (Figure 2.2). Each name of the ligament describes the bones and where its connected. The *anterior interosseous artery* (AIA) and *radial artery* (RA)

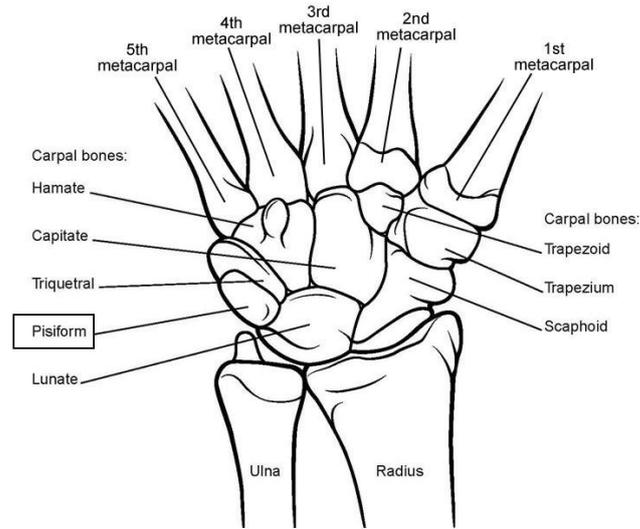


Figure 2.1: The drawing of human wrist anatomy. Image credited to [104].

have connected the *Radial* and the *Ulna* bones. The *capitohamate* (CH) connects bones of *Capitate* (C) and *Hamate* (H). The *long radiolunate ligament* (LRL) works around the *Radius* (R), and the *Lunate* (L). The *palmar radioulnar ligament* (PRU) works to join bones that form the palm, the *Radius* and the *Ulna*. There is also a ligament described based on the ligament's length size, such as *Short Radio Lunate* (SRL). The *Triquetrum* and the *Hamate* connected by *triquetrohamate* (TH) as well as the *Triquetrum* and the *Capitate* by *triquetrocapitate* (TC). The *Trapezoid* (TT) is a ligament to join the *Trapezium*. The *Ulna* to the *Carpal* connected by the *Ulnocarpal* (UC) ligament. The ligament between the *Ulna* and the *Lunate* is the *Ulnolunate* (UL) [104].

2.1.2 Wrist fractures

Wrist fractures problems

A wrist fracture happens when one or more of the bones of the wrist has cracked or broken. One type of wrist fracture is called the *distal radius fracture* [33]. This injury as resulting from a *fall onto an outstretched hand* (FOOSH) like accidental slip and falls [67].

The most common wrist fracture is Colles' fracture. Colles' fracture is described as a total break of the radius bone of the forearm resulting from a posterior displacement and changes in deformity (Figure 2.3 (a) and (d)) [25]. The broken

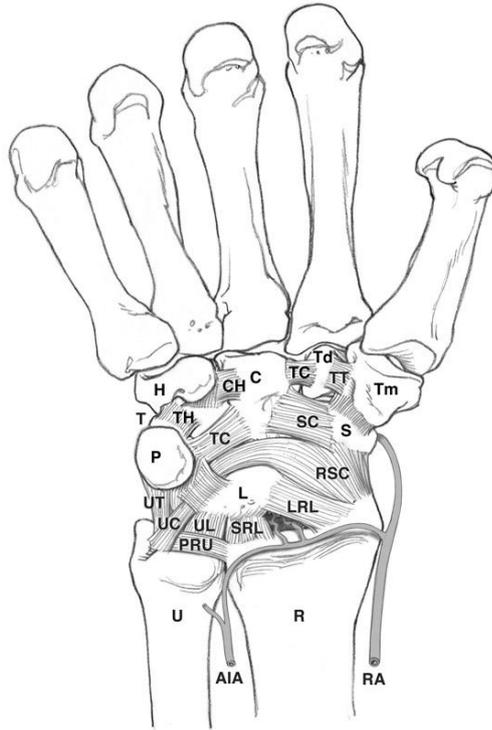


Figure 2.2: The drawing of human wrist anatomy completed with ligaments and bones. This image is taken from [104].

radio-carpal joint of a wrist (Figure 2.3 (b) and (e)) occurs within the *distal radius* is named Barton’s fracture. This injury is usually caused by a fall and wrist rotation of the hand and forearm so that the palm faces backward or downward. It increases the stress of the *Carpal* on the brink of the *Dorsal*. The Chauffeur’s fracture is another type of wrist fracture possibly caused by falling onto an outstretched hand, as shown in Figure 2.3 (c) and (f) [109, 55].

The position of image acquisition was also intended to see the side of the part observed. In Figure 2.3 (a), (b), and (c), there are several types of illustrations of wrist fractured X-ray images using the *Posterior-Anterior* (PA) view, where this position places the patient’s arm where the palms, wrists, and joints are shown. This PA position can also observe the actual conditions of the *Radial*, the *Ulna*, and the *Distal Radio-Ulnar Joint*. Small parts like the *Scapho-Lunate Distance* is also can be visually determined. The PA position is also can help on faults involving the *Radial* and the *Ulna* bones. The other position usually used is the *Lateral* (LA) position. The LA position takes an image from the side of the wrist and the palm. Figures 2.3 (d), (e), and (f), are taken on the lateral view and fractures that can be

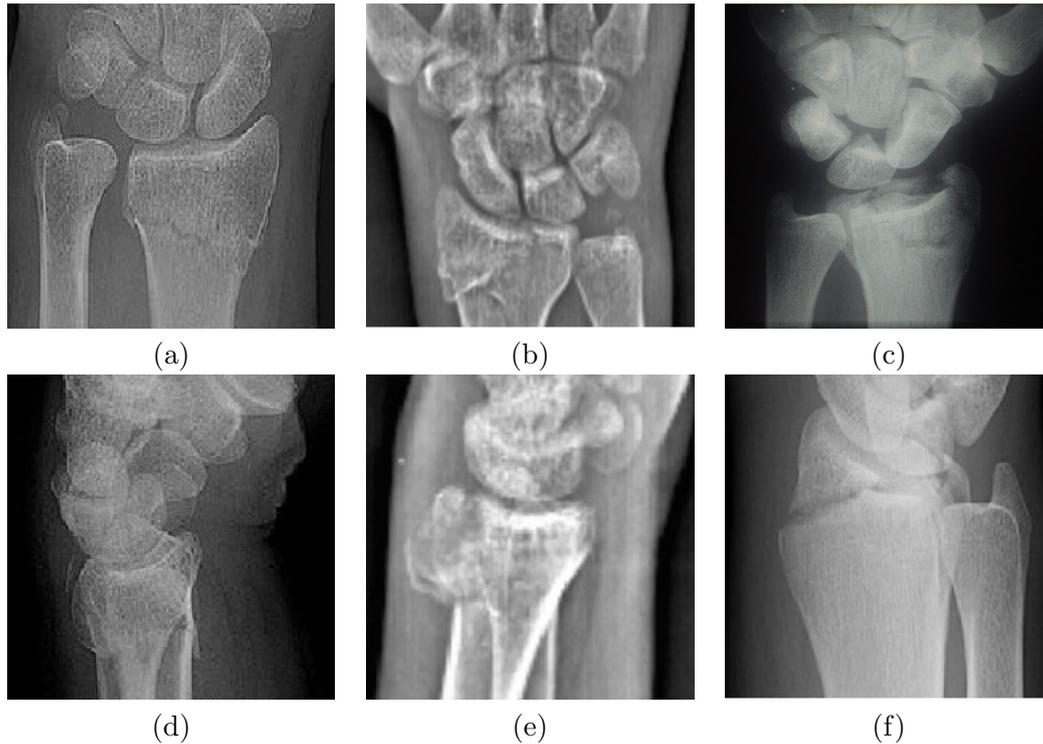


Figure 2.3: Illustration of three different types of wrist fractures. Colles' fracture (a) PA view [84], (d) Lateral view [84]. Barton's fracture (b) PA view [47], (e) Lateral view [47]. Chauffeur's fracture (c) PA view [123], (f) Lateral view [123]. Posterior-Anterior (PA) view and Lateral (LA) are two common positions the radiologist takes to determine conditions of wrist bone fractured visually.

seen from this side.

The *Osteoporosis* is usually described as a risk factor for bone fragility such as distal radial fractures. This bone fragility is studied by [101] where the *osteoporosis* could be distinguished if distal radius fractures occur early in life than hip and spinal fractures. Bone fragility is a global concern because it also has an impact on increasing mortality, decreasing the flexibility of movement, and increasing health costs [46, 134].

Wrist fractures diagnosis

When a patient arrives at a medical facility such as a hospital, clinic, medical centre, etc., medical staff will handle the patient. The pain is typically reduced with painkillers. A visual check is performed to look for swelling and deformity of the wrist. The decreased scope of dynamic movement is checked by moving the patient's wrist [109]. Unspecific pain of the hand and the wrist is a diagnostic challenge either by low specificity of clinical symptoms or due to inadequate imaging results [117].

A wrist fracture is categorised into two conditions: minor fracture and significant

fracture. A fracture is categorised as minor if there is no excessive bruise, wound, or open fracture that appears very clear. A plaster cast or splint will be applied to the patient's arm. They may also be given a wrist brace with or without an arm sling, painkillers, instructions on how to look after the injury, and probably follow-up appointments to re-check healing conditions. Major fractures are followed by attempting alignment-fractured bone. If needed, the surgery will follow the action of putting wires, plates, and screws inside the patient's wrist. It could be temporary, but there is a possibility of permanent placement. After the surgical procedure, the patient is asked for regular check fracture healing term-time. X-ray images are taken in the PA and the LA positions to check the post-surgery condition [109].

During the PA analysis check, there will be going through a subtle increase in density as impacted fracture, widening of the wrist bone joint, and the surface that is immediately adjacent is disrupted. The LA view check will be analysed by several self-asked questions such as the conditions of radial articular surface intact, the smoothness of the *dorsal cortex* of the *distal radius*, bone fragment lying posterior to the carpal bones and correct position of a cup of the *Lunate* where the bone placed. The *Scaphoid* series check will analyse two standards of the wrist views as *Scaphoid* fractures, mostly lucent and thin-line fractures, not sclerotic. This check also examines the *Scaphoid* intact, as if the *distal radius* looks pristine [109]. The patient's grip and forearm strength are assessed, and if needed splint may be fitted to the wrist.

The manual procedure of wrist treatment therapy varies based on the type, location, severity of the injury, age, and overall health status. A physical therapist can implement specific treatments and exercises for the injury. After the operation procedure, a physical therapist manages the rehabilitation process. A wrist brace or splint can help hold the bone fragments together during the healing process. After treatment, the patient wears the splint to avoid sprained or strained wrist. The splint should be left in place until it has completely healed, usually under the supervision of a doctor or physiotherapist.

Wrist fractures treatments

Treatment of fractures has not changed much in time, as Egyptian records describe the re-positioning of bones, fixing with wood, and covering with linen [34]. Incidents such as falling, slipping, and tripping may lead to fractures that sometimes are

ignored by patients who do not seek treatment [109]. Wrist fractures are a common injury, especially among older patients, in the United Kingdom and worldwide [29, 134, 143, 32, 83]. The bone commonly involved is the radius in its distal region, and these fractures are called Distal Radius Fractures (DRF) [34].

Distal radius fractures are widespread in older females [29, 44] and are related to *osteoporosis* [142, 101]. This type of fracture is also considered as a pathway to a high risk of poor functional recovery that would prevent the daily independent life of the patients [134]. In general, clinicians developed best-practice and standardised procedures to handle incidents, including wrist fractures. Clinicians, including paramedics in the Accident and Emergency room, are trained to diagnose and determine the best treatment they can provide to handle patients [34, 83, 109]. The treatment also depends on the characteristics of each case where particular methods of treatment are selected [80].

Wrist injuries surgery was also examined based on the severity of the injury experienced by the patient. An example of research by [93] helps the surgeon with joint congruence, which does not prevent stiffness, the main complication of these injuries, which the surgeon must know how to recognise and treat. Visual examination before treatment with Computed Tomography (CT) also allows Orthopaedists to understand the lesion better and help the surgeon propose the best possible treatment for the patient. This study provides sufficiently detailed knowledge of how to deal with wrist injuries. Surgery by performing Total Wrist Arthroplasty for the treatment of wrist fractures is also carried out mainly in severe conditions of wrist trauma [86]. Arthroplasty for wrist trauma is a relatively new technology, despite its use for hip, shoulder, and elbow surgery.

Not only paying attention to the variables inherent in patients such as age, gender, ethnicity, location of residence, and complexity level of injuries patients, but research from the medical side also looks at aspects of body weight on fracture sites. Research reported on [107], focusing on the relationship between fracture sites and obesity, especially in the male gender. In many studies, it is reported that the gender that often experiences bone injury problems is women. However, there was a connection between Body Mass Index and fracture sites in this study. The results of this study also indicate that obesity in men does not contribute to injury to the hip, wrist, forearm, pelvic but has an impact on increasing the risk of rib fractures.

Total Wrist Arthroplasty (TWA) is a surgical procedure by resectioning part of

the wrist and replacing it with artificial materials like metal implants as described in [57]. The wrist arthroplasty technique is used as an alternative or substitute if TWA fails as it is perceived as a safe, effective, and versatile surgical technique in the area [1]. Biaxial, Universal 2, and Re-Motion are several types of TWA [26]. Nevertheless, the TWA procedure still faces challenges with the complication component of losing Carpal and unstable joints [26, 57, 86]. Cement extraction techniques, stabilising *radius distal* with Cerclage wires, and further bone resection are several solutions to revise TWA [57]. Cerclage wires stabilise the radius during the treatment. The TWA procedure also faces the high-cost issue associated with complications and revisions [57]. The surgery procedure remains an option as it provides pain relief and motion recovery [26].

2.1.3 Manipulation under Anaesthesia

Overview

Manipulation under Anaesthesia (MuA) is an orthopaedic procedure performed to increase articular motion and relieve chronic pain and is minimally invasive [54]. MuA is an old but widely recognised procedure in musculoskeletal-based injured problems. It is used for treating acute and chronic musculoskeletal conditions with significant biomechanical dysfunction. The MuA belongs to a variety of therapies called conservative therapy which does not include a surgical procedure [56].

MuA procedures require a multidisciplinary team, including a chiropractor, but still lack standardisation and steps of post-treatment care [35]. A patient will be recommended for a MuA procedure based on the cause of their pain such as chronic recurrent pain, loosening, or unstable joint [54]. However, these criteria have not been thoroughly explored [35].

There are no widely accepted procedures on the standard for practice and therapy manipulation under anaesthesia. Doctors will consider this procedure based on comparing the case with other similar levels of evidence for support and making consensus [54]. The well-established procedure of MuA is based on generally accepted by an interdisciplinary team such as an anesthesiologist, an operating room nurse, and a qualified chiropractor. It is also widely accepted that the MuA phase is sedation, manipulative procedures, additional stretching or traction procedures, and no sedation in inpatient care [30].

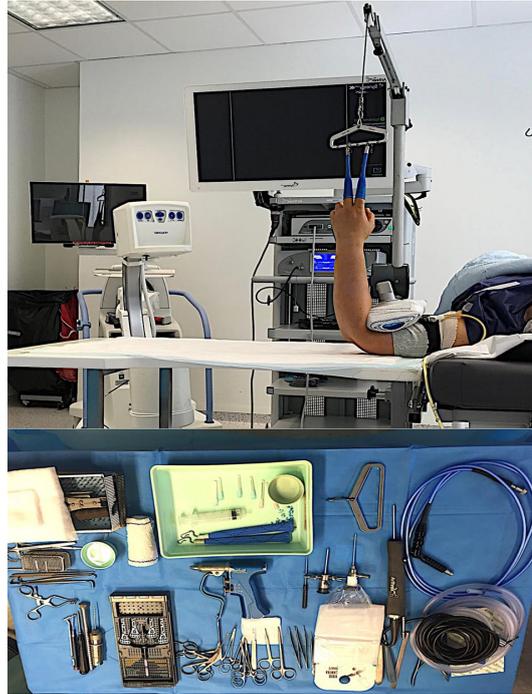


Figure 2.4: Operation room set-up for distal radius fracture including instruments for surgery [9]

Acceptable Patient's conditions for MuA

Gordon, Cremata, and Hawk [54] explain factors to qualify a patient for MuA clinical candidacy:

- The patient continues to have pain, a burden in daily activities, and/or biomechanical dysfunction after appropriate care such as spinal manipulation and medical treatment.
- Prior to recommending MuA, sufficient care within a minimum of 4 to 8 weeks should have been given to the patient. Nevertheless, this also depends on the condition of each patient. A sooner MuA procedure is considered for patients with more severe symptoms or little response to surgical procedures.
- Patients have undergone physical medical treatment procedures for 6 to 8 weeks before recommending MuA.
- The pain is intolerable for the patient that interferes with or even disabled personal daily living activities.
- Conditions of diagnosis must fall within the recognised condition categories for MuA.

Recognised condition categories for MuA's utilisation

The following disorders are classified as acceptable conditions for utilisation of MUA [54]:

- MuA is a treatment of choice, the patient chooses MuA treatment when the pain is personally intolerable which hinders the effectiveness of conservative manipulation.
- Patients choose the MuA when during a minimum of 4 to 8 weeks of care after a surgical procedure, the pain still exists and a greater degree of movement of the affected joint(s) is needed to obtain patient progress.
- The doctors chose the MuA due to the severity of the problem and the unsatisfactory healing progress of the patient.
- MuA is only being used as an alternative and/or an interim treatment when the patient is considered for surgical intervention and may be used as a therapeutic or diagnostic tool in the overall patient's condition.
- Based on the doctor's opinion, MuA being utilised when there are no better treatment options available for the patients.

MuA Procedures

Three stages of MuA procedure are patient sedation, chiropractic adjustments in distinct positions, and passive traction and stretching of specified bones [27]. The set of bone maneuvers explicitly determined for each patient before execution depends on the region of pain and range area of decreased motion [27]. Typically, MuA procedures are repeated over three days, and monitored by using X-ray images taken before and after treatment. The first LA view X-ray image was used to specify the fixation location, and the second image was taken after the second day of MuA to observe improvements. These two images allow the physicians to adjust the treatment approaches after executing a set of treatments [27]. Several research articles studied the MuA procedure for patient care [54, 35, 27].

Protocols and standards of MuA were developed by [54] based on evidence-informed and consensus-based guidelines. This guideline is intended for practitioners. A course of action decided and adopted made by trained practitioners based on resource availability and patient needs.



Figure 2.5: Situation inside the operation room where manipulation under anaesthesia performed [115]

Determining MuA progress

This procedure's success can be seen from subjective visual changes, level of pain, patient's ability to engage in active range of motion, changes in daily life activities, change in mass function, and strength. A fracture may slip or not; if it does not slip, then the MuA procedure was successful. The procedure's success can also determine using diagnostic tools such as X-rays, CT, and MRI. The following items are the parameters needed but not limited to seeing the success of MuA [54].

2.2 Computational background

2.2.1 X-ray image

X-ray was found as a form of radiation that led to a significant shift in the medical imaging field when it was discovered in the late 19th century [52] by German physicist Wilhelm Conrad Röntgen. The image shown in Figure 2.6 is believed as the first acquired X-ray image. This image was part of the observation publication he mailed to his colleagues at the end of 1895. This image is believed as his wife's left hand. An X-ray image is acquired by a beam of X-rays transmitted through the part of the body scanned. The light absorbed by the scanned body part and the light that is not passed through X-ray film in specific amounts of density is recorded on the X-ray film [13].

Although the rapid development of new technologies such as ultrasound imaging, magnetic resonance imaging, computed tomography, and plain X-rays remain a vital

tool for the radiologist and the cheapest option [52].

In digital radiography, digital images are obtained by digitising X-ray film or passing the X-rays through the patient and captured by a light-sensitive digitising system, as illustrated in Figure 2.8. The acquisition of the image starts when the energy flows from the generator to the output of the system which is the X-ray image. Energy moved from an X-ray tube where the electrons are converted in it and pass through the patient on the way to the image intensifier. X-ray is accelerated to the output *phosphor* where previously it was converted to light and then to electrons. The result of the image intensifier then catches by devices like camera and video camera to produce visualisation of an X-ray image.



Figure 2.6: The X-ray image of Wilhelm Conrad Röntgen wife's left hand. He took this image on 22 December 1895 and believed as the first X-ray image. Image is taken from [15].

Röntgen's X-ray was not the only X-rays-related article published within a year of discovery; at least four English texts dealt with this, as mentioned in [10]. The articles are The ABC of the X Rays by William Henry Meadowcroft [82], Practical Radiography by Henry Snowden Ward [138], Röntgen Rays and Phenomena of the Anode and Cathode by Edward Pruden Thomson [131], and The X-ray of Photography of the Invisible and Its Value in Surgery by William James Morton [87]. The latter article is the only text authored by a physician and probably an early medical article that directly pointed to the utilisation of X-rays for surgical procedures [10]. The patient usually lies down under the X-ray tube and is set in the position needed for the intended part of body being acquired using the X-ray machine, as illustrated in Figure 2.7.

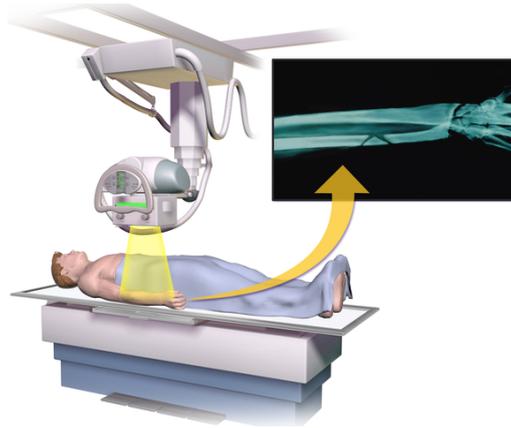


Figure 2.7: X-ray machine illustration where the patient lay down under the X-ray tube. This image is credited to [16].

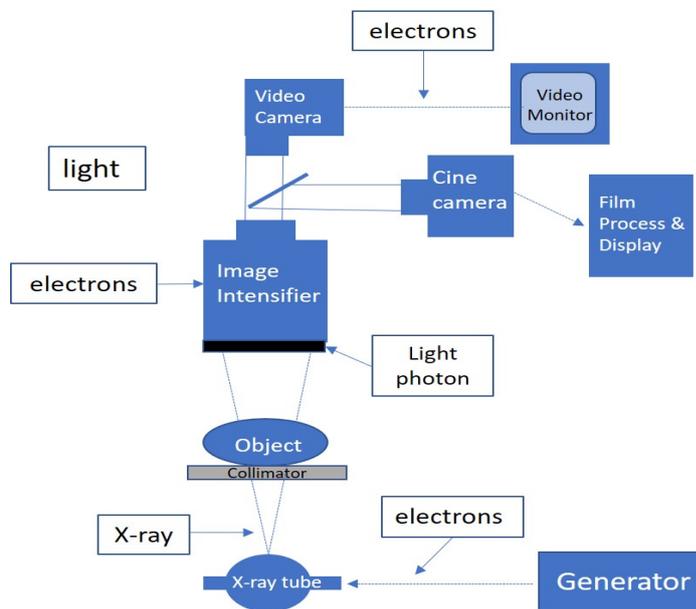


Figure 2.8: Basic X-ray imaging system illustration adapted from [13].

2.2.2 X-ray image of Wrist

An X-ray image of the wrist is a type of X-ray image taken from the hand part of the body. This image was usually taken by putting the to-be observed hand above the *collimator*, illustrated in Figure 2.9 above the X-ray beam tube.

DICOM standard is applied to the acquired X-ray image. It is an abbreviation for Digital Imaging and Communications in Medicine. This type of image has standardised metadata [90]. A DICOM file consists of information on the header and image data. The header of the DICOM file is organised as standardised series of attributes. An example of several attributes of the DICOM file is shown in table 2.1.

Attributes in Table 2.1 are some of the metadata attributes found in an im-

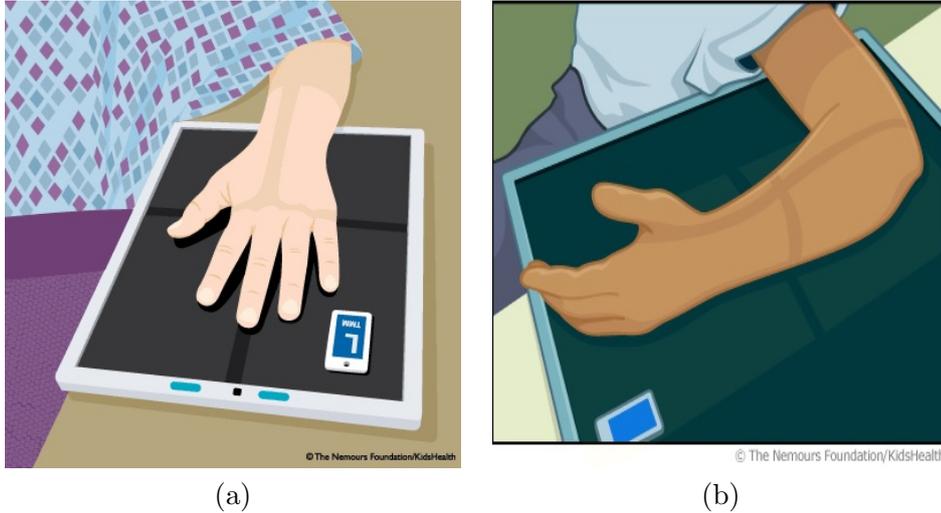


Figure 2.9: The illustration cartoon of (a) Posterior-Anterior (PA) and (b) Lateral (LA) of the wrist was scanned using an X-ray machine. These two images are taken from [38].

age with the DICOM standard applied. Complete details can be further explored in [90]. The currently identified metadata attributes that are important in this research are `BodyPartExamined`, `SeriesDescription`, `ViewPosition`, `PatientSex`, and `PhotometricInterpretation`. The value of these attributes described the body position when the image was acquired. The ‘sex’ is the identification of which gender each image belongs to and the background contrast of the image. These metadata attributes can be used in image processing, especially when we try to identify the object inside the X-ray image.

2.3 Literature review

2.3.1 Imaging Technology for wrist fracture problems

Clinical factors to predict incident fracture by a performed algorithm that uses clinical risk factors, patients’ bone mineral density, and fall history [40]. Falls and fractures are two factors usually perceived as causes of mortality in adults. [6] has explored this assumption where a systematic approach to screening and prevention is proposed.

In the medical field, problems related to bone fragility have been investigated in prescribing oral anti-*osteoporosis* drugs. The issue was studied by [132] where they covered the area of research in the United Kingdom between 1990 and 2012. Research that looked at the effect of giving this drug on the impact of bone fragility that led to bone fracture was distinguished based on variations in age, sex, geograph-

ical location, and ethnicity. Research by van der Velde et al. also shows that women experience susceptibility to *osteoporosis* compared to men. Furthermore, the same research shows that the process of bone fragility is suspected to be more likely to occur in the female sex, elderly (over 85 years), and are of white ethnicity followed by South Asian ethnicity. At the same time, other variables such as the location of residence do not influence the administration of this drug.

Inventions in medical imaging such as X-rays, Computed Tomography scans, Magnetic Resonance Imaging, and various other forms of medical imaging have provided much progress toward treatment solutions for bone-related problems, especially in this case, wrist injuries.

As it aims to restore the normal conditions of the patient, imaging technology was used as a pre-operative assessment to enable the best possibility of understanding the fracture as reported by [9]. *Scapholunate* is a ligament of the wrist, located in the wrist joint and it is very important for carpal stability. Ardouin et al. explored the Arthroscopy technique as the standard to identify and grade *scapholunate* injuries. Imaging technology is also adapted to manage the increasingly complex trauma of the wrist fracture with more critical functional requirements.

Research conducted by [89] tries to provide a solution by using computer vision to distinguish between *osteoporosis* patients and healthy patients. X-ray technology is used to visually diagnose abnormalities in the body and as a material for the utilization of more advanced X-ray images. Images can be used as material for image processing. Images can be used as data sets to implement artificial intelligence methods such as neural networks and other machine learning algorithms. Diagnosing *osteoporosis* by utilising machine learning methods such as deep learning technology to obtain information from images at pixel intensity levels is an example of utilising images in the data set. This research has a challenge with the small data set size for deep learning.

The new technology approach has been explored through deep learning. Fully automated learning from the X-ray data of bone proposed by [75] showed the robustness of the deep learning technology to measure bone age. They studied how to perform automated Bone Age Assessment (BAA) by distinguished regions of interest, standardise and preprocess input radiographs into the convolutional neural network (CNN) pipeline.

2.3.2 Medical Image Analysis

An X-ray image is part of the solution to observe the inner body condition before taking further steps toward the patient. Clear interpretation through a structured, systematic approach to describing fractures is vital for clinicians, including practitioners, to correlate clinical perspective with fracture identification on a plain radiograph [18]. Large amounts of clinical data have been generated and increasingly used for Biomedical and Health-care Informatics research [77, 74, 144]. Clinical benefits from vast patient data for making intelligent decisions and better treatment. Imaging in Informatics could adapt by integrating cloud computing technology to make paramedic work more ubiquitous [77].

Medical image analysis is the systematic evaluation of data with an algorithmic approach, which is similar to evaluating a human expert such as a radiologist, medical doctor, or paramedic. A medical image such as X-ray imaging technology is a long-time favourite by a radiologist to observe a patient’s medical problem, including a fracture in the wrist bone and another part of body organs [36]. Imaging technology like X-rays image is widely used and central support for doctors to diagnose patients [89, 2, 75].

Image segmentation for medical image analysis is not just for human cases but also has been applied to the veterinary area, such as diagnosis X-ray images of a mouse [3, 97]. In these two works, medical image processing techniques have been implemented for skeletal phenotype where precise segmentation and quick counting ribs are essential. The accurate identification of any perturbation in skeletal structure is significant biological relevance.

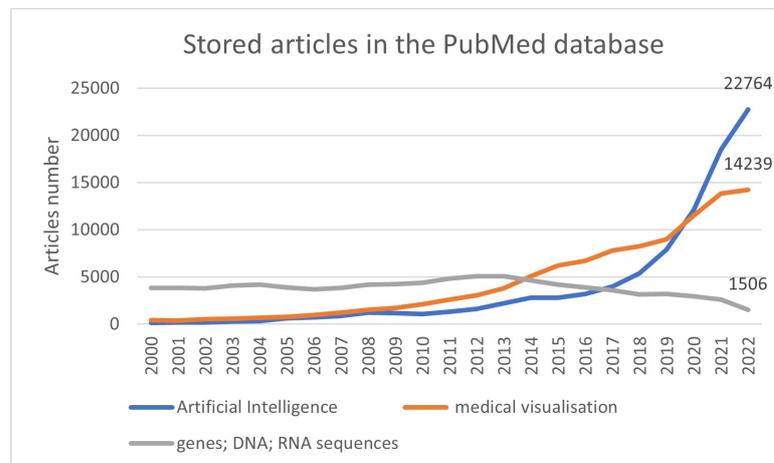


Figure 2.10: A diagram shows a comparison between certain topics of articles stored in the PubMed database from 2000 to 2022.

Analysis of medical problems is always an intriguing issue. It is because many parameters and conditions should be considered before deciding on further treatment for the patient. Rapid technological developments and medical problems always collaborate to tackle everyday challenges faced by people from the medical field. Medical problems always start from symptoms and are solved by the sequence of procedures. This type of work is aligned with the development of computer algorithms, where every solution came from procedure sequences. Combining these two will provide a more precise choice of treatment and a higher prediction of recovery success. Not only in terms of treatment for patients but also could build up preventive actions before disease or problems appear.

Artificial intelligence is a field in the technological world that has been heavily utilised in solving many medical tasks. The computational ability to mimic human intelligence, rapidly fast learning, and predict and reduce errors in diagnosing has tremendous contributions and benefits to medical fields. In the academic world, as shown in Figure 2.10, this condition is reflected in the increase of published articles where they put artificial intelligence and medical problems as the centre front to answer their research problem.

Big data technologies are used increasingly for research in biomedical engineering and informatics. Large amounts of biological and clinical data have been generated and collected at an unprecedented speed and scale [77].

Artificial intelligence based on extensive data sources can be utilised through this. The world of medical practice has begun to apply it in Biomedical Engineering, such as neurological diseases [122], to a virtual and physical branch of the medic [58]. Most applied along to understand the medical data that hugely produced [74, 118]. The use of X-ray films is still widely used because this technology is stable, inexpensive, and available in many hospitals and healthcare centres. Nevertheless, due to the frequent conditions of misdiagnosis, because the images are inaccurate due to some circumstances, the results of the acquired image were useless. Big data, Deep Learning, and a variety of other software algorithm techniques exploit this [77].

Table 2.1: This table is an example of metadata attributes from the DICOM image. These attributes are standard for communicating and managing medical imaging information and related data [90].

No	Attributes	Value
1	Filename	IMG0
2	FileModDate	'12-Oct-2018 17:17:19'
3	FileSize	1567164
4	Format	'DICOM'
5	Width	608
6	Height	1287
7	BitDepth	12
8	ColorType	'grayscale'
9	AcquisitionDate	'20160811'
10	AcquisitionTime	184202
11	Modality	CR
12	ManufacturerModelName	Philips Medical Systems'
13	StudyDescription	XR Wrist Lt
14	SeriesDescription	Lateral
15	InstitutionalDepartmentName	Accident & Emergency'
16	ManufacturerModelName	DigitalDiagnost
17	PatientName	[1×1 struct]
18	PatientID	ANON9731'
19	PatientBirthDate	20100713
20	PatientSex	F'
21	PatientAge	006Y'
22	PatientSize	0
23	PatientWeight	0
24	PregnancyStatus	4
25	BodyPartExamined	HAND'
26	ProtocolName	Wrist L'
28	SpatialResolution	0.1440
29	ExposureTime	10
30	ImagerPixelSpacing	[2×1 double]
31	Grid	NONE'
32	ViewPosition	'LL'
33	PatientOrientation	A\H'
34	Laterality	L'
35	SamplesPerPixel	1
36	PhotometricInterpretation	MONOCHROME2'
37	Rows	1287
38	Columns	608
39	RequestedProcedureDescription	'XR Wrist Lt'
40	RequestedProcedureCodeSequence	[1×1 struct]
41	PerformedProcedureStepDescription	XR Wrist Lt

Chapter 3

Materials

In this chapter, detailed information about image materials used in this research is explained. The first subsection explains the University of Exeter’s DICOM Wrist data set and the second subsection is a detailed explanation of the Musculoskeletal Radiographs data set from the Stanford Machine Learning Group. These two data sets are the data sets used in the thesis.

3.1 University of Exeter’s DICOM wrist data set

This data set is acquired at the University of Exeter. The data consisted of 1,007 X-rays, 884 from wrist fractures, and 123 from healthy controls funded by The Royal Devon and Exeter Small Grant Scheme:1802595.

The Normals folder consists of two folders named N1 and N2. The N1 folder contains thirty folders named ‘PATx’, where x is the number of anonym patients. This naming pattern is also considered for the second folder named N2 which contains forty patient folders. Each folder has the same pattern of data-filled folder naming which is ‘STDx’, where x is the number of procedures applied to the patients. The ‘STDx’ folder contains one or more folders named ‘SERx’ where x is the number of the SER applied. The ‘SER’ is a series label of an image in a specific position. An identical labelling pattern was also applied to the N2 folder.

The second major category defined by the expert in this data set is the Patients folder. This folder contains collections of DICOM images identified as abnormal patients. This condition means that the expert assumed fractures or any other abnormalities in an X-ray image. This folder also contains a diverse variation pattern of folder naming, but generally, the name ‘Txx Tranche X’ was used as the folder

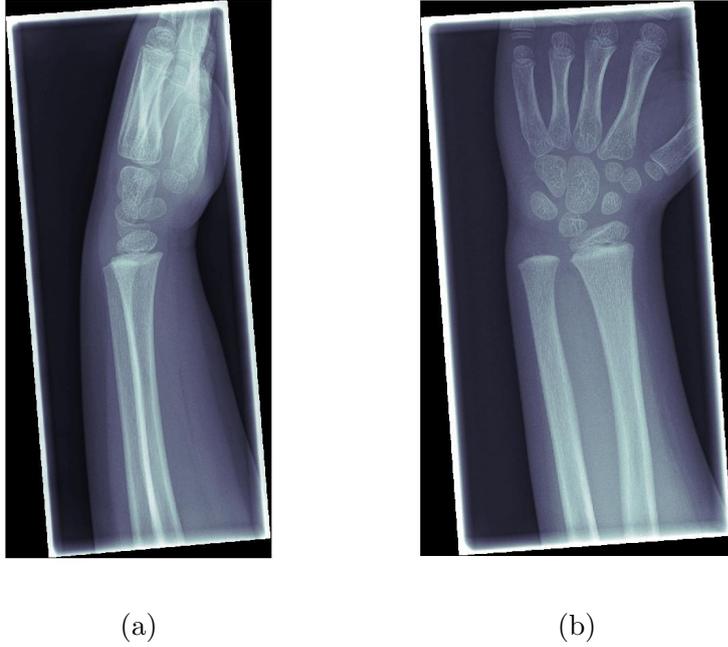


Figure 3.1: These two X-ray images are examples of normal wrist X-rays images from The University of Exeter data set. Image (a) is a type of X-rays position named Lateral. Image (b) is the type of X-ray taken in the Posterior-Anterior(PA) position.

name pattern. Each ‘Txx Tranche X’ folder consists of the same pattern as ‘PATx’ of the Normals folder. There are fifteen of these folders, and each folder contains several folders of anonym patient image folder. The rest of the folder naming pattern is the same as the Normals folder.

Two folders in ‘Patients’ have annotated images. A folder named ‘T01 Tranche 1’ contains 23 annotated Wrist X-ray images and 31 wrist X-ray images in a folder named ‘T02 Tranche 2’. The data set of images in a folder named Normals does not have images with annotations. Those are the conditions of the data set received.

The Annotations folder contains images in the JPG image file format that had previously been manually annotated for fractures or abnormalities. However, not all images in the Patients’ data set had been annotated. There were no further descriptions of these annotated images provided. Based on all the annotated X-ray images, it seems that the location of the fractures was identified in the X-ray image. Annotated X-ray images also had changed to the JPG file type rather than their original in the DICOM file type.

The DICOM’s X-ray images in these two folders are raw images. The patient’s name was the only information that had been removed for anonymisation purposes. The DICOM images have the identical pattern of names, which are ‘IMGx’, and

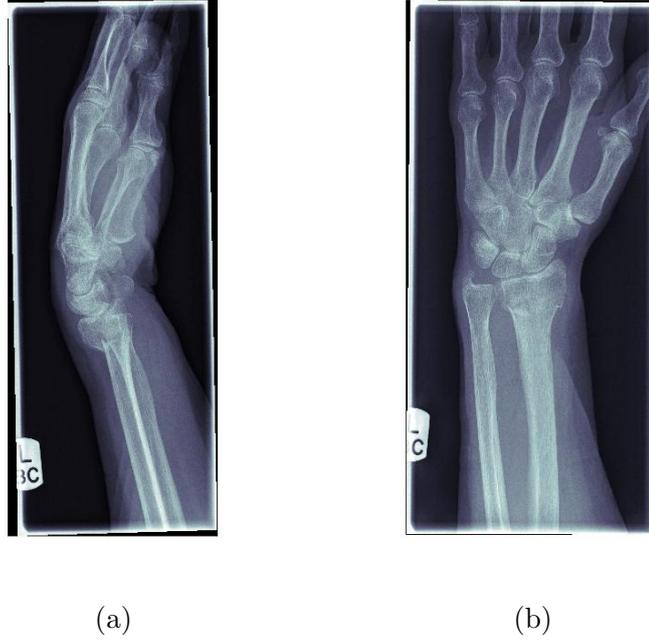


Figure 3.2: Image (a) is a type of X-rays position taken used to identify abnormalities of the bone in the lateral taken position. Image (b) is the type of X-ray taken in the Posterior-Anterior (PA) position. These are wrist X-rays images with fractures identified from The University of Exeter data set.

on average, each ‘SERx’ contains one ‘IMGx’ DICOM file where x stands as an expression of number, e.g. IMG1, IMG2, etc. The size of each DICOM file is at least 2 MB.

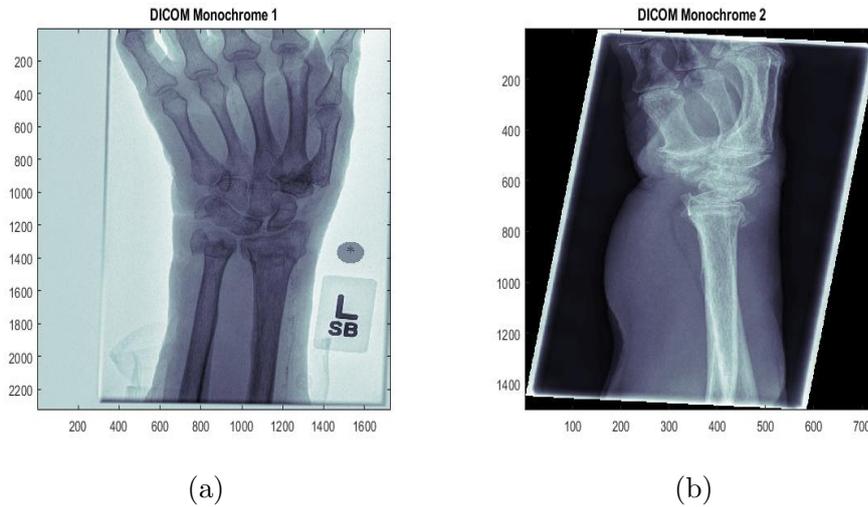


Figure 3.3: These are examples of wrist X-rays images with different values of Photometric Interpretation from The University of Exeter data set. Photometric interpretation is one of the standard attribute metadata of DICOM. Image (a) has a monochrome value of 1, where intended to be displayed on white background. Image (b) has a monochrome value of 2, where intended to be displayed in black background [65]

After manual inspections of the 1,005 images, 681 images were selected for further

processing. Wrist X-ray images were selected based on a single position in an image, excluding the child’s wrist and non-wrist images. These criteria were used because the research only focused on the wrist. All data is anonymous and approved to be used. Figure 3.1 shows the example of a normal X-ray image of a wrist. The image is from the data set. Images in Figure 3.1 (a) are a type of X-ray image position taken to identify abnormalities of the bone in the Lateral taken position. From this position, the radiologist will observe the intact of the *radial* articular surface, the smoothness of the *dorsal cortex*, and the sitting bone still in the cup of the *Lunate*. The type of X-rays taken in the *Posterior-Anterior* (PA) position as shown in Figure 3.1 (b). Using the Lateral position, the radiologist will observe the whole intact of the *radial articular surface* and the *ulna styloid*.

Photometric Interpretation is one of the attribute metadata of DICOM. X-ray images in this data set came in one of two values of *Photometric Interpretation*. Figure 3.3 (a) has a monochrome value of 1, intended to be displayed on white background. Image 3.3 (b) has a monochrome value of 2, intended to be displayed on a black background. These two types of images in different values of *Photometric Interpretation* are shown in Figure 3.3. In this data set, several data have been annotated by the Radiologist. The wrist X-ray images have been marked as identification to annotated which means the image contains a problem of interest. In this case, the Radiologist annotated with a circle to point out the fracture’s location based on their image observation. The example of the expert’s annotated images is shown in Figure 3.4.

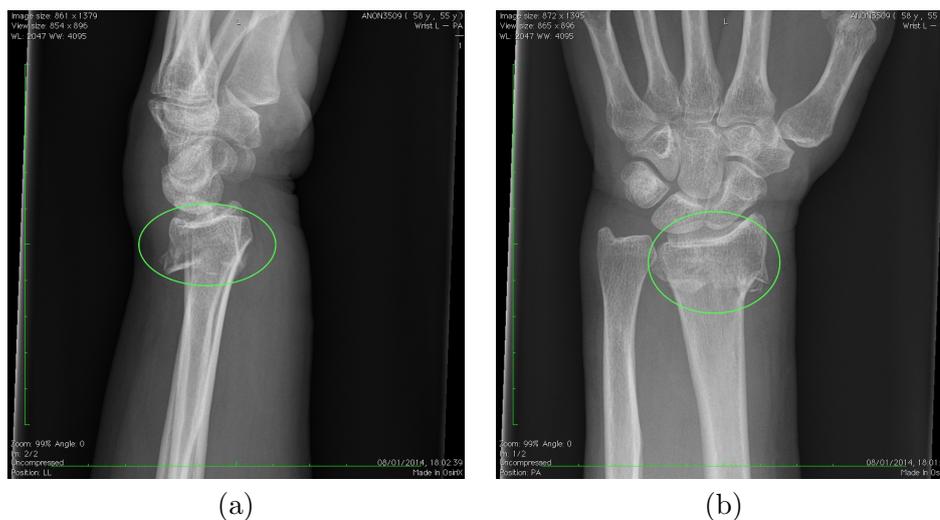


Figure 3.4: These are annotated wrist X-ray images with fractures identified from The University of Exeter data set. A green circle is an abnormality annotated by a radiologist.

Table 3.1: This table shows the exact number of data from Exeter’s data set, which has a value of the PatientSex metadata attribute.

PatientSex	Normals	Patients
Females	43	423
Males	27	50

Table 3.2: This table shows the exact number of data from Exeter’s data set with a value of SeriesDescription, ViewPosition metadata attribute related to Lateral and Posterior-Anterior which are focus positions of this work.

X-ray positions	Normals	Patients
Lateral	63	272
Posterior-Anterior	45	284

Based on the gender attribute values contained in the DICOM image in this data set, in general, the amount of data coming from the female gender is greater than the male gender. The values shown in Table 3.1 are the exact values contained in the image files in the data set. There are 43 data with female sex in the Normals folder, and 27 data have male gender attribute values. The largest data was in the Patients folder with 423 female sex attribute values compared with 50 male data. The total amount of data that can be detected based on the gender attribute value is 543 data.

Another attribute used for this study related to the position of object acquisition. This position is obtained from reviewing the values in the DICOM attribute related to the Posterior-Anterior and Lateral positions. Table 3.2 shows the exact values of the amount of data contained in the data set related to this. The total data that can be detected based on the values of this attribute is 664 data, divided into data contained in the Normals folder and the Patients folder. Table 3.2 shows that the amount of combination data between Normals and Patients with the Lateral position is larger than the data with the Posterior-Anterior position. Nevertheless, with a deeper look, there is a difference where the amount of Lateral data in the Normals folder is greater than the data with Posterior-Anterior position attribute values; the inverse thing occurs in the Patients folder.

Photometric Interpretation is a DICOM attribute that is also used in this study. Based on the detection of the value contained in the DICOM image owned in the data set, we get the data in Table 3.3. The combined total data from the Normals and Patients folder shows that most of the data in this data set are Monochrome 1 compared to Monochrome 2. This number is based solely on the values detected

Table 3.3: This table shows the exact number of data from Exeter’s data set that has a value of the PhotometricInterpretation metadata attribute.

Photometric Interpretations	Normals	Patients
Monochrome 1	24	331
Monochrome 2	46	143

from the Photometric Interpretation attribute in the DICOM file.

3.2 Stanford Machine Learning Group’s MURA data set

The data set has been shared publicly by the Stanford Machine Learning Group named Musculoskeletal Radiographs (MURA) [111]. This data set consists of 40,561 images from 14,863 studies, where radiologists manually label each study as either normal or abnormal. To evaluate models robustly and to get an estimate of radiologist performance, they collect additional labels from six board-certified Stanford radiologists on the test set of 207 musculoskeletal studies, a majority vote of a group of three radiologists serves as the gold standard [111]. It is not stated explicitly in Rajpurkar et al. about what happened to the image if six annotators disagree. The table represented the distribution of X-ray images based on upper body human anatomy from the MURA data set is shown in Table 3.4.

Table 3.4: Distribution of cases of the Stanford MURA (musculoskeletal radiographs) data set [111] for studies of the upper body.

No.	Study	Train		Validation		Total
		Normal	Abnormal	Normal	Abnormal	
1	Elbow	1,094	660	92	66	1,912
2	Finger	1,280	655	92	83	2,110
3	Hand	1,497	521	101	66	2,185
4	Humerus	321	271	68	67	727
5	Forearm	590	287	69	64	1,010
6	Shoulder	1,364	1,457	99	95	3,015
7	Wrist	2,134	1,326	140	97	3697
	Total	8,280	5,177	661	538	14,656

MURA data set consists of two folders named Train and Valid. The train folder is a collection of musculoskeletal X-ray images from parts of the Elbow, Finger, Forearm, Hand, Humerus, Shoulder, and Wrist. The Valid folder also uses similar structures. This data set also has trained and valid files of labelled images. The

train and valid path of the image also provided this data set.

The MURA data set is divided into several image folders that are parts of the body. The total data in this data set is 40,005 images divided into 36,808 images in the Train folder and 3,197 in the Valid folder. The largest data held by this data set is Wrist data, where there are 9,752 images in the Train Folder and 659 data in the Valid folder as shown in Table 3.5.

Table 3.6 shows the exact amount of Normal and Abnormal images identified in each part of the body from the MURA data set. Unlike the data shown in Table 3.5, in Table 3.6 the number of images displayed is based on the amount of data with Abnormal and Normal status. There is a combination of data from the Train and Valid folders in each condition. Each body image, data consists of a combination of Train and Valid data then differentiated into Abnormal and Normal conditions. The total number of images with Abnormal conditions in this MURA data set is 16,403. Normal conditions are 23,602, so the total image in the data set based on this condition is 40,005, equal to the total number of images contained in Table 3.5. The Wrist images remain the most significant amount of data identified based on this condition. The amount of data in the Normal condition is larger than in the Abnormal condition. In the Wrist data set, the amount of Abnormal data detected was 4,282, and the data with normal conditions was 6,129. The Shoulder data set is the second largest with 8,942 images, and the Finger data set is in the third-largest position with 5,567 images.

Table 3.5: Distribution of images in MURA (musculoskeletal radiographs) data set for training and validation.

No.	BodyParts	Number Images in Train Folder	Number Images in Valid Folder
1	Elbow	4,931	465
2	Finger	5,106	461
3	ForeArm	1,825	301
4	Hand	5,543	460
5	Humerus	1,272	288
6	Shoulder	8,379	563
7	Wrist	9,752	659
Total per condition		36,808	3,197
Total actual images in data set		40,005	

An example of images labelled as Train positive is shown in Figure 3.5. The

image labelled as a Train positive is based on the training process by the MURA team shows the image is an X-ray with an abnormal condition. Figure 3.5 (a) shows abnormality detected in the elbow X-ray image. Abnormality is also detected in the X-ray image of fingers, as shown in Figure 3.5 (b). Figure 3.5 (c) is an example of abnormal forearm labelled as ‘Train positive’. As shown in Figure 3.5 (d), the X-ray image of fingers is identified as a Normal image, but it is not. The *humerus* X-ray image in Figure 3.5 (e) is an example of a Train-positive image that is not straightly aligned during image acquisition. X-ray image with high contrast is also set as Train positive in MURA data set as its example in the shoulder image of Figure 3.5 (f). An image with a metal plate is labelled as Train positive in this data set, as seen in Figure 3.5 (g) which is the wrist image. ‘Train-positive’ is also labelled to an image in low contrast like the Lateral position of wrist X-ray image in Figure 3.5 (h).

Table 3.6: Distribution of images in the Stanford MURA (MUsculoskeletal RAdiographs) data set into Abnormal and Normal groups. This work concentrated on the wrist radiographs.

No.	BodyParts	Abnormal (Train + Valid)	Normal (Train + Valid)
1	Elbow	2,236	3,160
2	Finger	2,215	3,352
3	ForeArm	812	1,314
4	Hand	1,673	4,330
5	Humerus	739	821
6	Shoulder	4,446	4,496
7	Wrist	4,282	6,129
Total per Condition		16,403	23,602
Images total in data set		40,005	

MURA data set examples of valid positive are shown in Figure 3.6. A valid positive is an abnormal detected image with labeled approved by the MURA team of experts. Most of the images in Figure 3.6 show any obvious abnormality. The Elbow image in Figure 3.6 (a) shows a clearly visible fracture as well as in Figure 3.6 (b-e). The wrist images of Figure 3.6 (g) and (h) were taken in two different positions where the fracture can be seen as slightly prominent and these images labelled as valid positive by the MURA’s experts.

The negative result of training labelled as Train negative image in the MURA data set. Images in this label contain only a normal-condition image assumed for the training process. The collage images in Figure 3.7 (a-h) contain X-ray images of

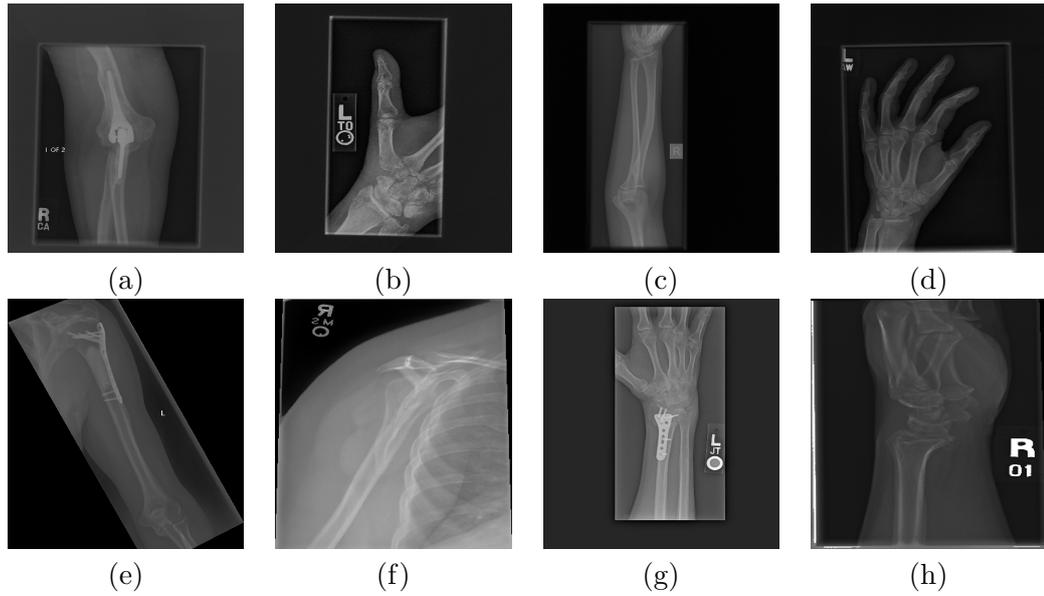


Figure 3.5: Examples of Train positive X-ray images from MURA data set [111]. These are images that are positively labelled and have an abnormality in them. These images are part of the images in the Train folder by MURA. The subfigures show train-positive X-ray images of the Elbow (a), Fingers (b), Forearm (c), Hand (d), Humerus (e), Shoulder (f), and Wrist in PA position (g), and Wrist in Lateral position (h).

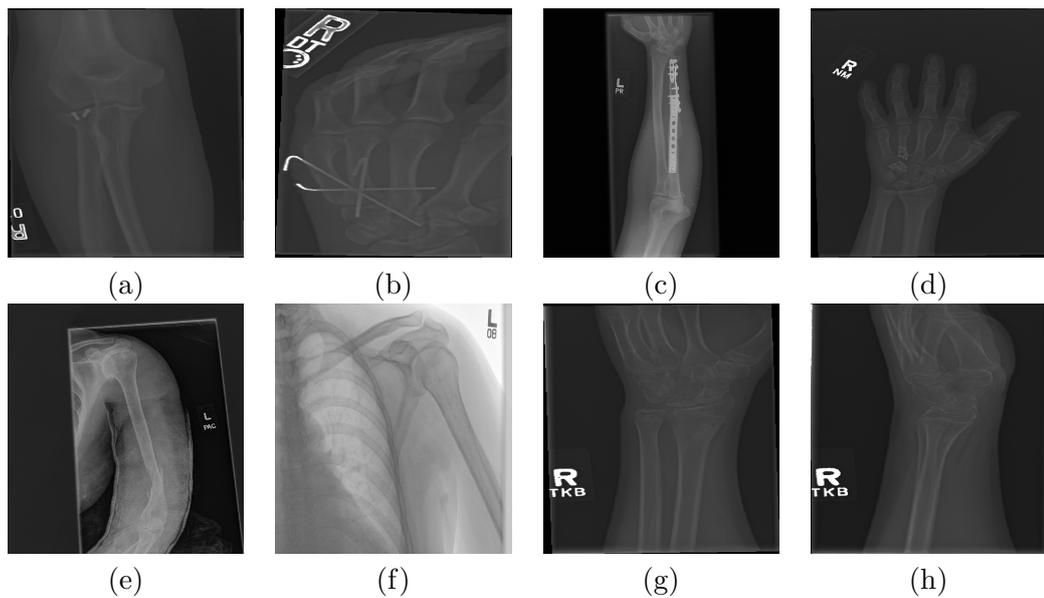


Figure 3.6: Examples of valid positive X-ray images from MURA data set [111] labelled by MURA team experts. Example of valid positive X-ray image of Elbow (a), Fingers (b), Forearm (c), Hand (d), Humerus (e), Shoulder (f), Wrist in PA position (g), and Wrist in lateral position (h).

Elbow (a), Fingers (b), Forearm (c), Hand (d), *Humerus* (e), Shoulder and Posterior-Anterior (g) and Wrist on Lateral side positions (h).

The expert team of MURA had put a label of valid negative to an image that is diagnosed as a normal condition in their observation. Images in this label will

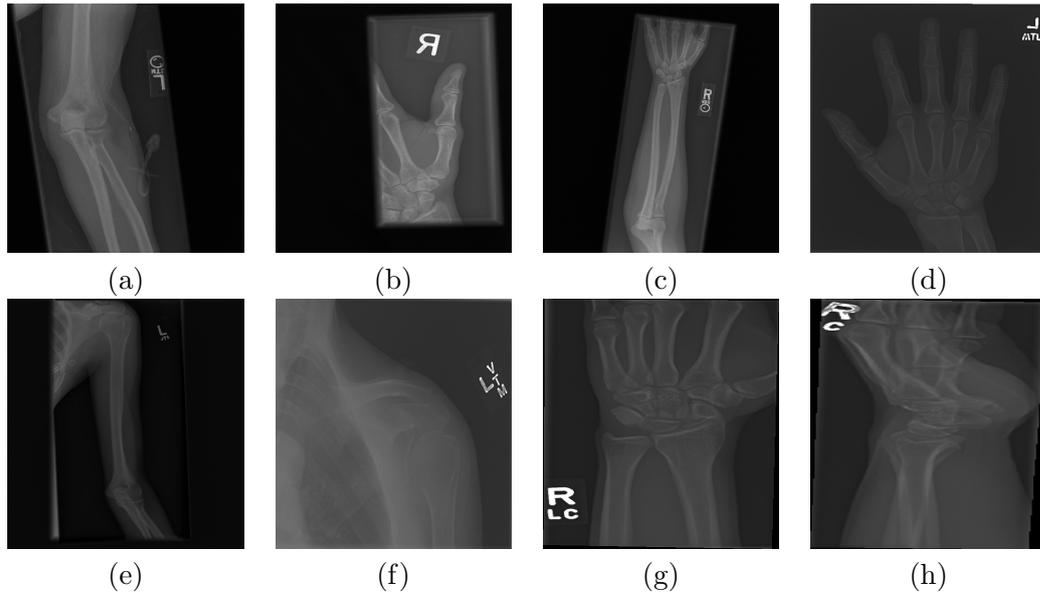


Figure 3.7: Examples of Train negative X-ray images from MURA data set. These images are labelled for the training process of normal data. The subfigures show train negative X-ray images of Elbow (a), Fingers (b), Forearm (c), Hand (d), Humerus (e), Shoulder (f), Wrist in PA position (g), and Wrist in Lateral position (h).

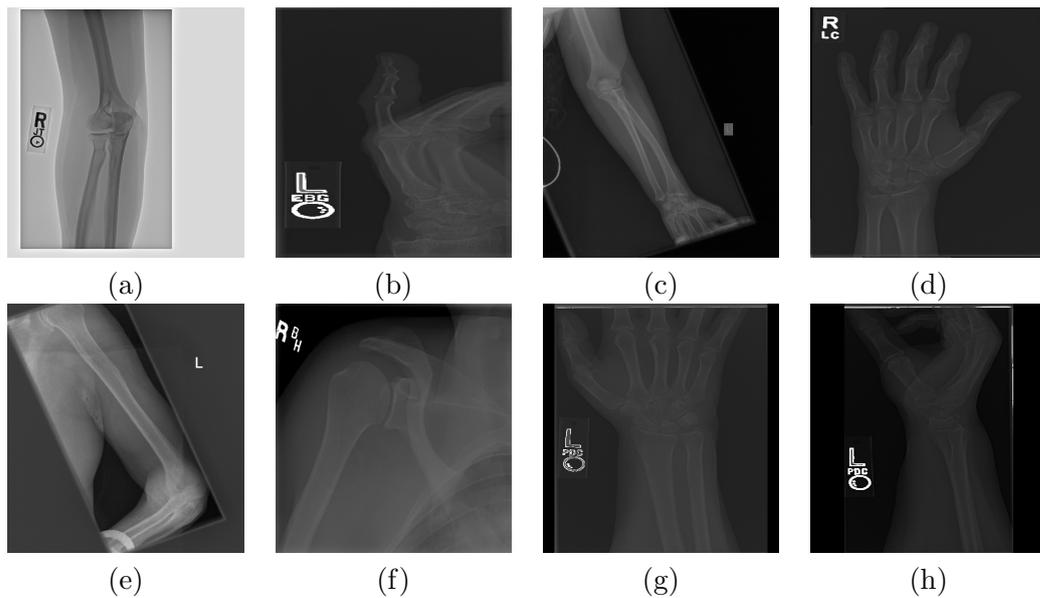


Figure 3.8: Examples of valid negative X-ray images from the MURA data set. The MURA's expert justified that these images in the data set are normal. Images in this label will be used as validity for the negative result of data training. The subfigures show valid negative X-ray images of the Elbow (a), Fingers (b), Forearm (c), Hand (d), Humerus (e), Shoulder (f), and Wrist in PA position (g), and Wrist in Lateral position (h).

confirm training results from data in the Train folder. The image for each part of the body in the data set can be seen in Figure 3.8(a-h). Visually, all of these images look normal to the MURA experts.

There are some repeating labels in this data set. The label is per image. The

Table 3.7: MURA’s Wrist train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Wrist Train data set	Positive (Abnormal)	Negative (Normal)
Study 1	3,920	5,282
Study 2	64	425
Study 3	3	45
Study 4	0	13
Total	3,987	5,765
Total Data (Positive + Negative)	9,752	

labels are Study 1, Study 2, Study 3, and Study 4. These labels correspond to each patient who may have visited the hospital several times, and each study corresponds to one visit. There are also labels related to anonymous subjects’ interpretation conditions: Positive means Abnormal and Negative means Normal. The interpretation of these labels is based on confirmation with the MURA team. No further information was provided by the MURA team regarding whether the fracture healed or not, for example, an image which has labels Study 4 whether the fracture healed or not. The fracture’s condition of healed or not healed is not the focus of this thesis.

Each label study is then broken down to determine how many Positive means Abnormal suspect conditions and Negative means Normal suspect condition data are contained in this data set. Table 3.7 and Table 3.8 show the amount of detailed data in the data set for the wrist parts of the body with differentiation based on the research label and the diagnosis conditions of the subjects obtained during the acquisition process.

The largest amount of data is obtained in Table 3.7 when anonymous subjects scan the body at the first visit. This is shown by the large amount of data identified in Study 1. The abnormal data identified at this first visit was 3,920, and normal diagnoses were found at 5,282 data. Furthermore, only a few anonymous subjects carry out the process of image acquisition data up to four times. The Study 4 label shows that no patients identified as being in an Abnormal condition, and all patients who took data at the fourth visit were all in Normal condition. The total image data placed in this Train folder is 9,752 images, consisting of 3,987 Abnormal data and 5,765 Normal data.

In line with Table 3.7, the data shown in Table 3.8 is a valid data set of the Wrist images. The total amount of data is not as large as the Train folder’s amount

Table 3.8: MURA’s Wrist valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Wrist Valid data set	Positive (Abnormal)	Negative (Normal)
Study 1	287	293
Study 2	5	59
Study 3	3	9
Study 4	0	3
Total	295	364
Total Data (Positive + Negative)	659	

Table 3.9: MURA’s Elbow train image data set breakdown based on the number of studies taken for the patient and condition detected on the image

Elbow Train data set	Positive (Abnormal)	Negative (Normal)
Study 1	1,993	2,818
Study 2	13	105
Study 3	0	2
Study 4	0	0
Total	2,006	2,925
Total Data (Positive + Negative)	4,931	

with only 659 data images. In this valid data set, the largest number remains in the ‘Study 1’ label with the total Abnormal data of 287 and the normal number of 293. The total number of the Train folder and the Valid folder is in line with the Wrist images data in Table 3.5.

The following detailed data is the Elbow train data set shown in Table 3.9 and the valid Elbow data set shown in Table 3.9. The Elbow train data set in Table 3.9 has a large amount of data in the Normal data of 2,818 and 1,993 Abnormal. Similar to the previous pattern in Table 3.7 and Table 3.8, the largest part is the data obtained at the first anonymous subjects visit the hospital for image scanning as it labels as ‘Study 1’. Total train data with Abnormal conditions is 2,006 images, and Normal is 2,925 images. In the Elbow valid data set, the data with the ‘Study 1’ label contained 228 Abnormal images and 222 Normal images. These tables do not have anonymous subjects who must visit the hospital for up to 4 visits, labelled as ‘Study 4’.

The collections of forearm images in the MURA data set have been divided into two: the Forearm Train and the Valid data set. Table 3.11 for the train data set and Table 3.12 for the Valid data set. These second tables also have the same pattern

Table 3.10: MURA’s Elbow valid image data set details the number of studies taken for the patient and condition detected on the image.

Elbow Valid data set	Positive (Abnormal)	Negative (Normal)
Study 1	228	222
Study 2	2	13
Study 3	0	0
Study 4	0	0
Total	230	235
Total Data (Positive + Negative)	465	

Table 3.11: MURA’s Forearm train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Forearm Train data set	Positive (Abnormal)	Negative (Normal)
Study 1	657	1,144
Study 2	4	20
Study 3	0	0
Study 4	0	0
Total	661	1,164
Total Data (Positive + Negative)	1,825	

as the data set of body parts that were previously owned. The forearm train data set has a total image of 1,825, and the valid forearm data set has 301 images. In the Forearm data set the number of Normal images is greater than the Abnormal, 1,144 versus 657. However, the inverse is seen in a valid data set with the number of images.

The following data set is the details for the data set shown in Table 3.13 and Table 3.14. These two tables are the details for the Finger body of the MURA data set. The train images data set is shown in Table 3.13 based on the amount of data detected on the labels contained in the data set. The largest data remained in the data labelled as Study 1 data, namely 1,949 abnormal images and 2,974 Normal images. There is no data obtained from Study 4, only 11 data images with Normal conditions in Study 3, and for Study 2, there are 19 Abnormal data images and 153 data images with Normal conditions. The total finger train data set data is 5106 data images.

Table 3.14 shows a smaller amount of data than in Finger’s training data set contained in Finger’s valid data set. The total data obtained is only 461 data images. The largest number of labels was still obtained in Study 1, with Positive

Table 3.12: MURA’s Forearm valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Forearm Valid data set	Positive (Abnormal)	Negative (Normal)
Study 1	149	144
Study 2	2	4
Study 3	0	2
Study 4	0	0
Total	151	150
Total Data (Positive + Negative)	301	

Table 3.13: MURA’s Finger train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Finger Train data set	Positive (Abnormal)	Negative (Normal)
Study 1	1,949	2,974
Study 2	19	153
Study 3	0	11
Study 4	0	0
Total	1,968	3,138
Total Data (Positive + Negative)	5,106	

data of 247 images and Negative data of 191. In the data labelled Study 2, there are no Abnormal data and 20 Normal data. Study 3 only has 3 data with Normal conditions and none with Abnormal conditions. In this valid data set for Finger, there is no data labelled Study 4.

MURA’s Finger valid image data set breakdown is based on the number of studies taken for the patient and the condition detected on the image. Details for the Hand part of the body are shown in Table 3.15 for the Train data set and Table 3.16 for the Valid data set. This follows the same pattern as the data sets described, from

Table 3.14: MURA’s Finger valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Finger Valid data set	Positive (Abnormal)	Negative (Normal)
Study 1	247	191
Study 2	0	20
Study 3	0	3
Study 4	0	0
Total	247	214
Total Data (Positive + Negative)	461	

previous parts of the body, where the most amount of data acquired labelled in Study 1. The total data of Hand images acquired for the Train data set is 5543 data images and for a valid data set is 460 data images. There are 3870 Normal data and 1,470 abnormal data labelled as Study 1. Fourteen data with Study 2 labelled as Abnormal conditions and 178 data with Normal conditions. In Study 3, there are no data with Abnormal conditions, and there are 11 data with Normal conditions. No data were obtained in Study 4 of this Hand train data set as well.

In the Hand valid data set, there are 189 data identified with Abnormal conditions and 248 data with Normal conditions. In this data set, only the Study 1 label has Abnormal conditions, while the data with the Study 2 label and Study 3 only have Normal data with a value of 20 data and 3 data. There is no data labelled Study 4 contained in this data set. A body part named Humerus is included in this data set as well. The detailed conditions of images identified by this body part are shown in Table 3.15 and Table 3.16. The total Humerus data in this data set is 1,560 images consisting of 1,272 images from the Train data set and 288 images from the Valid data set. In this data set, data is only found on Study 1 and Study 2 labels, whereas for Study 3 and Study 4, there is no data labelled between these two.

In this Humerus data set, most of the data are still dominated by data with Normal conditions and labelled as Study 1. There are 1,470 images with Positive abnormal conditions in the Train data set and 3,870 under normal conditions, both for the Study 1 label. The valid data set contains 138 images identified as positive and 145 as Negative conditions, both for the Study 1 label. Data with the Study 2 label contained in the training data set are two Abnormal images and nine images for Normal. There are two Abnormal images and three Normal images in the valid data set.

The last body part in this data set is the shoulder. This body part has the amount of data that includes a more significant part with the other body parts in this data set: as many as 8,942 images where 8,379 images are Train data, and 563 are Valid data. The Shoulder Train data set consists of 4,063 images and 3,930 images of Normal, and both were labelled as Study 1. For images that were labelled as Study 2 in this data set consist of 90 images of positive and 244 images negative. Study 3 in this data set consists of 10 positive images and 27 negative images. The Study 4 label obtained as many as five images with positive and ten images with

Table 3.15: MURA’s Hand train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Hand Train data set	Positive (Abnormal)	Negative (Normal)
Study 1	1,470	3,870
Study 2	14	178
Study 3	0	11
Study 4	0	0
Total	1,484	4,059
Total Data (Positive + Negative)	5,543	

Table 3.16: MURA’s Hand valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Hand Valid data set	Positive (Abnormal)	Negative (Normal)
Study 1	189	248
Study 2	0	20
Study 3	0	3
Study 4	0	0
Total	189	271
Total Data (Positive + Negative)	460	

negative conditions.

For the valid data set, there are images labelled as Study 1 where 261 images with positive conditions and 241 images with negative conditions. Study 2 labelled images for abnormal conditions in 17 images and 36 images for normal conditions. For Study 3 labelled images only can be found in eight suspect normal conditions and zero for abnormal conditions.

Table 3.17: MURA’s Humerus train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Humerus Train data set	Positive (Abnormal)	Negative (Normal)
Study 1	597	664
Study 2	2	9
Study 3	0	0
Study 4	0	0
Total	599	673
Total Data (Positive + Negative)	1,272	

Table 3.18: MURA's Humerus valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Humerus Valid data set	Positive (Abnormal)	Negative (Normal)
Study 1	138	145
Study 2	2	3
Study 3	0	0
Study 4	0	0
Total	140	148
Total Data (Positive + Negative)	288	

Table 3.19: MURA's Shoulder train image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Shoulder Train data set	Positive (Abnormal)	Negative (Normal)
Study 1	4,063	3,930
Study 2	90	244
Study 3	10	27
Study 4	5	10
Total	4,168	4,211
Total Data (Positive + Negative)	8,379	

Table 3.20: MURA's Shoulder valid image data set breakdown based on the number of studies taken for the patient and condition detected on the image.

Shoulder Valid data set	Positive (Abnormal)	Negative (Normal)
Study 1	261	241
Study 2	17	36
Study 3	0	8
Study 4	0	0
Total	278	285
Total Data (Positive + Negative)	563	

Chapter 4

Methods

This chapter explains the methods associated with this thesis research. Image processing techniques are explained from the concept of an image, geometric transformations, and image enhancement techniques to edge detection. A brief explanation of artificial intelligence is also covered, and the concept of machine learning and deep learning techniques utilised in the experiments are later explained in this chapter.

4.1 General Methodologies

In this thesis, digital images are used as materials. An image according to [52] is defined as a two-dimensional function, $f(x, y)$, where x and y are spatial coordinates and the intensity of any pair of coordinates as the amplitude of f . A digital image can be defined when x , y and the value of intensity in f are all finite and discrete quantities. The value of each element in a digital image is called a *pixel*.

Each pixel can be manipulated according to the needs of the end user. The manipulation of each pixel will lead us to an area called image processing. The field of digital image processing includes input and output processing in the form of images. Extraction of attributes from an image, not limited to recognising the objects contained in it, is also included in this area. The earliest application of digital images was in the newspaper industry when they were first transmitted via an underwater cable between London and New York [52].

Since the 1970s, image processing has thrived and has been used in a broad range of applications. In the early 1970s, digital image processing techniques began to be used in medical imaging, astronomy, and space applications. Using an X-ray source, the invention of *computerised axial tomography* or *Computerised Tomography* (CT)

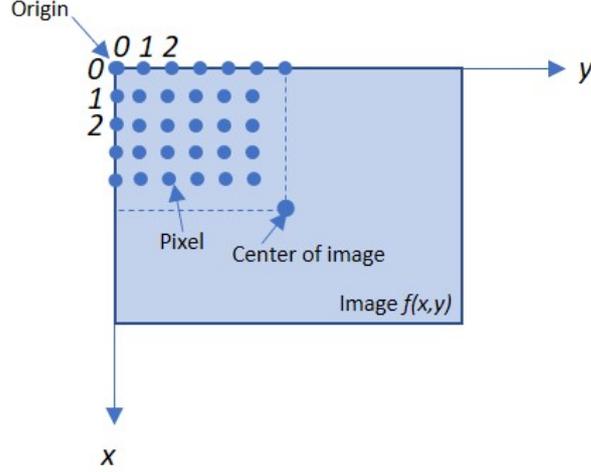


Figure 4.1: The illustration of the coordinate convention used to represent a digital image. This image is adapted from [52].

becomes the most critical event in applying image processing in medical diagnosis. In this study the focus is on images produced by X-rays, hence the emphasis of explanation in this chapter will always be related to this field.

4.1.1 Geometric Transformations

The images contained in the data set used in this study came in a variety of conditions such as shape, contrast, and brightness. The only visual attribute we have is the edge of the image, which often also varies in shape and condition. These images need to be aligned. The process of transforming different sets of pixels data into a spatial arrangement in an image is called *Geometric Transformations*. Registration is necessary in order to be able to compare or integrate data from different measurements. This research mainly focuses on linear transformation models of image registration. It includes rotation, scaling, shearing, and translation.

Two basic operations of geometric transformations of digital images are coordinates changes in the spatial domain and the intensity interpolation that assigns values to the spatially transformed pixels. Coordinates transformation may be expressed as

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \mathbf{T} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (4.1)$$

where (x,y) are pixel coordinates in the original image and (x',y') are the coordinates of pixels corresponding to the transformed image. The Affine transformations are able to do identity, scale, rotate, translate, and shear operations on an image. The

thing to do is to transform certain parts of the image by maintaining several things such as points, straight lines, and planes. Equation 4.1 is an example of how this transformation process is carried out while Table 4.1 is some example of a transformation that can be done by using equation 4.2.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4.2)$$

The operations in Table 4.1 can be performed depending on the values used on the elements of matrix A. In the Identity operation, each value will be processed in proportion to the values in the existing matrix in the original image so that the same image results are obtained from the original image. Each pixel in the original image will be changed according to the magnitude of the multiplier given to the image in the Scale operation. In this operation, the image will appear bigger or smaller than the original image.

The pixels in the image can also be processed to get a different result based on changes in the pixel's degree position on the image. Changes are applied by processing the pixel's value in the degree direction through all regions of interest. The name of the operation is Rotation. The rotation process outlines appear in Table 4.1. The image is also can be manipulated by making a shift impression. In this case, the image seems like doing a movement. This case is referred to as the Translation. Translation operations are carried out by processing each pixel in the image by changing the location value by the desired number of values—for example, tx for pixels on the x-axis and ty for the y-axis.

Changing one of the values of the image axis is also possible to apply. This process will maintain one axis value and change the value of the other axis to the value that comes from multiplication with a certain value. This operation is called Shear. The effect resulting from this operation is an image that appears to have a certain depth even though the image is actually in two-dimensional conditions. In Table 4.1, two types of Shear operations are displayed. First is the Shear operation which changes the **x-axis** value with a certain value Sv multiplied by **they-axis** value. The shear effect obtained is vertical shear. Second, then as it's also shown in Table 4.1, the reverse condition from the previous operation will result in an image in horizontal shear condition.

Table 4.1: These are types of geometric transformation usually applied to an image in the image processing field [52].

Transformation Name	Affine matrix	Coordinate Equations
Identity	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= x \\ y' &= y \end{aligned}$
Scaling	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= c_x x \\ y' &= c_y y \end{aligned}$
Rotation	$\begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= x \cos \theta - y \sin \theta \\ y' &= z \sin \theta + y \cos \theta \end{aligned}$
Translation	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= x + t_x \\ y' &= y + t_y \end{aligned}$
Shear (vertical)	$\begin{bmatrix} 1 & s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= x + s_v y \\ y' &= y \end{aligned}$
Shear (horizontal)	$\begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{aligned} x' &= x \\ y' &= s_h x + y \end{aligned}$

Figure 4.2 (b) is an example of Rotation applied to wrist X-ray images. Image in Figure 4.2 (a) is an original X-ray image in a lateral position, then its 45° rotation resulted image in Figure 4.2 (b). Figure 4.3 shows an example of the translation operation of an image. In this case Figure 4.3 (b) is resulted from the translation coordinate of a vector of $x = 250.3$ and $y = -300.1$ from the original position of Figure 4.3 (a). An example of scale operation shows in Figure 4.4 (b). This image was 50% scaled-down, as it can be observed from the size of the images shown in the image coordinates. Figure 4.4 (b) appearance still looks like the original image, but the scale of the image has changed.

4.1.2 Edge Detection

The edge in an image is a condition when the intensity of the pixels of an image changes abruptly. In many cases of image processing, the detection of this edge can be a crucial part of the process. Edge detection needs a combination of point, line,

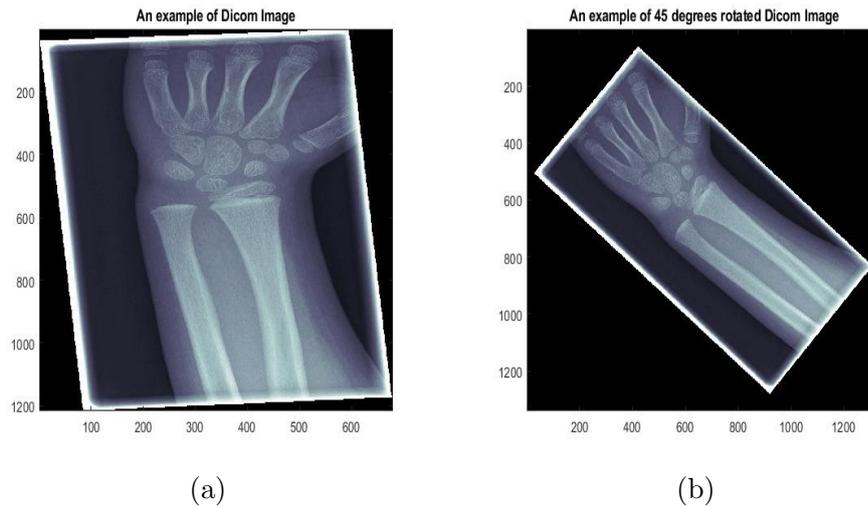


Figure 4.2: These are examples of applied Rotation of an image for wrist X-rays images. Image (a) is an original wrist X-ray image in a lateral taken position. Image (b) results from 45° rotations from the original position of the image (a).

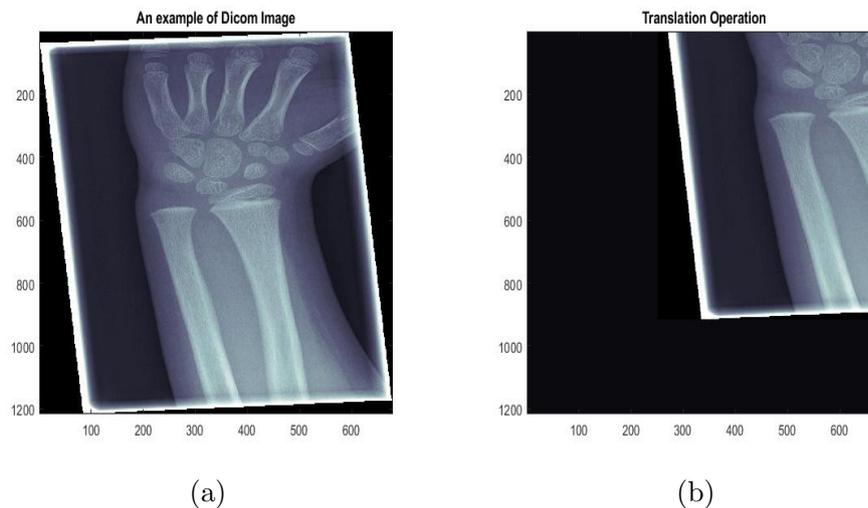


Figure 4.3: These are examples of applied image translation of an image for wrist X-rays images. Image (a) is an original wrist X-ray image in a lateral taken position. Image (b) is the result translation vector of $x = 250.3$ and $y = -300.1$ from the original position of image (a).

and edge detectors to determine sharp, local changes in pixel intensity. Edges are determined according to the forms of their pixel intensity.

According to [52], there are three edge models in a digital image: Step Edge, Ramp Edge, and Roof edge models. These three models are illustrated as shown in Figure 4.6. The ‘Step Edge’, as shown in Figure 4.6 (a), is a type of Edge that has a transition shape between two steep pixel intensities or can be called vertical differences. This edge model can be ideal because the differences in values between pixels are substantial, so differences between pixels can be obtained easily. Edges with this model do not require further processing because of the shape of the edges

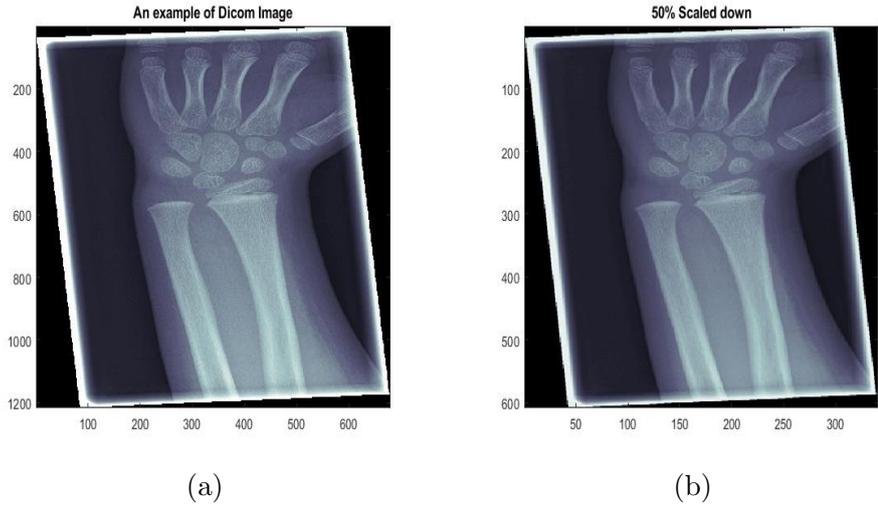


Figure 4.4: These are examples of applied image scaling of an image for wrist X-rays images. Image (a) is an original wrist X-ray image in a lateral taken position. Image (b) is 50 per cent scaled down from the original position of the image (a), which could be observed from the value image coordinates.

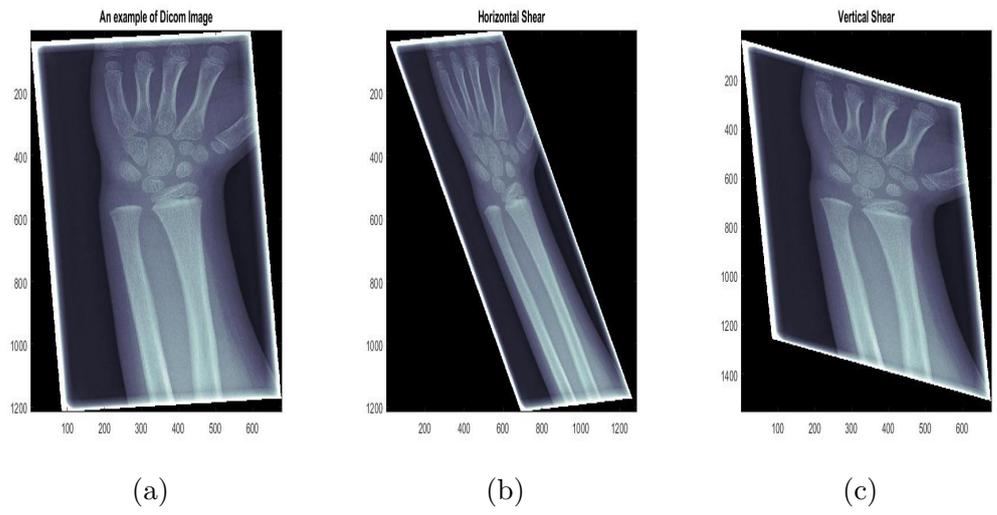


Figure 4.5: These are examples of shear operation to an image of wrist X-rays. Image (a) is an original wrist X-ray image in a lateral taken position. Image (b) is the result of the horizontal Shear of the original position of image (a), and image (c) is the result of the vertical Shear of the original image(a).

visible in the image. The second form of Edge is the Ramp Edge, as shown in Figure 4.6 (b). The ‘Ramp Edge’ is an edge model with gradually changing values between pixels to a significant difference. In this model, no point directly becomes the difference between image pixels, but it has a range of changes with a specific value. Figure 4.6 (c) illustrates that a third type is Roof Edge. Roof edges are models of lines where the width of the Edge is determined by the thickness and sharpness of the line. This type of Edge usually appears in the digitisation of line drawings, which contain delicate features like satellite images. In this image, delicate features such

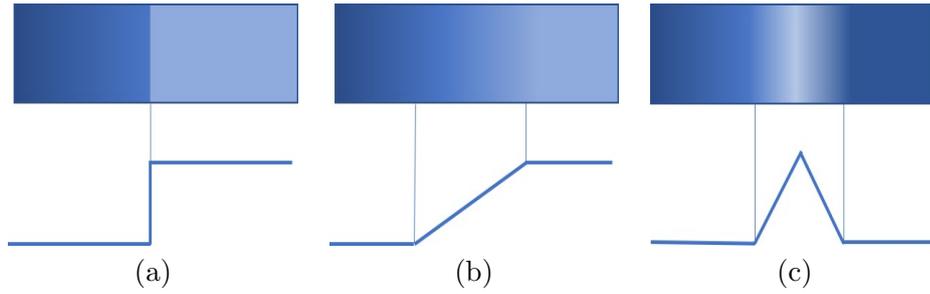


Figure 4.6: These are the ideal representations of edges. From left to right, a step-edge model (a), a ramp-edge model(b), and a roof edge model(c). Below each model is corresponding intensity profiles. These illustrations are adapted from [52].

as roads can be identified as this type of Edge. Three steps commonly performed for edge detection are image smoothing, Edge points detection, and localisation of the Edge [52].

How to precisely detect changes in the image pixels' intensity is the basic principle in edge detection. Calculating the gradient of each pixel of the image is a way to detect the edges of an image. Image gradients provide information about the strength and the direction of the change in the image value at each pixel.

Image Thresholding

Image thresholding is one common solution to detect the edge in an image. This method is based on pixel(s) value to differentiate between the foreground and the background of an image. The basic idea of thresholding is illustrated in Figure 4.7. These two graphics illustrate how the thresholding point can be put in the sudden spike difference of value between pixels. Figure 4.7 (a) illustrates the contrast stretching of the corresponding pixel value that transformed the image from dark contrast to light colour contrast.

The difference point can be taken as a threshold to transform an image into one the case needed. If the difference value between the pixels totally contrasts like Figure 4.7 (b) illustrated, the threshold point can be easily defined. Suppose the intensity histogram in Figure 4.8 corresponds to an image $f(x, y)$, we can define T as a threshold and the region above or below this threshold can be determined by using this value. In general this case denoted by $g(x, y)$ where

$$g(x, y) = \begin{cases} 1 & f(x, y) > T \\ 0 & f(x, y) \leq T \end{cases} \quad (4.3)$$

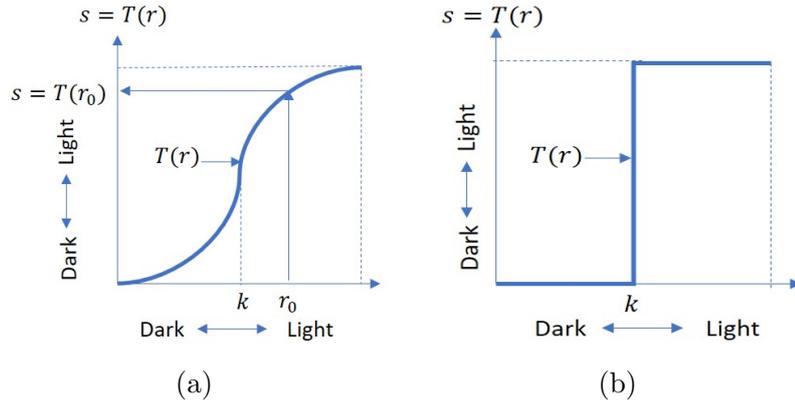


Figure 4.7: (a) Contrast stretching function (b) Thresholding function. These graphics are adapted from [52]

T can be applied over an entire image and able to refer as *global thresholding*, but if the value of T dynamically changes over the process of the image, it can be also referred to as *variable thresholding*. In some cases, there can be more than one threshold defined along the process of the image, like denoted in following *multiple thresholding*. The following classified equation can be changed depending on the situation applied to the case of the image, where a , b , and c are any three distinct intensity values.

$$g(x, y) = \begin{cases} a & f(x, y) > T \\ b & T_1 < f(x, y) \leq T_2 \\ c & f(x, y) \leq T_1 \end{cases} \quad (4.4)$$

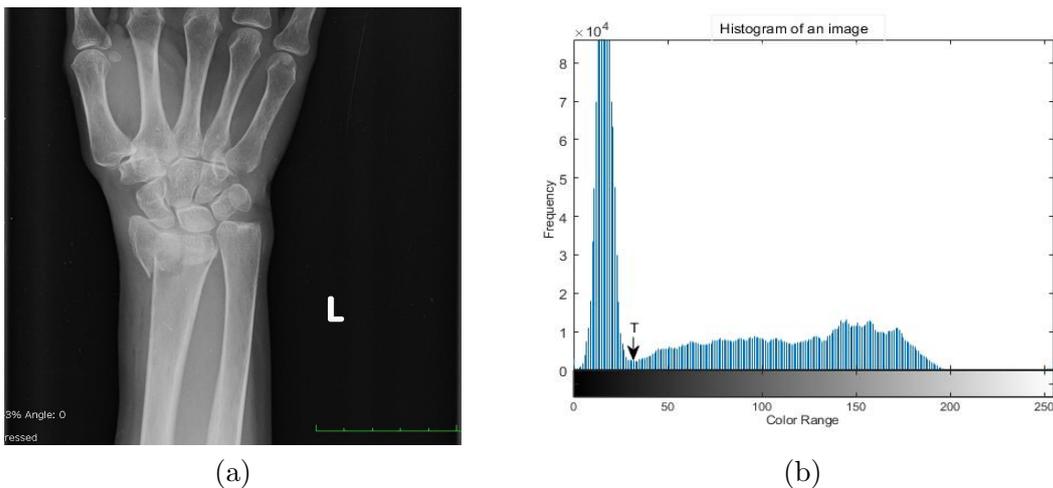


Figure 4.8: (a) An example of the wrist X-ray image, (b) Intensity histogram that can be partitioned by T (threshold).

Otsu's method [99] and image filter are several methods related to this field,

like Canny's edge detector [21]. An automatic threshold was proposed by Otsu's method, which was selected for image segmentation without any parameters and unsupervised methods. This method maximises the difference of grey-level in the image. The method utilises image separation of foreground and background. The separation is based on the threshold of cumulative intensity pixels in the grey-level area colour space histogram. This method is quite general as it covers a broad scope of unsupervised decision procedures. Details of Otsu's method are explained in [99].

Canny's edge detector [21], named after the inventor of this method John Canny, was introduced in 1986 through his paper. Canny's approach is based on three primary objectives to find the edges. First, there must be a comprehensive set of goals for the computation of edge points whilst making minimal assumptions about the form of the solution. Second, edge points should be well localised close to the actual edges. The third objective is that the detector should only return one response to a single edge. It means that the detector should not identify multiple edge pixels where only one exists.

Examples of implementing Canny's edge detector [21] and Otsu's method [99] in digital images are shown in Figure 4.9. In this image, Figure 4.9(b) is the processing result of an image in Figure 4.9 (a) as the implementation of edge detection using the Canny edge detector as well as Figure 4.9 (c) is the result of implementing edge detection using Otsu's method. Based on the results shown in Figures 4.9 b and c, these methods provide an opportunity for edge detection of objects in the image. This is because these two exemplified can reduce the pixels that are suspected to be not the edges of the image. This capability provides a way for fracture research on digital wrist X-ray images.

Linking Edge Points

Canny's and Otsu's are applicable when pixel information belonging to certain conditions is available. If the image does not provide unique information or we only have an edge map and no knowledge of where objects of interest might be, we need global processing of the image. The idea that each pixel has its unique information and is possibly connected can be utilised as linking candidates for us to reserve or eliminate during edge detection. Hough's transforms are used to detect lines in the image [60].

Figure 4.10 is an illustration of the basic idea of the connections between two

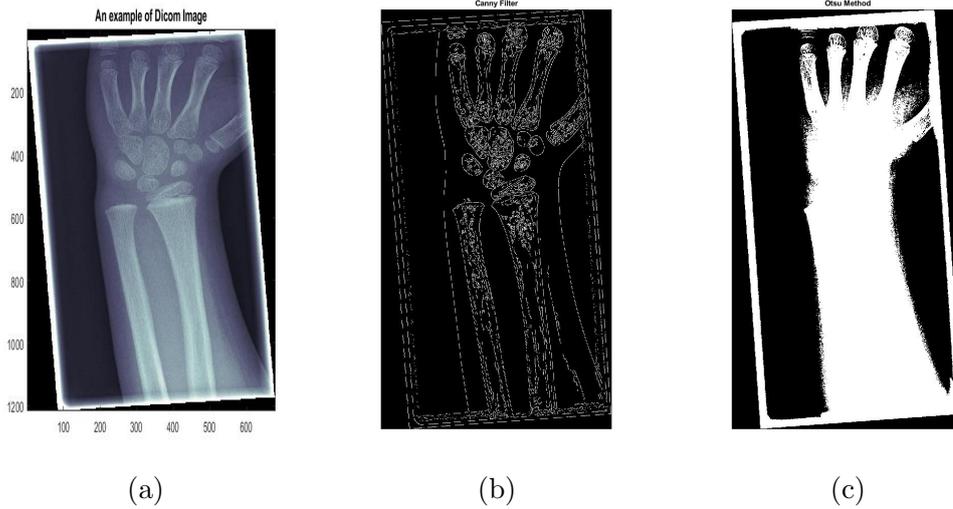


Figure 4.9: These are applied edge detection techniques examples to an image for wrist X-rays images. Image (a) is an original wrist X-ray image in a lateral taken position. Image (b) is the result of the Canny edge detection method [21]. Image(c)is the result edge detection using Otsu’s method [99].

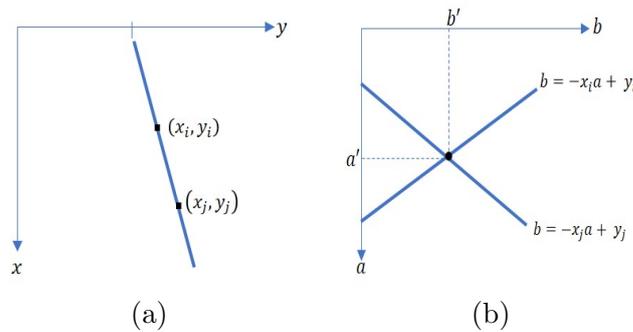


Figure 4.10: (a) xy -plane as an illustration of the idea of a connection between two points. (b) An example of parameter space to describe the general equation of a straight line where infinitely lines pass through. These graphics are adapted from [52].

points that could link each other. In this picture, there are two illustrations of basic interactions between points and in the image, these kinds of points will consider as lines and could probably lead to the edge(s) of the object in the image. Figure 4.10(a) explains about two points (x_i, y_i) and (x_j, y_j) in the xy -plane. Figure 4.10(b) illustrates the parameter space to describe the general equation of a straight line of $y_i = ax_i + b$, where infinitely many lines pass through (x_i, y_i) and values of variables a and b . Each point represents and probably consists of various line intersects and these lines are equivalent to the original line. However, a difficulty with this approach is that a as the slope of a line, approaches infinity as the line approaches the vertical direction. Hough’s transform uses the normal representation of a line: $x \cos \theta + y \sin \theta = \rho$. Figure 4.11(a) illustrates the geometrical interpretation of the parameters ρ and θ . Illustrations of Figure 4.11 are adapted from [52]. A horizontal

line has $\theta = 0^\circ$, with ρ being equal to the positive x -intercept. Similarly, a vertical line has $\theta = 90^\circ$, with ρ being equal to the positive y -intercept. For Figure 4.11(b), each sinusoidal curve represents the collection of lines that pass through a particular point (x_k, y_k) in the xy -plane. Figure 4.11(b) describes the intersection point (ρ', θ') that corresponds to the line that passes through both (x_i, y_i) and (x_j, y_j) in Figure 4.11(a).

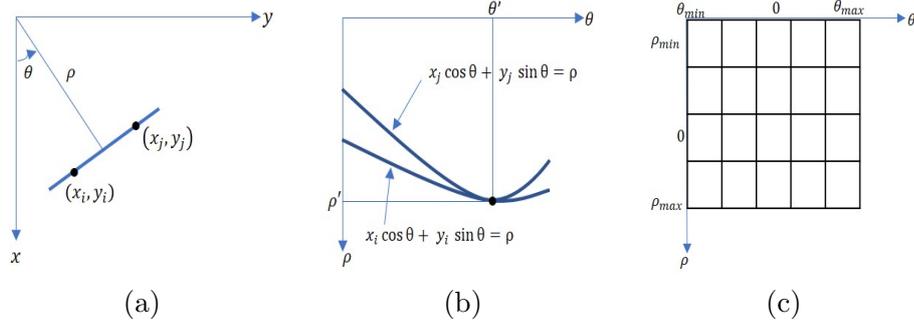


Figure 4.11: (a) The geometrical interpretation of the parameters ρ and θ . (b) Each sinusoidal curve represents the collection of lines that pass through a particular point (x_k, y_k) in the xy -plane. (c) An example of the accumulator cells, where every non-zero pixel, this sinusoidal curve will be created and counted in these cells. These graphics are adapted from [52]

In assumption of θ range between -90° to 90° or also can describes as $-90^\circ \leq \theta \leq 90^\circ$, aligned with the $\rho\theta$ parameter in Figure 4.11(c), the *accumulator cells* can be defined by subdividing this the $\rho\theta$ parameter. The expected range of the parameter are (ρ_{min}, ρ_{max}) and $(\theta_{min}, \theta_{max})$. The range for this parameter is $-90^\circ \leq \theta \leq 90^\circ$ and $-D \leq \rho \leq D$, where D is the maximum distance between the opposite corners in an image. The looping process operates through all the image pixels. Initially, these cells are set to zero and identified this value as not an edge during the looping process of entire pixels in the image. If the pixel is non-zero, a sinusoidal curve is in the $\rho\theta$ space. The $\theta = -90$ and calculate the corresponding value of ρ using equation $\rho = x_k \cos \theta + y_k \sin \theta$, where k is relate to (x_k, y_k) of non-background point. Every single line through this procedure counts as a 'vote' in *accumulator cell* (θ, ρ) and increases the value of the cell by 1. For every non-zero pixel, this sinusoidal curve will be created and counted. A lot of points 'voted' for this spot which means as the parameters to describe the lines in the original image.

Figure 4.12(a) is a wrist X-ray image that we try to identify its Edge through Hough's transform. Visually, the interest of the image that wants to identify is the square shape edge created by the collimator plate during image acquisition. Once Hough's transforms technique is applied, Figure 4.12(b) will appear to represent

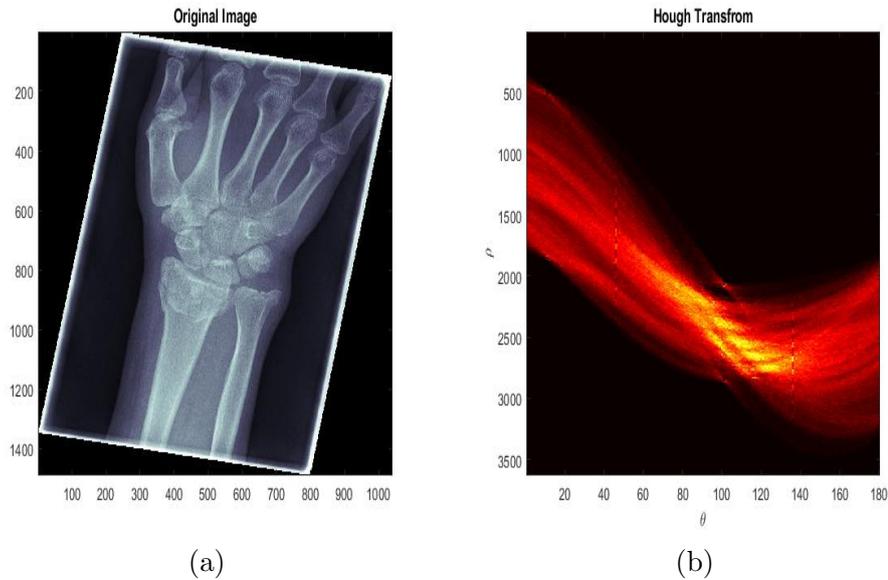


Figure 4.12: An example implementation of Hough's transform to align the X-ray wrist image in a certain angle degree. (a) An example of the wrist X-ray image, (b) The result of vote pixels in accumulator cells between ρ and θ on the image proceeds through the transform where the brightest colour represents potential lines in the input image.

accumulator cells between rho and theta value of the image. The bright colour parts of this graphic represent the potential lines in the image. The value of potential lines can be traced to get the maximum theta value in the image. This value can be used to align the image vertically.

4.1.3 Contrast Limited Adaptive Histogram Equalisation

An image often comes in unpredictable conditions. Poor contrast and a low level of brightness could be reasons for an imperfect image condition. It needs the effort to make it visually better. Image enhancement could be translated as the effort to make an image better visually from the viewer's point of view. One classic strategy is manipulating the image's pixels. Pixels within an image could be manipulated by modifying their values. The pixel's value could give the different effects of contrast and level of brightness.

Contrast is a part of the image which gives a nuance of luminance or colour that makes an object distinguishable. The difference in luminance can give a contrast effect to the image. The contrast gives colour and sometimes further interpretation when it combines with brightness. The idea to enhance image visualisation by adjusting the grayscale of the image's pixels based on its pixel's distribution in a histogram was penned by [63] and referred it as Histogram Equalisation.

Histogram equalisation aims to obtain an even distribution of the histogram, such that each degree of grey has a relatively equal number of pixels. This strategy then widens the range of grayscale values to increase the contrast of the image. It changes the current grayscale value of pixel r with a new grayscale value s based on its equalised histogram transformed an image as a function T .

$$s = T(r). \quad (4.5)$$

and the probability of a pixel with a certain grayscale value within the image as

$$P_r(r_k) = \frac{n_k}{n} \quad (4.6)$$

where n_k is a number of pixels with a certain value of r_k . n is the total number of pixels within an image. For instance, the average r value of grayscale k is also normalised to the maximum length gray's L value of the image as it calculated as

$$r_k = \frac{k}{L-1}, 0 \geq k \leq L-1 \quad (4.7)$$

for example if $L = 8$, the values of r_k can be calculated as follow

Table 4.2: Values of r_k if $L = 8$

k	r_k
0	$0/7 = 0$
1	$1/7$
2	$2/7$
3	$3/7$
4	$4/7$
5	$5/7$
6	$6/7$
7	$7/7 = 1$

A transformation of pixels using histogram equalisation of a range of colour-scale and a size of an image could be expressed as follows :

$$s_k = T(r_k) = \sum_{j=0}^k \frac{n_j}{n} = \sum_{j=0}^k P_r(r_j) \quad (4.8)$$

where $0 \geq r_k \leq 1, k = 0, 1, 2, \dots, L-1$

A modest example of an image like Figure 4.14 has size 4×4 and consists of 16 pixels. It has a 10 level of gray level which can be spread from 0 to 9 grayscale level.

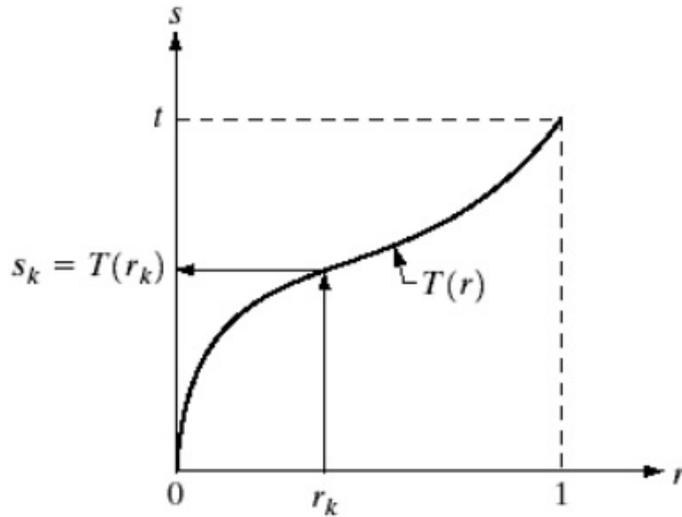


Figure 4.13: An graphic illustration of transformation T grayscale r value of an image to a new grayscale value of pixels s

Histogram distribution of pixels to its gray level is shown in Figure 4.15 and it is used as an explanation of how histogram equalisation applied to an image.

2	3	3	2
4	2	4	3
3	2	3	5
2	4	2	4

Figure 4.14: An example of an image which has size of 4×4 and 16 pixels.

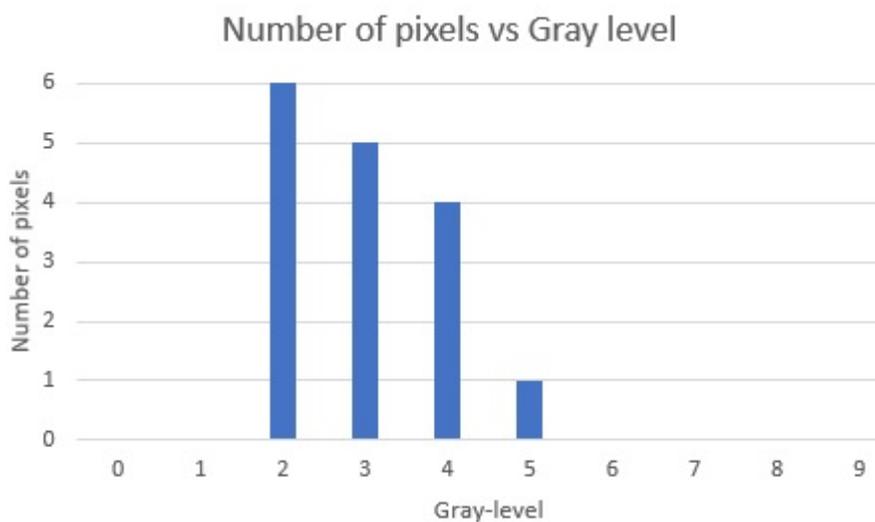


Figure 4.15: A distribution of pixels within image in Figure 4.14. In assumption, this image has ten levels of the grayscale colour map. This histogram will be the basis of the histogram equalisation process.

After the distribution of pixels on the image calculated as shown in Figure 4.15, then formula 4.8 applied to pixels within Figure 4.14. Details are explain below in Table 4.3.

Table 4.3: Detail calculation of histogram equalisation of image Figure 4.14. Abbreviations: Frequency (freq), Cumulative Distribution (Cumltv Dist).

Gray level	Pixel's value Freq	Cumltv Dist	Pixel's grayscale probability	Pixel's new gray level value ($s = \sum_{j=0}^k \frac{n_j}{n}$)	New pixel's value ($s * 9$)
0	0	0	0/16	0	0
1	0	0	0/16	0	0
2	6	6	6/16	6/16	3.9 ~3
3	5	11	5/16	11/16	6.1 ~6
4	4	15	4/16	15/16	8.4 ~8
5	1	16	1/16	16/16	9
6	0	16	16/16	16/16	9
7	0	16	16/16	16/16	9
8	0	16	16/16	16/16	9
9	0	16	16/16	16/16	9

The new pixel's values from the equalisation procedure are then mapped through the related pixels within the image. The comparison of images before and after the histogram equalisation process is shown in Figure 4.16.

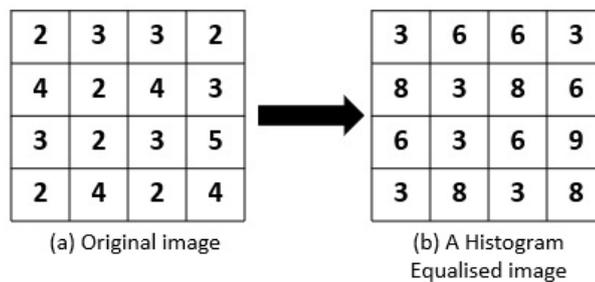


Figure 4.16: These are comparisons between an original image (a) and the result (b) after proceeding through the histogram equalization technique.

This phenomenon then developed as an image enhancement technique named Adaptive Histogram Equalization (AHE) as it has been published by [106]. This method works well if the pixel values distribution is similar throughout the image. Figure 4.20 (b) shows a result of the histogram equalisation process of Figure 4.20 (a). Comparison of pixel distribution between the original image and after histogram equalization applied to the image shows in Figure 4.21. The distribution of pixels before and after equalisation spreads through all levels of the grayscale colourmap, and it gives a brighter looks image, as shown in Figure 4.18 (b). In this case, the

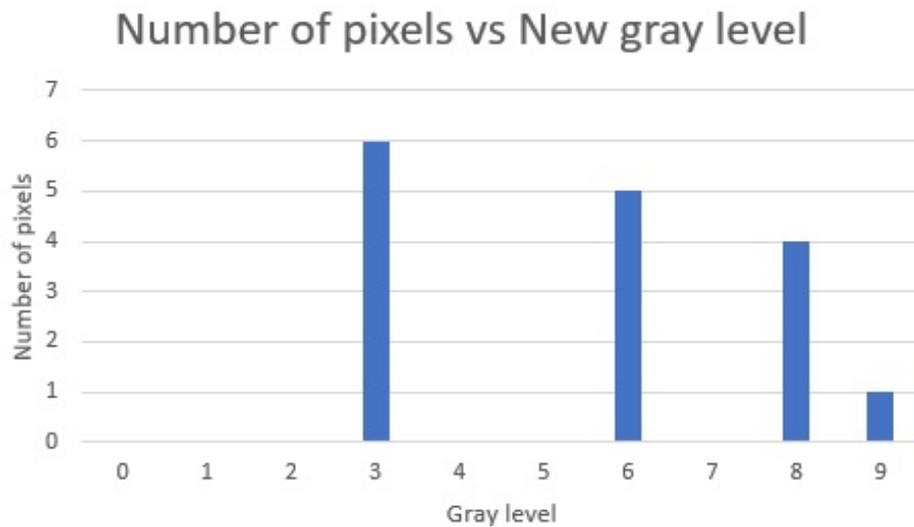


Figure 4.17: A distribution of pixels within image in Figure 4.16(b). In assumption, this image has 10 levels of the grayscale colour map. This histogram is the result of image (a) as the basis of the histogram equalization process.

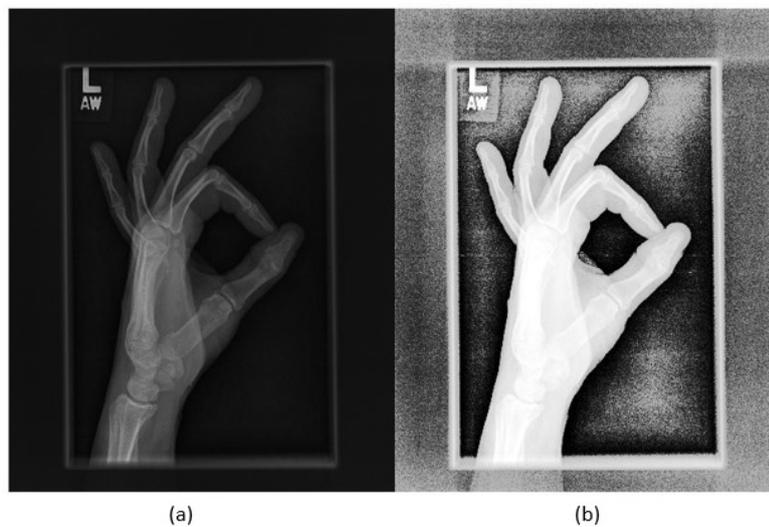


Figure 4.18: An example of histogram equalisation applied to a medical image. Image (a) is an original image and (b) is an image resulting from the classic histogram equalization process.

classic histogram equalisation does not work well because, from a visual perspective, much of the bone details within the X-ray image were lost as pixel values turned significantly white-coloured. This is the weakness of the original histogram equalisation where the significantly lighter or darker pixel values dominate the image as such an X-ray image can affect the result of pixels distribution after the equalisation process. The need for slightly mild or adaptive contrast-related enhancement of pixel values within the image that had been developed relating to the image contrast. Transforming each pixel from a neighborhood region with a transformation function then named Adaptive Histogram Equalization (AHE) and its specific work

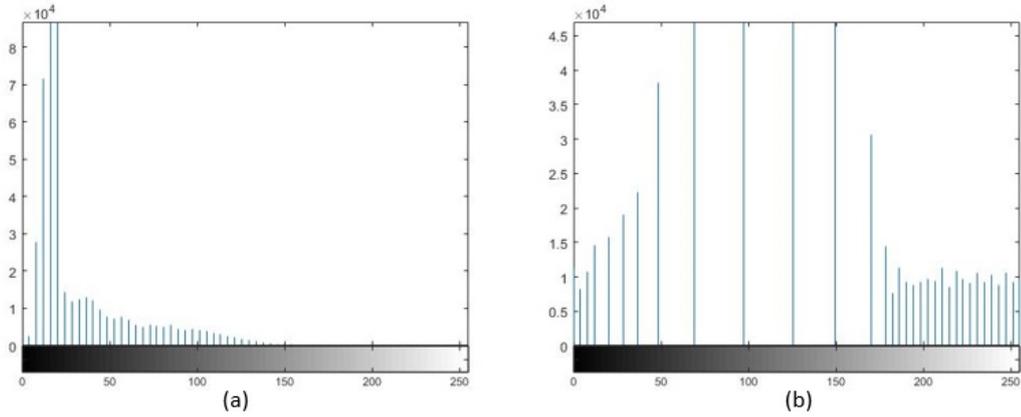


Figure 4.19: Two histograms applied to medical image in Figure4.18. Image (a) is a histogram of the original image of Figure4.18(a) and (b) is a histogram of Figure4.18(b) resulting from the histogram equalization process.

with contrast updated to The Contrast-Limited Adaptive Histogram Equalization (CLAHE).

The Contrast-Limited Adaptive Histogram Equalization (CLAHE) is an image processing method that addresses the loss of local contrast by dividing the image into small regions called tiles or blocks and performs histogram equalization independently on each tile [106]. The result of applying CLAHE is an image with increased contrast and increased detail, particularly in areas of low contrast or uneven illumination. This thesis' colour range is primarily binary black and white and grayscale colour range.

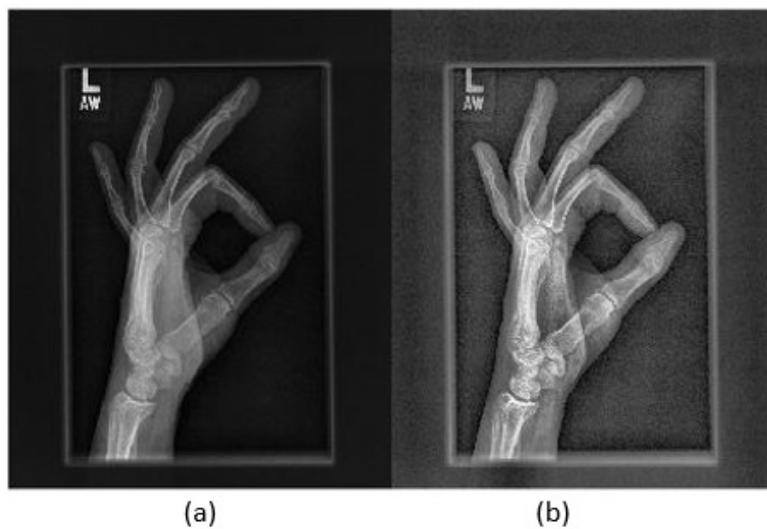


Figure 4.20: An example of contrast limited histogram equalisation applied to a medical image. Image (a) is an original image and (b) is an image resulting from CLAHE process.

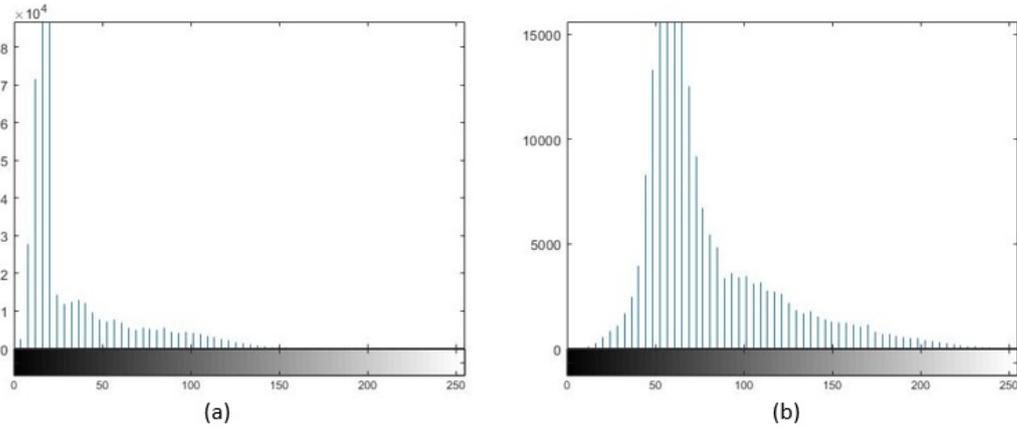


Figure 4.21: Two histograms of a wrist X-ray image in Figure 4.20. Image (a) is a histogram of original image of Figure 4.20(a) and (b) is a histogram of Figure 4.20(b) resulted from contrast limited histogram equalisation process.

4.1.4 Local Binary Patterns

An image is built by pixels. Each pixel contains a value in the range of colour space. Pixel has relations with other pixels around it. Pixel neighbourhood is usually used to describe the relationship between a pixel and the pixels surrounding it. Pixel neighbourhood has relative intensity but not a strict relation as a whole image. Relation between pixels and their neighbourhood has an intensity to construct a texture. The small intensity value of the neighbourhood of pixels explores using a technique named Local Binary Patterns (LBP).

Local Binary Patterns was originally proposed by [137]. It is based on the Texture Unit's foundation and the Texture Spectrum concept, described in [136]. The standard LBP operator is defined as a grayscale invariant texture measure derived from a general texture definition in a local neighbourhood. An example of the relationship between pixels of a 3×3 neighbourhood is proposed by [136]. Pixels neighbourhood in small units called Texture Units. It connects eight directions from the central pixel and is considered the smallest complete texture unit within this neighbourhood.

For instance, a neighbourhood of given 3×3 pixels contains a set of nine elements $V = \{V_0, V_1, \dots, V_8\}$, where V_0 represents the central pixel's intensity value and V_i represents the pixel i^{th} intensity value of this neighbourhood. The set of texture units is defined by these eight elements as $TU = \{E_1, E_2, \dots, E_8\}$ where E_i is the

threshold by

$$E_i = \begin{cases} 0 & \text{if } V_i < V_0 \\ 1 & \text{if } V_i = V_0 \\ 2 & \text{if } V_i > V_0 \end{cases} \quad (4.9)$$

for $i = 1, 2, \dots, 8$ and E_i is the same position as the i^{th} pixel. Each element of the texture unit (TU) has three possible values. If we considered those values as 0, 1 and 2 obtained from a neighbourhood of 3×3 pixels then we have $3^8 = 6561$ possible Texture Units. [136] has found a way to label all texture units and named the Texture Unit Number (N_{TU}). It was found by using the following formula:

$$N_{TU} = \sum_{i=1}^8 3^{i-1} E_i, \quad N_{TU} \in \{0, 1, 2, \dots, 6560\} \quad (4.10)$$

Pixels are clockwise ordered starting from the top-left, to the middle-left as shown below.

a	b	c
h		d
g	f	e

Figure 4.22: An illustration of eight pixels neighbourhood. These pixels are ordered in clockwise as the eight elements of Texture Unit. The first element is a as element of E_1 clockwise-way through h as the element E_8

A square shape of a neighbourhood is not a strict rule for Local Binary Patterns. It could be other shapes such as circularly symmetric neighbourhoods as proposed by [95] as an updated technique of their work [96].

The method proposed by [137] as it explains in [96] is using a two-level version as it shows in Figure 4.23. The possibility of 3^8 equals 6561 possible Texture Unit could have resulted from 3×3 pixels neighbourhood then elaborated in Figure 4.23 as an example. A region of 3×3 pixels in Figure 4.23 (a) is then thresholded to two-level of the binary distinction of 0 and 1 by comparing it to the value of the centre pixel which has a value of 6. It gives results in 4.23(b). If we put an assumption that Figure 4.23 (c) as weights given to the corresponding pixels then multiplied it with those in 4.23 (b). The results are shown in 4.23 (d). The total values of pixels in Figure 4.23 (d) is 180 as the number of texture units of this neighbourhood.

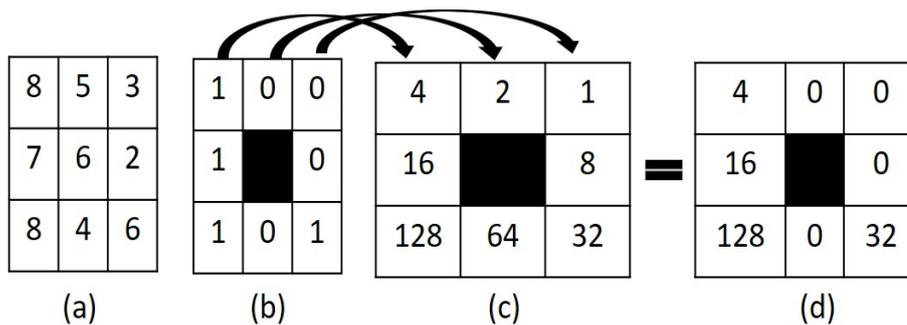


Figure 4.23: An illustration to describe Local Binary Patterns (LBP) in a texture. A 3×3 neighbourhood (a) is thresholded by the center pixel's value. The values of each pixel in (c) then multiplied with two-level thresholding results in (b) resulted (d). Texture unit of this neighbourhood is obtained by sum up the eight pixels which got $TU = 180$. Images inspiration from [96].

An image is usually composed of texture elements and random noises. More immense proportions of texture than noise may be seen by human vision. This certainty to differentiate between the wanted and unwanted value of pixels becomes the advantage of the LBP. This technique could apply in a grayscale colour range and be combined with a contrast measure.

Different distributions of texture spectra within particular texture units could lead to various image textures [137]. This technique was first described by [96] that were using the grey-level differences with classification based on feature distributions to determine image texture measures. A Texture Unit was calculated by differentiating the grey level of a central pixel from the grey level of its neighbours. The occurrence of the distribution of texture units computed over a region is called the texture spectrum. In the Texture Spectrum, an increase in the percentage of texture components in an image will tend to form a particular distribution of peaks. This method is further developed by [95], which generalised grayscale and rotation invariant operator which allows detecting similar patterns for any quantisation of the angular space and any spatial resolution and presents a method for combining multiple operators for multiresolution analysis.

Classification of an image could benefit from the texture spectrum of an image and become its descriptor. In a recent development, texture and colour as descriptors for visual recognition would be possible to cooperate in a data-driven approach to potentially replace the manual-driven approach as it has been overview historically by [14].

In a medical image, texture interpretation is crucial to determine the abnormality

Table 4.4: Details of convolutional neural networks (CNNs) that were used in this work.

No.	Network	Depth	Image Input Size	Reference
1	GoogLeNet	22	224-by-224	[126]
2	VGG-19	19	224-by-224	[121]
3	AlexNet	8	227-by-227	[70]
4	SqueezeNet	18	227-by-227	[64]
5	ResNet-18	18	224-by-224	[59]
6	Inception-v3	48	299-by-299	[128]
7	ResNet-50	50	224-by-224	[59]
8	VGG-16	16	224-by-224	[121]
9	ResNet-101	101	224-by-224	[59]
10	DenseNet-201	201	224-by-224	[61]
11	Inception-ResNet-v2	164	299-by-299	[127]

observed in the image. The opportunity to use Local binary pattern variants as texture descriptors for medical image analysis was evaluated by [88].

Information within the image could be seen from the value of features of the image. Features of the image contain a specific value that builds up the image itself. Fractures, in general, are challenging to distinguish from an X-ray image. It happened because of various levels of colour intensity related to image texture. Fractures are connected, and separating abnormality within a medical image is a difficult task. Image noise such as a blurred result of scanned images could be another challenge. The opportunity of using this technique for a specific analysis of wrist X-ray images is explored in Chapter 6.

4.1.5 Convolutional Neural Networks (CNN)

Convolutional Neural Networks are a subclass in the hierarchic terminology that includes artificial intelligence (AI), machine learning, and deep learning. CNN is a specialised neural network for processing data with a grid-like topology [53].

Overview of CNN architectures

The convolutional neural network research design process includes defining the clinical question, choosing a predefined computer vision task, generating data acquisition and preprocessing, selecting hardware and software solutions, developing a network architecture, and validating the algorithm performance. In this report, eleven convolutional neural network architectures have been tried to adapt to a given data set. The following section explains the eleven CNN architectures that were used.

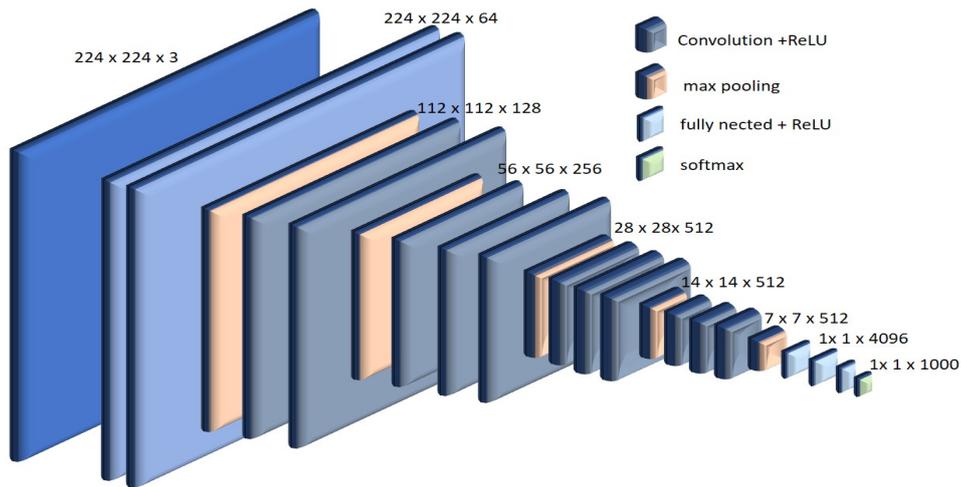


Figure 4.25: A diagram showing VGG-16 CNN Architecture which is adapted from [121].

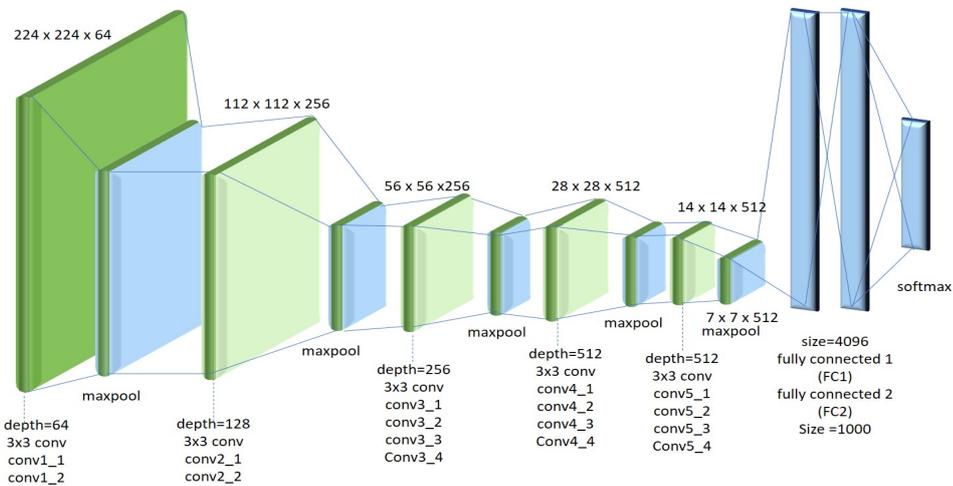


Figure 4.26: A diagram showing VGG-19 CNN Architecture which is adapted from [121].

SqueezeNet

SqueezeNet is a convolutional neural network that is trained on more than a million images from the ImageNet database [64]. The network is 18 layers deep and can classify images into 1000 object categories. The network has an image input of size 227-by-227

ResNet-18

ResNet-18 is a convolutional neural network that is trained on more than a million images from the ImageNet database [59]. The network has an image input size of 224-by-224. The network is 18 layers deep.

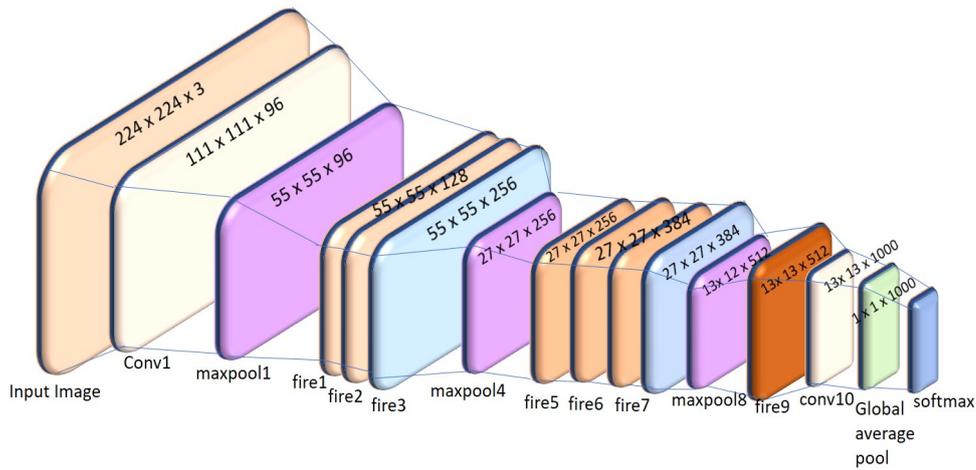


Figure 4.27: A diagram showing SqueezeNet CNN Architecture which is adapted from [64].

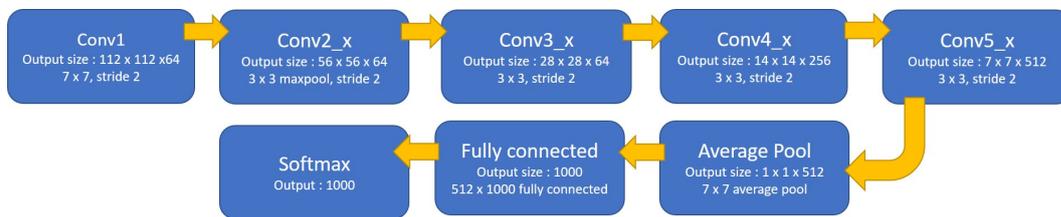


Figure 4.28: A diagram showing ResNet-18 CNN Architecture which is adapted from [59].

GoogLeNet

GoogLeNet is a convolutional neural network that is trained on more than a million images from the ImageNet database [126]. The network has 22 layers deep. It has an image input size of 224-by-224.

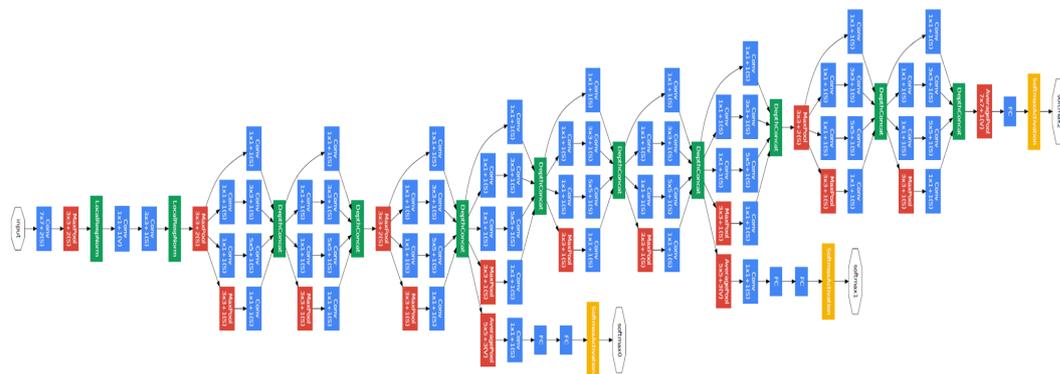


Figure 4.29: A diagram showing GoogLeNet architecture which is adapted from [126].

ResNet-50

ResNet-50 is a convolutional neural network that is trained on more than a million images from the ImageNet database [59]. The network has 50 layers deep and can

classify images into 1000 object categories. It has an image input size of 224-by-224.

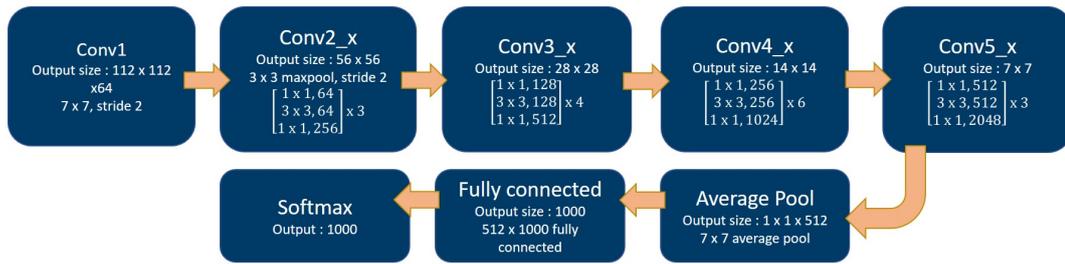


Figure 4.30: A diagram showing ResNet-50 CNN Architecture is adapted from [59].

Inception V-3

Inception V-3 is a convolutional neural network that is trained on more than a million images from the ImageNet database [128]. The network has 50 layers deep. It has an image input size of 299-by-299

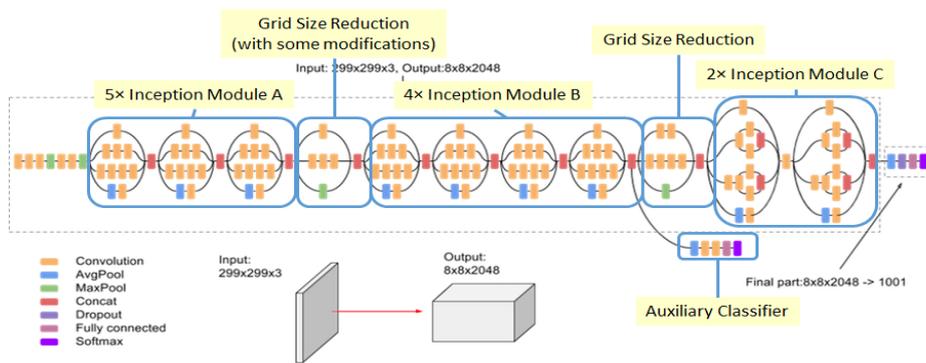


Figure 4.31: A diagram showing Inception-ResNet-v2 CNN Architecture which was taken from Inception-V3 on Google Cloud as it adapted from [127].

ResNet-101

Resnet101 is a convolutional neural network that is trained on more than a million images from the ImageNet database [59]. The network has 101 layers deep. It has an image input size of 224-by-224.

DenseNet-201

DenseNet-201 is a convolutional neural network that is trained on more than a million images from the ImageNet database [61]. The network is 201 (or 709 as shown in MATLAB[®]) layers deep. It has an image input size of 224-by-224.

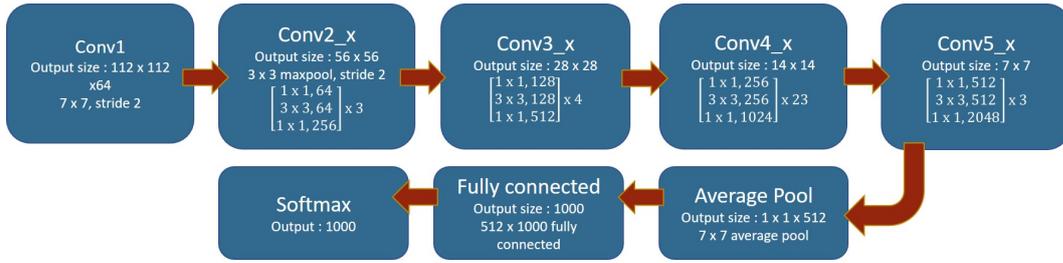


Figure 4.32: A diagram showing ResNet-101 CNN Architecture is adapted from [59].

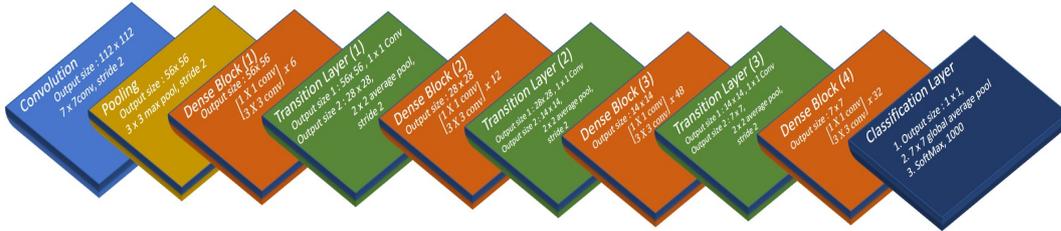


Figure 4.33: A diagram showing DenseNet-201 CNN Architecture is adapted from [61].

Inception-ResNet-v2

Inception-ResNet-v2 is a convolutional neural network that is trained on more than a million images from the ImageNet database [127]. The network has 164 (or 825 as its shown in MATLAB[®]) layers deep. The network has an image input size of 299-by-299.

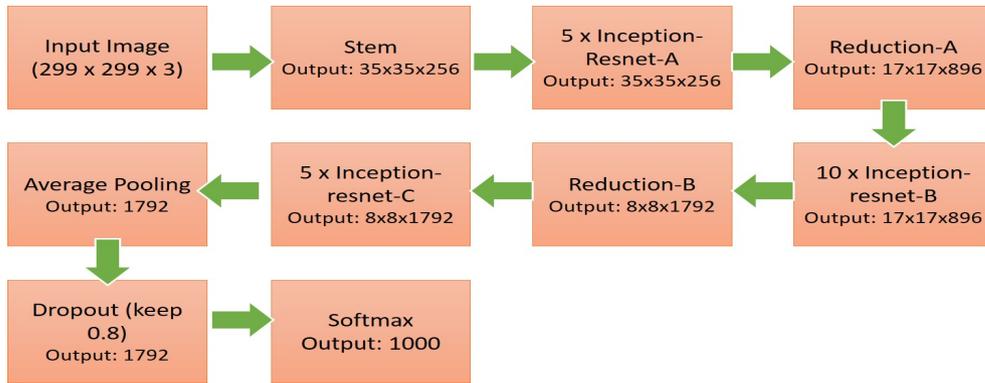


Figure 4.34: A diagram showing Inception-ResNet-v2 CNN Architecture. The diagram is adapted from [127].

All of these architectures are designed for colour images. Modifications were applied to the pre-training CNN architecture so that it can be used for grayscale images which are colour-scaled wrist X-ray images used in this thesis. To address this challenge, several modifications were applied to these architectures:

- Input layer modification. The input layer of pre-trained CNN architecture is

designed to accept three-channel colour (Red, Green, Blue) images as input. A modification was made to the first layer which is usually an input layer to accept single-channel grayscale images. Input The network would adapt to accept one channel input instead of three.

- Transfer learning. Once the architecture is adapted to grayscale images as input, then the network can adjust to retraining the architecture on a grayscale data set. Replace the final layers or the layer that is usually used to classify the image with the classes needed for this research which are normal and abnormal. Sub-chapter 4.1.7 Radiograph Transfer Learning explains further how the transfer learning method is applied to pre-trained CNN architectures.

4.1.6 X-ray data set augmentation

Machine learning architecture with a specific purpose also depends on what data set has been trained for it. The size of the data set is always an obstacle to getting a better result from the machine learning architecture. The medical-related data set is not commonly available in large size due to many reasons such as data security, privacy protection or rarity of the disease or abnormality. On the other hand, the best way to improve machine learning is to train it on more data.

An approach to making a data set seems bigger than reality by creating synthesized data and summing it to the training set [53]. Data augmentation is widely adopted as a solution to the overfitting problem due to limited data available and a common problem in medical image computation [120]. Data augmentation creates various images from existing images within a data set. It is aligned to adapt with a general notion of better results of deep learning architectures relies on more images within data set [124] which is not easy to achieve in the real world, but it could make variations from an existing image within a data set.

Image data augmentation could be done in two ways. First, manipulate the data set using basic image manipulation techniques such as filters, colour space transformations, texture, and other image processing techniques. Second, by using deep learning approaches to make data augmented. This strategy could be adversarial training, neural style transfer or even training the data set through indirect training, in which the data set itself is updated dynamically such as a generative adversarial network [120]. A taxonomy of data augmentation proposed by [120] based on survey

methods could be used to enhance the performance of the architecture where limited data or diverse data is needed.

A tree-structured image data augmentation that illustrates the variation in image data augmentation is shown in Figure 4.35. Kernel filters are a common technique in image processing. The filter directly manipulates the image, which usually aims to sharpen or blur an image. Changing the contrast value of the image is also a kind of kernel filter technique. One popular technique relates to this, such as Canny's edge detector [22] and the Histogram Equalisation method [106]. *Random Erasing* is a technique specifically designed to give a solution to a few data sets by randomly erasing some part of an image within the data set [145]. This technique is believed could enhance image varieties within a limited-size data set. The image has a specific colour space when it has been acquired. Transformed the image's colour space could be counted as the augmentation of the image data. The pixels of an image could be manipulated by using geometric transformation. The result of manipulation could change the visualisation of the image whilst it is the same image but geometrically manipulated. Pixel-based geometric transformations such as rotation, shift, and mirroring are some of the geometric transformation strategies which could be applied to augment the data set.

In this study, the initial dataset contained a small number of images, however, during the study, other datasets were found to be publicly available and relevant to the case study. Mixing images together to create a new image could be counted as the image data augmentation as well as geometric transformations, kernel filters, random erasing, and colour space transformations [120]. Basic image manipulation techniques which can be used to create image data augmentation have been explained in Sub-chapter General Methodologies.

Creating synthesized data to tackle a few data owned to train the deep learning architecture is an exciting approach to the image-based case. This approach is suitable for image and classification problems. Teaching a machine to learn a specific task related to image cases needs a classifier where it takes high dimensional input x and encapsulates it to an identity category y . A new pair of (x, y) could be produced by transforming the x inputs in the training data set.

Data augmentation is also suitable to work with medical image data sets. Radiograph images are obtained under various conditions though it has a standardised acquisition procedure. The augmented data gives a broad possibility of data points.

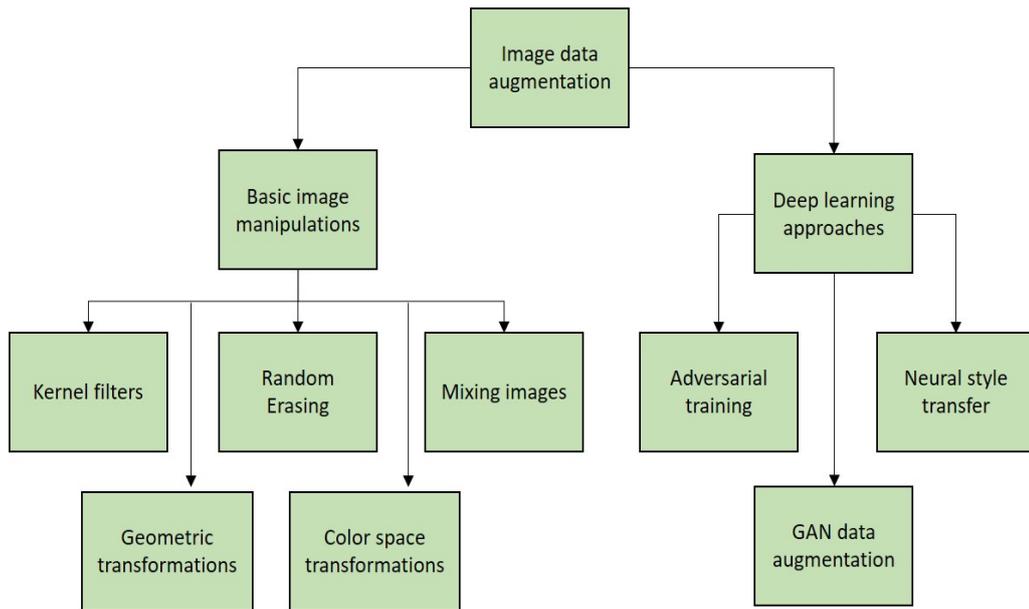


Figure 4.35: An illustration in the tree model of image data augmentation varieties. The graphic is adapted from [120].

It could minimise the gap between the training, validation set, and testing set. An example of augmented images resulting from an X-ray of the wrist is shown in Figure 4.36. An original X-ray image could produce several synthesized images through the augmentation technique. These images in Figure 4.36 can be counted as six new images. When specific portion images within the data set are duplicated through this technique, it can significantly increase the amount of data.

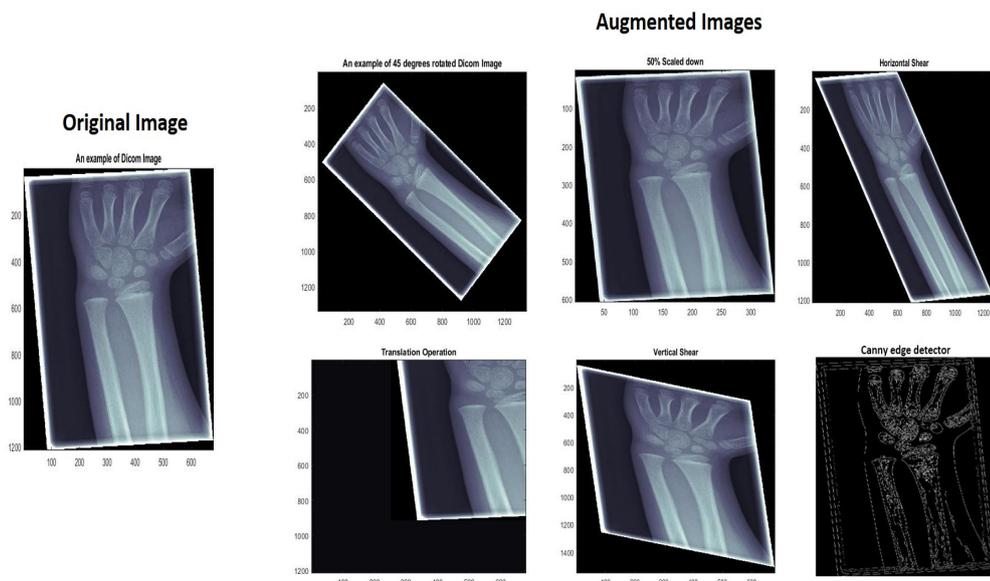


Figure 4.36: An example of the wrist being augmented into several techniques to make a variety of images within the limited size of a data set through image data augmentation strategy.

4.1.7 Radiograph Transfer Learning

Transfer learning is a deep learning method in transferring knowledge and then fine-tuning it to a task from a previous task within a network. The most obvious difference between conventional machine learning and transfer learning is the learning process. Conventional machine learning techniques learn a task from scratch, whilst transfer learning tries to get knowledge from previous tasks and then apply it to a new task [103]. Domain adaptation and knowledge transfer are aimed to exploit what has been learned in a specific setting and then use it to improve results in another setting [53]. It is used as a starting point to learn a new task. It has been a solution to reduce the need for the annotation process. One of the first ideas to use transfer learning was to adopt pre-trained architectures of the ImageNet data set instead of training from the ground [110].

Annotating the image is part of the pre-processing data set before training it to the deep learning network. The annotation process is time-consuming, easily mistaken, and high-cost as well. The transfer learning method was introduced to reduce this pressure. It is believed to help with the problem of limited sources of images within a data set to train [5].

Transfer learning is widely employed from pre-trained architectures in many medical image classification methods. A pre-trained network is being reused as a starting point to learn a new task. Reusing a pre-trained network to get a fine-tuning network with a transfer is usually faster than training the new architecture network that has not been trained. A newly trained network always depends on randomly initialized weights from scratch.

The transfer learning process is illustrated in Figure 4.37. Low-level image features such as edges and colours are learned in the early layers of the network. Specific task features to determine and decide the intention of what network is being trained is learned in the last layers of the network. Final layers replace to adapt the network for new knowledge from the data set. It could be more data or fewer classes and, hopefully, learn faster. By changing the final layers, the network could be trained in new training options and a new data set. A trained network could give predictions, and the network's accuracy assessed whether it fulfills its aim. The performance of the network to determine the abnormality possibility within the image could be seen when the network deployed. These learned features then reiterate to retrain

the network with existing knowledge from the previous iteration.

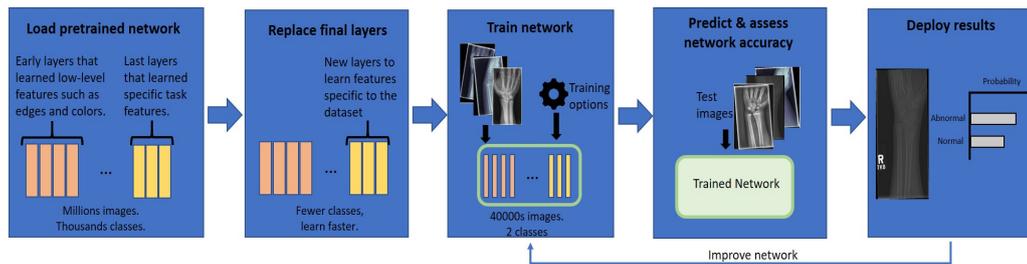


Figure 4.37: A generic illustration of the transfer learning process from a pre-trained convolutional neural network. This image has been adapted from MATLAB[®] documentation materials.

The combination of pre-trained CNN architectures with medical images affects learned features transferred during the transfer learning process. It happens because pre-trained CNN has different ‘knowledge’ than the intended result of working with medical images. Modification by replacing the final layers of the pre-trained network with new layers to accommodate a new data set to be learned gives opportunities to utilise many pre-trained networks to explore in this thesis. Transfer learning is applied in this work. This thesis used eleven pre-trained convolutional neural networks to explore this method. X-ray images were used as the new data set to be trained to the existing pre-trained networks.

4.2 Class Activation Mapping

Class activation mapping generates a visualisation of the intention of the targeted class on the image. This technique was first introduced by [146]. This technique uses global average pooling to perform object localisation without box annotations. Global average pooling on the convolutional feature maps is used for the fully-connected layer that produces the output.

Class activation mapping allows us to visualise the predicted class scores on any given image, highlighting the discriminative object parts detected by the CNN. It benefits from exploiting and visualising what the network thinks when deciding on a subject during its process. In general, the convolutional neural network unit has various layers. Despite a lack of supervision, these layers act as object detectors in the provided location. The pipeline proposed by [146] to generate this map is shown in Figure 4.38.

Global average pooling on the input image to perform class activation mapping is applied to ResNet-50 and Inception-ResNet-v2 at the end tail of `softmax`. This layer

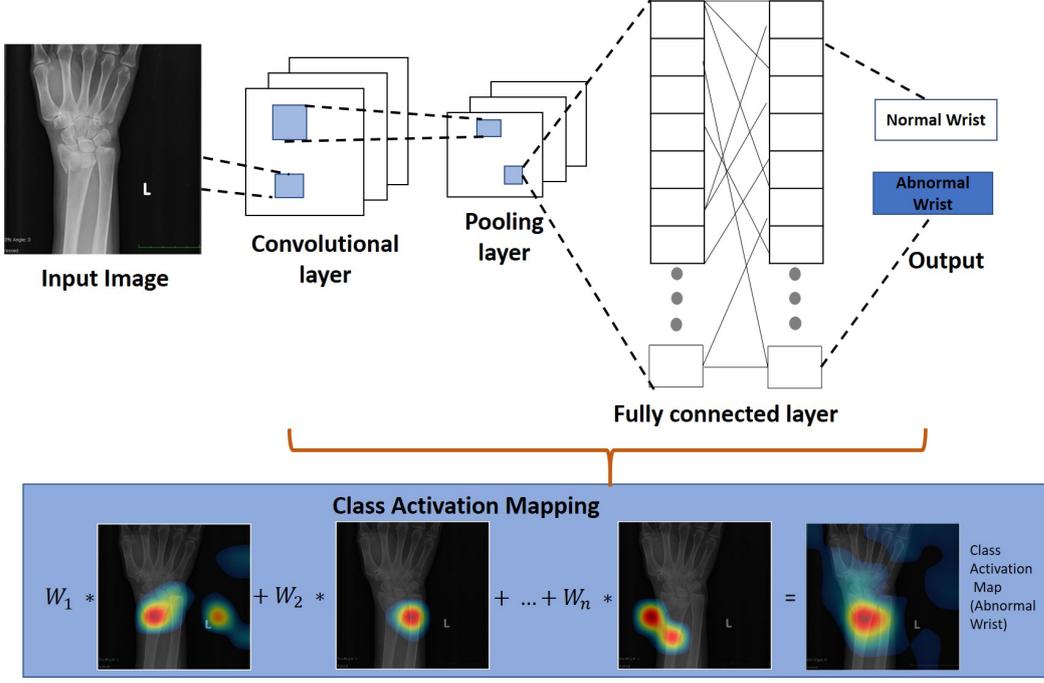


Figure 4.38: Illustration of the Class Activation Mapping. It generates the highlight of the class intention regions within an X-ray image of the wrist.

categorises the image input. As illustrated in Figure 4.38, The last convolutional layer consists of the result of global average pooling which spatial average of each unit feature map.

Based on [146], an input image $f(x, y)$ added with k to represent the activation of unit k at a specific spatial location of a pixel (x, y) . The global average pooling is counted as

$$F^k = \sum_{(x,y)} f_k(x, y) \quad (4.11)$$

For a given class c , the softmax input as S_c where weight of class c for unit k defined as w_k^c , thus

$$S_c = \sum_k w_k^c F_k \quad (4.12)$$

the output of class c is named as P_c where

$$P_c = \frac{\exp(S_c)}{\sum_c \exp(S_c)} \quad (4.13)$$

in [146], the input bias of the softmax set to 0. In their assumption, it has almost zero impact on the performance of classification.

Then the global average pooling of formula $F^k = \sum_{(x,y)} f_k(x, y)$ is join into the class score S_c , then obtain

$$S_c = \sum_k w_k^c \sum_{x,y} f_k(x, y) = \sum_{x,y} \sum_k w_k^c f_k(x, y). \quad (4.14)$$

where [146] define M_c as the class activation map of class c , then each pixel is given by

$$M_c(x, y) = \sum_k w_k^c f_k(x, y). \quad (4.15)$$

As it defines on the softmax of the networks S_c as $\sum_{x,y} M_c(x, y)$. It is obvious that the $M_c(x, y)$ gives an indication of activation importance at the spatial layout. This leads to how class c is decided for the identified image.

This method is a visual pattern at different spatial locations resulting from a weighted linear sum of the presence of f_k . The network has been trained to classify abnormalities within wrist X-ray images in this research. The presentation map allows it to identify the most intention within the image to a specific category that the network has trained.

It provides what the network precisely focuses its attention on. The mapping is represented with a rainbow-coloured or jet colour map where the intensity spectrum ranges from the lowest represented by blue, then green to red as the highest possibility of activation of the targeted class on the image. Exploration of Class Activation Mapping (CAM) as a localiser for abnormality within the wrist X-ray image explains further in Chapter 8.

4.3 Performance evaluation

Performance evaluation is used to determine the possible characterisation of the classification output. The ground truth in this research is the wrist X-ray images that classified Patients as images with abnormalities within the image. The possible characterisation of the output of image classification, given its output and the value of the same image's pixel in the ground truth.

Four possible classifications, depending on the comparisons: (i) True positive (TP) as the pixels where both the output and ground truth show a positive or '1'; (ii) True negative (TN) in pixels where both images show a negative or 0; (iii) False

positive (FP) where the classifiers show a ‘1’ but the ground truth shows a ‘0’, and (iv) False negative (FN) where the classifiers show a ‘0’, but the ground truth shows a ‘1’. In this research case, each pixel whose class is correctly determined by the classifiers will be counted as Correct; every pixel assigned a different class will be considered Incorrect.

In the work by Fawcett [45], two measurements are presented. First, *precision*, which measures how many detected pixels are relevant, by computing the ratio of true positives and the overall number of detected pixels, $TP + FP$. Second, the work presents the *recall* measurement, which computes the ratio of true positives with the sum of the true positives (TP) and the false negatives (FN) — or points that should have been detected, but was missed—, this measurement can be seen as a ratio of the relevant image’s pixels that were detected.

Randen and Husoy [112] present a calculation of *accuracy* which compares the number of correct detection, whether positives or negatives against the overall number of image’s pixels.

The accuracy as Ac , precision as $Prec$, and recall are calculated in the following way.

$$Ac = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (4.16)$$

$$prec = \frac{TP}{TP + FP} \quad (4.17)$$

$$recall = \frac{TP}{TP + FN} \quad (4.18)$$

Cohen’s Kappa (κ) [24] is also calculated as it is the metric used to rank the MURA challenge [111]. This method is considered more robust as it takes into account the possibilities of random agreements [81]. Cohen’s Kappa κ is calculated in the following way. With

$$Tot = (TP + TN + FP + FN), \quad (4.19)$$

being the total number of events, the probability of a ‘yes’ or TP is

$$P_Y = (TP + FP)(TP + FN)/Tot, \quad (4.20)$$

the probability of a 'no', or TN is

$$P_N = (FN + TN)(FP + TN)/Tot, \quad (4.21)$$

and the probability of random agreement $P_R = P_Y + P_N$, then

$$\kappa = (Ac - P_R)/(1 - P_R). \quad (4.22)$$

Chapter 5

Wrist Fractures Semi Automatic Analysis

5.1 Introduction

This chapter explores a geometric semi-automatic image analysis algorithm to analyse and compare the X-rays of healthy controls and patients with dorsally displaced wrist fractures (Colles' fractures) presents. A series of 161 posterior-anterior radiographs from healthy controls and patients with Colles' fractures were acquired and analysed. The semi-automatic analysis consisted of the manual location of three landmarks (finger, *lunate* and *radial styloid*) and automatic processing to generate 32 geometric and texture measurements. These three landmarks have been chosen because these could have been related to a certain condition such as wrist swelling and *osteoporosis* [17, 66, 139]. This work has been published as a journal [114].

The main objective is to determine if there are geometric differences between the successful and unsuccessful cases of Manipulation under Anaesthesia (MuA). The MuA is the main procedures for Colles' fractures and open surgery [11]. The use case of pre and post the MuA classification is to determine if the procedure was successful or unsuccessful. The semi-automatic comparisons extracted a series of measurements, e.g. widths of forearm and metacarpal, based on three manually-placed landmarks. In particular, texture measurements at the radial bone were also explored.

5.2 Methods

This study was using one hundred and sixty-one posterior-anterior radiographs of wrist fractures. The activity of Manipulation under Anaesthesia had supervised or undertaken by fully accredited emergency clinicians. These wrist X-ray images were acquired by clinical experts and were approved by the Health Research Authority through the Integrated Research Application System (IRAS). Data were anonymised which followed ethical procedures with Caldicott Guardian approval, from the Royal Devon and Exeter Hospital.

The data set is divided into two categories which are Patients and Normals. The controls corresponded to patients who required wrist radiographs, mostly following injury, to rule out fractures. As these cases did not present fractures, they were considered as healthy and used as controls.

The wrist fractures were divided by the acquisition time: before (Pre) or after (Post) MUA and the outcome of these: successful or unsuccessful therefore creating four classes pre-successful (n=50), pre-unsuccessful (n=31), post-successful (n=40), post-unsuccessful (n=18). These cases and the clinical outcome were retrospectively identified from electronic attendance logs and electronic records.

X-ray is the modality in this work. Images were obtained with five X-ray units which were DigitalDiagnost DidiEleva01 (Philips Medical Systems, Netherlands), Mobile tablet workstation (Thales, France), DirectView CR 975 and CD 850A (Kodak, USA), Definium 5000 (GE Healthcare, USA) with a variety of exposure factors and saved in DICOM format [90].

Images were obtained in various conditions. Most of those conditions are collimators and many types of angle degrees wrist position. Despite common positions being Postero-Anterior and Lateral but those could come in many types of PA and Lateral positions. So that the first pre-processing step removed the lines caused by the collimator and then aligned the forearm vertically. DICOM and its metadata were read and converted to the binary data container format that the MATLAB program uses which is the integrated development environment for this work.

The lateral position of the wrist contains the forearm bone. Area selection of the forearm bone was then selected as a focal point to align the image vertically. Identification of bone line to indicate the orientation of the arm was determined by using Canny edge detector [22]. How far is rotation required to align the image,

Hough’s transform [60, 37] was used in this matter. The collimator is detected as pixels with zero value and differentiated with the region inside the line of the collimator, which has a pixel value above zero. These regions are then processed with dilation operation and removed from the image.

The wrist X-ray images were analysed based on three manually chosen unique feature markers within each image. These three markers are *lunate*, middle finger, and *radial styloid*. Unique feature markers or landmarks of the wrist’s bone were used to extract bone textures. Each landmark was used to analyse its texture with a strong belief in its correlation with a clinical condition. *Lunate*’s textures was used to determine the typical condition of swelling. The middle finger was used to extract texture that correlates to bone thickness. This could correlate to an indication of *osteoporosis*. The third landmark, *Radial Styloid*, was used to extract texture that correlates with fracturing in the wrist area.

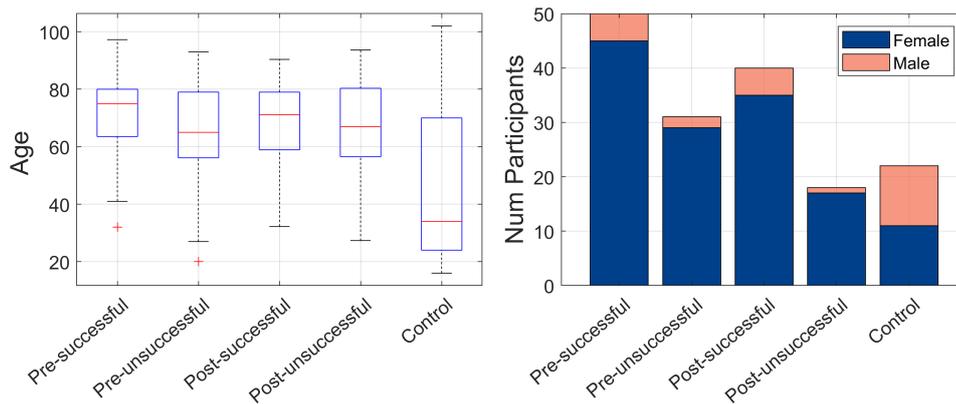


Figure 5.1: Graphical illustration of the demographic distribution of the population of the study. (a) The age distribution is shown with boxplots, one per group of the study. (b) Female and Male distribution is shown with bars, one per group of the study.

Variation of pixel values within observed bone could be seen as the intensity of bone texture. It was analysed through line profile intensity of the bones. The line is a trace between landmarks of the Lunate and the *Radial Styloid*. This step gained 30 degrees and 45 degrees lines from the *Radial Styloid* up to the edge of the radius. A drastic intensity drops into a darker region between bones is detected as the edge. Measurements were extracted both from the intensity profiles and also the profiles after these were adjusted by removing the slope with the idea that measurements like the standard deviation would not be biased by a line that increases its intensity. This analysis provided 10 measurements, e.g. length, slope, and standard deviation

of the profile.

The measurements are analysed through five groups within the data set. These groups are Pre-successful, Pre-unsuccessful, Post-Successful, Post-unsuccessful, and Control. The total measurements from this texturing process are 32. Lunate-based landmarks consist of 10 measurement categories. Finger-based measurement results in two measurements which are values from the Trabecular area and width of the finger. Local Binary Patterns are derived from Lunate to *Radial Styloid* parts. Slope profiles were obtained in four types of measurements which are two full lines profiles, two short segment profiles, and three standard deviation profiles. Two distance profiles are also obtained from *Radial Styloid*. These 32 measurements are the basis for the geometric analysis of wrist fracture radiographs.

For each of the measurements, the statistical difference between the following cases was tested with paired t-tests: (i) healthy controls against patients. (ii) Pre-intervention (successful and unsuccessful) against post-intervention (successful and unsuccessful). (iii) Successful against unsuccessful. (iv) Pre-intervention successful against pre-intervention unsuccessful. (v) Post-intervention successful against post-intervention unsuccessful. (vi) Pre-intervention successful against post-intervention successful. (vii) Pre-intervention unsuccessful against post-intervention unsuccessful.

5.3 Results

Geometrical analysis of wrist X-ray images was obtained through those 32 texture-based measurements of five groups.

Pre-processing steps are applied to remove objects that are unnecessary to the analysis. Removed the lines caused by the collimator and then aligned the forearm vertically. The forearm bones were selected by dividing the image into three parts and selecting the central region of the radiograph as the obvious part of the bones. It was then rotated at certain degrees to make the wrist look standing still. The edge was detected by applying the Canny edge detector [22]. It gives the bone's line and uses it to indicate the arm's orientation. It is not enough that only Canny is used to detect line orientation, then Hough's transform [37] is applied as well. The maximum peak value resulting from the Hough transform was used to determine the degrees of rotation required to align the image. The X-ray image's collimator lines

were detected as pixels by the drastic difference between the darkest region inside the collimator lines and beyond it. The outer region(s) of the line is then removed by dilating it. An example of how preprocessing steps has been affected in cases is shown in Figure 5.3.

Profile texture within the image was made by making a line across points of interest. An example of Figure 5.7's three profile lines was made to generate the specific texture along these lines. Lines were made on three points marked with green, red and blue. The green line is traced between *Lunate* and *Radial Styloid*. The red line is also in this directive but using 30 degrees increase from *Radial Styloid* to *Radius* edge. Blue is using a similar strategy but different degrees which are 45 degrees. The grey level intensity of these three profile lines has been observed with and without the slope. Profile lines captured bone intensity variations over a straight line. A drastic dropped in intensity is considered as the edge between bones. Standard deviation is not biased by an increased intensity of a line. This slope was then removed as shown in Figure 5.7.

Abnormalities on the X-ray images relating to the wrist or part of the forearm can be caused by swelling or *osteoporosis*. The boundaries of the forearm have been detected to indicate a swelling within the forearm on the image. The landmark of the lunate (Red dot in Figure 5.4 a) was used to determine the base of the wrist. The region of interest was determined from this point towards the forearm, and the region of the hand was removed.

Blue-coloured landmarks are then being used to get the region of interest from the middle finger part of the hand (Figure 5.5 a). This region's bone (Figure 5.5 b) was vertically aligned, obtained finger, *trabecular* and *cortical* regions edge. Those were calculated by using a vertical projection of the image intensities (Figure 5.6 c).

It was conjectured that the thickness of the cortical and trabecular regions of the bone would be an indication of *osteoporosis* [17, 66, 139]. The width of the finger and the ratio of the trabecular area to the total area were taken as measurements.

The potential correlation of X-ray-based texture measurements with clinical outcomes was given results with two ways of analysis. First, the texture was analysed by using a small region of bone selected from the radius area. The local binary patterns technique was applied to analyse this part. It explored the relationship between neighbourhood pixels within a concentration area. The local binary pattern is focused on the relative intensity relations between the pixels in a small neigh-

bourhood. A neighbourhood size of 3×3 pixels is the base of this texture analysis. The difference between the grey level between the central pixel with its neighbours. LBP's values are shown as histograms in Figure 5.6 c which gives 10 measurements.

The intensity of the bone line profile has given an alternative way to analyse the bone texture. An example is shown in Figure 5.7 gives a visualisation of how intensity profiles are extracted and compared their texture. Figure 5.7 shows profile lines of the Radial Styloid. In the beginning, a line (green) is traced across the lunate (red in Figure 5.7 a) to the radial styloid (green in Figure 5.7 a) landmarks. Two different degrees lines were produced to get texture information when it corresponds to the different angles of lines being taken as shown in Figure 5.7 b. Removed the slope of intensity profiles is shown in Figure 5.7 c.

Each X-ray image within the previously describe five groups has extracted 32 measurements. These measurements are presented in Table 5.1. Measurement results were tested through paired t-tests. Measurements were compared between the Healthy control data set against the Patients' collection of images. Condition before intervention named Pre-intervention and after intervention were also observed. Successful interventions and unsuccessful results were compared. Measurements of images with the pre-successful record were compared with measurements extracted from Pre-unsuccessful. Postconditions of successful and unsuccessful measurements were compared to each other. Pre-successful against Post-successful was tested as well. In the end, Pre-unsuccessful and Post-unsuccessful are the final comparisons in these paired t-tests.

Three anatomic landmarks which are *Lunate*, *Finger*, and *Radial Styloid* were chosen to extract texture-based measurements. *Lunate*-based landmark has extracted 10 measurements. Each measurement corresponds to the bone's width. The trabecular area is a focus for measurements extraction of the finger's landmark which also include the measurement of its width.

The combination between *Lunate* and *Radial Styloid* has resulted in 10 measurements using the Local Binary Pattern technique. *Radial Styloid*'s landmark has given 10 measurements which are three measurements of slope profile, 2 standard deviation intensity profiles, 2 adjusted standard deviation profiles, and two measurements of distance profile.

Boxplot chart on Figure 5.8 shows that between control and patients and both pre-and post-intervention, the difference is able to distinguish. The results of pre-and

post-intervention groups are near to each other.

Table 5.1: Measurements extracted from the radiographs. The third column corresponds to the landmark used to calculate the measurement. Columns 4 – 10 show the p-values result of paired t-tests between different groups. Values lower than 0.05 are highlighted in bold. Abbreviations: Ratio of width line 1 / width line 4 (W1/W4), Local Binary Pattern (LBP), Standard Deviation (Std), Measurement (Msrmt), Landmark (Ldrk), Control (Ctrl), Patient (Pat) Unsuccessful (Unsucc), Successful (Succ), Pre-Successful (Pre-Succ), Post-Successful (Post-Succ), Pre-Unsuccessful (Pre-Unsucc), Post-Unsuccessful (Post-Unsucc)

No	Msrmt	Ldrk	Ctrl v Pat	Pre- Succ v Post- Succ	Succ v Unsucc	Pre- Succ v Pre- Unsucc	Post- Succ v Post- Unsucc	Pre- Succ v Post- Succ	Pre- Unsucc v Post- Unsucc
1	W1/ W4	Lunate	< 0.01	0.01	0.55	0.16	0.51	0.01	0.74
2	W2/ W4	Lunate	< 0.01	0.21	0.96	0.53	0.44	0.09	0.99
3	W3/ W4	Lunate	0.04	0.36	0.37	0.37	0.68	0.19	0.78
4	W5/ W4	Lunate	0.10	0.30	0.38	0.19	0.33	0.94	0.09
5	W6/ W4	Lunate	0.22	0.37	0.39	0.30	0.76	0.88	0.27
6	W7/ W4	Lunate	0.92	0.31	0.41	0.53	0.68	0.49	0.51
7	W8/ W4	Lunate	0.45	0.25	0.65	0.50	0.77	0.54	0.33
8	Min width/ Max width	Lunate	0.22	< 0.01	0.16	0.39	0.41	< 0.01	0.11
9	W1+W8/ W4+W5	Lunate	< 0.01	0.01	0.72	0.21	0.72	< 0.01	0.53
10	W1+W2/ W7+W8	Lunate	0.06	0.50	0.58	0.35	0.60	0.23	0.73
11	Trabecular Area / Total Area	Finger	< 0.01	0.07	0.13	0.11	0.49	0.07	0.43
12	Width Finger	Finger	0.85	0.02	0.66	0.86	0.52	0.04	0.40
13	LBP 1	L+Rad Sty	0.02	0.01	0.28	0.96	0.07	0.01	0.54
14	LBP 2	L+Rad Sty	< 0.01	< 0.01	0.41	0.83	0.09	< 0.01	0.01
15	LBP 3	L+Rad Sty	< 0.01	< 0.01	0.22	0.68	0.13	< 0.01	0.03

5.4 Discussion

This chapter aimed to analyse wrist fractures with minimal user intervention. A series of measurements were obtained to achieve this. Three manually selected body

16	LBP 4	L+Rad Sty	< 0.01	< 0.01	0.14	0.45	0.26	< 0.01	< 0.01
17	LBP 5	L+Rad Sty	< 0.01	< 0.01	0.16	0.46	0.13	< 0.01	< 0.01
18	LBP 6	L+Rad Sty	< 0.01	< 0.01	0.08	0.24	0.30	< 0.01	< 0.01
19	LBP 7	L+Rad Sty	< 0.01	< 0.01	0.17	0.60	0.07	< 0.01	< 0.01
20	LBP 8	L+Rad Sty	< 0.01	< 0.01	0.06	0.22	0.23	< 0.01	< 0.01
21	LBP 9	L+Rad Sty	< 0.01	< 0.01	0.15	0.57	0.11	< 0.01	< 0.01
22	LBP 10	L+Rad Sty	< 0.01	< 0.01	0.46	0.78	0.09	< 0.01	0.01
23	Slope profile 1 (full line)	Radial Sty- loid	< 0.01	< 0.01	0.88	0.29	0.35	< 0.01	0.09
24	Slope profile 2 (full line)	Radial Sty- loid	0.04	< 0.01	0.39	0.78	0.52	< 0.01	0.02
25	Slope profile 1 (short seg- ment)	Radial Sty- loid	< 0.01	< 0.01	0.91	0.53	0.85	< 0.01	< 0.01
26	Slope profile 2 (short seg- ment)	Radial Sty- loid	< 0.01	0.06	0.82	0.74	0.85	0.11	0.29
27	Std pro- file 1	Radial Sty- loid	< 0.01	< 0.01	0.92	0.39	0.59	< 0.01	0.05
28	Std pro- file 2	Radial Sty- loid	< 0.01	< 0.01	0.74	0.84	0.77	< 0.01	0.06
29	Std pro- file 1 ad- justed	Radial Sty- loid	< 0.01	< 0.01	0.50	0.80	0.47	< 0.01	< 0.01
30	Std pro- file 2 ad- justed	Radial Sty- loid	< 0.01	< 0.01	0.10	0.42	0.22	< 0.01	< 0.01
31	Distance profile 1	Radial Sty- loid	< 0.01	0.23	0.78	0.99	0.75	0.28	0.63
32	Distance profile 2	Radial Sty- loid	< 0.01	0.25	1.00	0.98	0.86	0.42	0.38



Figure 5.2: Six representative radiographs that were collected from previous clinical activity at Royal Devon and Exeter NHS Foundation Trust Emergency Department. The images present considerable variability in the quality, positioning of the arm and presence of lines caused by the collimator also caught during X-ray image acquisition. The images were anonymised and metadata such as age, date of acquisition, gender and clinical outcome was available.

markers were selected to get their texture value extracted. Images came in a variety of qualities which resulted in all measurements being validated visually however this did not effect measurement extraction.

Comparisons scenarios to get the statistical differences between groups such as successful and unsuccessful, pre-intervention successful and pre-intervention unsuccessful, and post-intervention successful and post-intervention unsuccessful have shown an indication of interesting perspective. Within the patient groups, the texture features, both those extracted from the profile lines as the LBP features, showed a statistical difference between controls and patients, as well as between X-rays of pre-and post-intervention. Twenty-five of the 32 measurements indicated statistical differences between controls and patients. The local binary pattern measurements were showing distinct results in four out of the seven groups. Similar indications were followed by values from intensity profile measurements.

Results from the Local Binary Patterns could be extended for further studies, specifically texture analysis from different regions. Larger area's texture differences could be diverse and might relate to bone mineral density as reported on [116].



Figure 5.3: Automatic pre-processing of the radiographs. The six representative cases shown were automatically rotated so that the forearm's position aligned vertically. In addition, the artifacts due to the collimator were removed.

None of the measurements indicated a statistical difference between the following MuA groups as shows in Table 5.1 : successful and unsuccessful, pre-intervention successful and pre-intervention unsuccessful, post-intervention successful and post-intervention unsuccessful. Texture-based measurements of fractures within wrist X-ray images have not given significant results between successful and unsuccessful MuA categories.

5.5 Summary

Texture-based measurements which were derived from geometrical calculation within an X-ray image have been experimented with in this chapter. Pixels within an image are sources to explore in image processing and image analysis fields. Works in this chapter have given a prospect of replacing half-manual wrist fracture identification procedures with automatic procedures. The collection of images in this work can be used as the basic learning for convolutional neural networks.

The results were then explored with statistical and machine learning techniques to determine the most discriminant features and consider new features to be extracted [91]. Local Binary Patterns measurements were giving a better performance

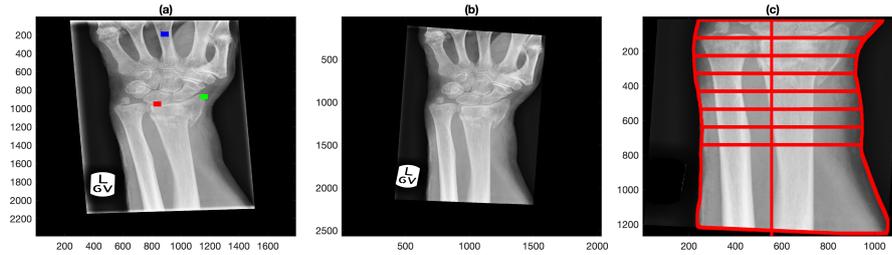


Figure 5.4: Semi-automatic extraction of measurements of the forearm. (a) Original radiograph that presents rotation of the arm and artefactual lines due to the collimator. Three landmarks have been manually located in the base of the lunate (red), *Radial Styloid* (green), and centre of the middle finger (blue). (b) Automatic pre-processing of the image where the forearm was aligned vertically and the lines removed. (c) Using the lunate landmark as a guide, the boundaries of the forearm were automatically delineated and lines were traced between the boundaries. The distance between the lines is 1 cm and was being used to derive swelling measurements of the wrist.

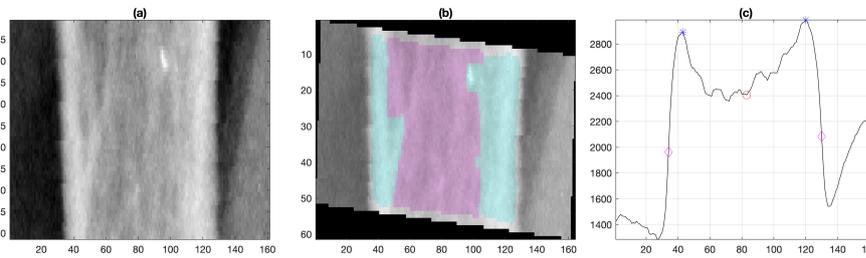


Figure 5.5: Semi-automatic extraction of measurements of the finger (a) Region of interest (ROI) of the central finger generated from the landmark, blue dot in Fig 5.4 (a). (b) Identification of regions of cortical bone (shaded in cyan) and trabecular bone (shaded in pink) from which the ratio of cortical to total area was calculated. Notice that the finger was rotated to align vertically as the previous rotation aligned the forearm but the fingers are not necessarily vertical. (c) Intensity profile of the ROI with the following key points: edges of the bone (magenta diamond), the peak of cortical bone (blue asterisk), and centre of the bone (red circle).

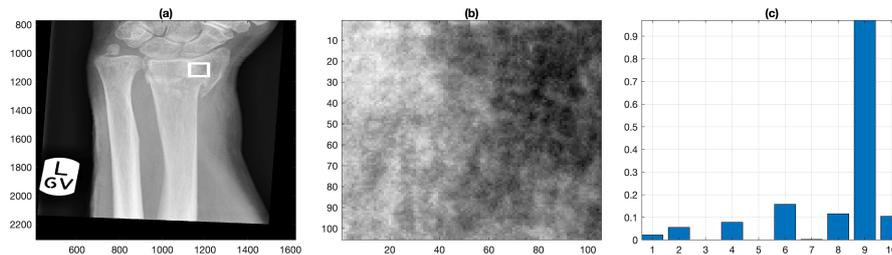


Figure 5.6: Semi-automatic extraction of texture measurements of a region of interest. (a) To analyse the texture of the radius, an ROI is automatically located by traversing a fixed distance from the *Radial Styloid* landmark. (b) Zoom of the region of interest. (c) Texture coefficients generated by Local Binary Pattern analysis.

to indicate differences between pre-intervention and post-intervention among 32 measurements acquired. Combination of two anatomical landmarks of Lunate and

Radial Styloid. This study includes a low number of cases but has given encouraging statistical differences in measurements.

Extraction of image measurements based on its texture gives a channel for further work in combining this with a complex computational technique such as convolutional neural networks. Even though experiment results show a slight difference between pre-and post-intervention. The Local Binary Patterns have given the best results to distinguish differences among compared groups. This technique could be further used to distinguish abnormality within wrist X-ray images and furthermore other areas of bone diseases. This work has been published as an article on [114]. The next chapter describes an analysis of abnormality within wrist X-ray images by exploring this case through a deep learning methods perspective.

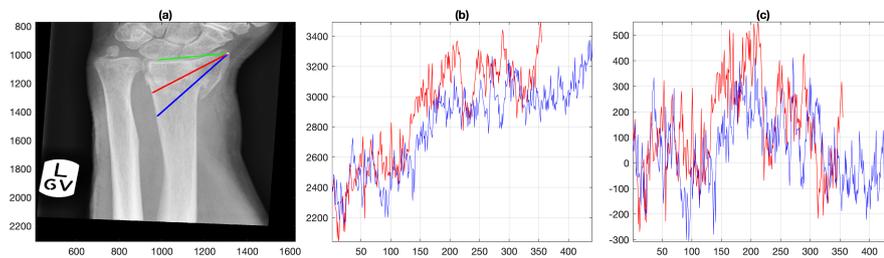


Figure 5.7: Semi-automatic extraction of texture measurements from intensity profiles. (a) Profile lines from the *Radial Styloid*. Initially, a line (green) is automatically traced between the lunate (red in Figure 5.4 a) the *Radial Styloid* (green in Figure 5.4 a) landmarks. Two lines are automatically derived from the first, one at 30 degrees (red) and one at 45 degrees (blue) from the *Radial Styloid* up to the edge of the radius, which is automatically detected. (b) Intensity profiles correspond to the lines traced in (a). Notice the increasing slope. (c) Intensity profiles are adjusted by removing the slope.

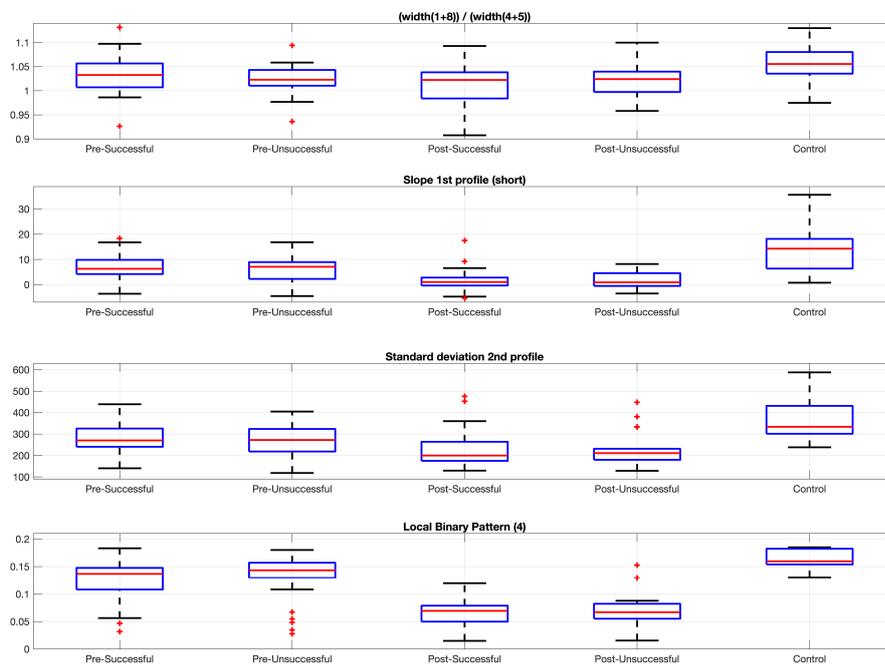


Figure 5.8: Boxplots corresponding to distributions of four representative measurements.

Chapter 6

Wrist Fractures Classification using Deep Learning

6.1 Introduction

In this chapter, the abnormality identification within the X-ray image as the first step to assist a radiologist is explored. A supervised machine-learning technique was used to classify X-ray images of the wrist into normal and abnormal conditions like fractures. This work has been published as a paper at a conference [7].

Identification of wrist abnormality depends on the integrity and function of bones and muscles. The internal condition of this could be checked using an X-ray. The abnormality within an X-ray image is observed by a certain region area of intention. In this work, a method is developed to identify fractures only based on their related region.

The wrist region is a region between the hand and the forearm. An X-ray image taken of this region often includes anatomical features of the forearm and hand. The data sets often came with diverse conditions. In this work, we used a publicly available data set named MURA. Details of this data set have been explained in Chapter 3. The data set was used in this work, particularly the wrist subset of MURA. The data set used in this experiment has labels based on anatomy conditions observed by MURA's expert team. It has two major labels which are positive and negative. Abnormality is labelled as positive and normal is labelled as negative. Detail of these two labels is shown in Table 6.1.

The data set has been trained to classify wrist X-ray images into two categories

which are normal and abnormal. The actual labels from the MURA data set were based on the times of patient visited the clinical facility. Each visit labelled as *Study_X* and if experts diagnosed the image as containing abnormality then a Positive label is added at end of the *Study_Xpositive*. A similar fashion with negative conditions was identified within the investigated X-ray image of a patient.

The strategy of changing the value of hyperparameters and optimisers and mixing domain knowledge of anatomy data sets has been explored in experiments. Computation times were recorded to explore the effect of adding more domain knowledge into the train data set on CNN's performance.

Radiography classification as the foundation of this work takes benefit of convolutional neural networks. Eleven widely-known CNN architectures were used to assess the classification of wrist fractures into these two classes. The classification of wrist radiographs into two categories (Normal / Abnormal) was considered with eleven CNN architectures.

There architectures considered were: GoogLeNet, VGG-19, AlexNet, SqueezeNet, ResNet-18, Inception-v3, ResNet-50, VGG-16, ResNet-101, DenseNet-201 and Inception-ResNet-v2. In addition, the training process of the architecture was tested with different numbers of epochs (10, 20, 30), different mini-batch sizes (16, 32, 64) and with and without data augmentation. The details of the architectures are displayed in Table 4.4. The experiment pipeline is illustrated in Figure 6.2. No pre- or post-processing was applied in any case.

In humans, the *scapula*, *humerus*, and clavicle form the shoulder joint; the *humerus* is the single bone that forms the arm and it articulates with the ulna and radius at the elbow. Each region related to wrist anatomy has been added to the training data set. Wrist fractures have often been seen on other regions of scanned X-ray images such as the hand and forearm. Train networks with many types of anatomic regions as part of strategies to increase network performance but the diversity of its form. We experiment with hyperparameters, various values of epoch, and all three solvers for training networks such as Adam, RMSProp, and SGDM.

The time of computation was recorded as well as the accuracy and area under the curve for each training process. These measures were applied to all regions of the added anatomic region of the training data set.

Fractures surrounding the area of the wrist are not just in a specific location but

also could be seen through another part of upper-body anatomy specifically in the region of the hand and forearm. The wrist itself has two different parts of the angle of image acquisition, as those have been explained in Chapter 2. This is because alike information within different parts of upper body anatomy relates to the wrist. These additional data came also come from the MURA data set. Added data from different regions of anatomy to identify abnormalities within an X-ray image was explored in this chapter.

Table 6.1: MURA’s wrist radiographs are subdivided into four studies.

Wrist-Train data set	Abnormal	Normal
Study 1	3,920	5,282
Study 2	64	425
Study 3	3	45
Study 4	0	13
Total	3,987	5,765
Total Wrist Train Images	9752	
Wrist-Valid data set	Abnormal	Normal
Study 1	287	293
Study 2	5	59
Study 3	3	9
Study 4	0	3
Total	295	364
Total Wrist Valid Images	659	
Total Images of Wrist	10411	

6.2 Methods

This study analysed the wrist radiographs from the public data set Musculoskeletal Radiographs (MURA) [111]. The data set has been manually labelled by board-certified radiologists between 2001 and 2012. The radiographs ($n = 14,656$) are divided into images for training ($n = 13,457$), and validation ($n = 1,199$). Furthermore, the radiographs belong to a group called abnormal (i.e. fracture, foreign body, etc.) ($n = 5,818$) or normal ($n = 9,045$). The distribution per anatomical region is shown in Table 3.4 and selected cases are illustrated in Figure 7.1. Of these, the subset of the **wrists** was selected for this study. In experiments, the actual numbers of data have been checked as it is shown in Table 3.5. Furthermore, this study emphasises in the classification of wrist X-ray images to abnormal and normal categories. Table 3.6 shows the actual distribution of labelled images in the

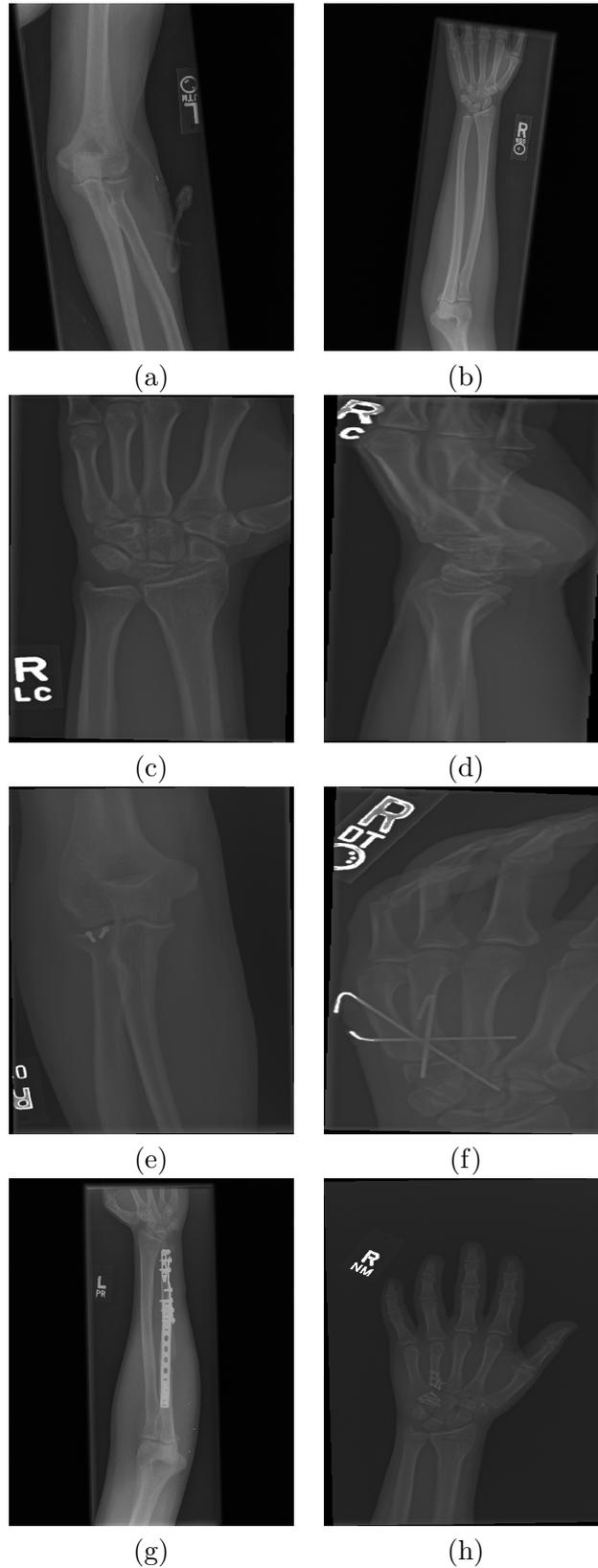


Figure 6.1: Illustration of different radiographs of the Musculoskeletal Radiographs (MURA) data set corresponding to the training set and negative (no abnormalities) in the top row, and positive (abnormalities) in the bottom row. (a) Elbow, (b) Forearm, (c) The Postero-Anterior view of Wrist, (d) Lateral view of Wrist, (e) Elbow, (f) Fingers, (g) Forearm, (h) Hand.

Table 6.2: Summary of convolutional neural networks (CNNs) hyperparameters.

1	GoogLeNet	Optimizer	SGDM	Adam	RMSProp
		Epoch	30	30	30
		Mini batch size	64	64	64
		Init. Learn. R.	0.01	0.001	0.001
		Momentum	0.9000	-	-
		L2 Reg.	0.0001	0.0001	0.0001
2	VGG-19	Optimizer	SGDM	Adam	RMSProp
		Epoch	30	30	30
		Mini batch size	64	64	64
		Init. Learn. R.	0.001	0.001	0.001
		Momentum	0.9000	-	-
		L2 Reg.	0.0001	0.0001	0.0001
3	AlexNet	Optimizer	SGDM	Adam	RMSProp
		Epoch	50	50	50
		Mini batch size	128	128	128
		Init. Learn. R.	0.001	0.001	0.001
		Momentum	0.9000	-	-
		L2 Reg.	0.0001	0.0001	0.0001
4	SqueezeNet	Optimizer	SGDM	Adam	RMSProp
		Epoch	30	30	30
		Mini batch size	64	64	64
		Init. Learn. R.	0.001	0.0001	0.0001
		Momentum	0.9000	-	-
		L2 Reg.	0.0001	0.0001	0.0001
5	ResNet-18	Optimizer	SGDM	Adam	RMSProp
		Epoch	30	30	30
		Mini batch size	64	64	64
		Init. Learn. R.	0.001	0.0001	0.0001
		Momentum	0.9000	-	-
		L2 Reg.	0.0001	0.0001	0.0001
6	Inception-v3	Optimizer	SGDM	Adam	RMSProp
		Epoch	10	10	10
		Mini batch size	64	64	64
		Init. Learn. R.	0.001	0.0001	0.0001
		Momentum	0.9000	-	-
		L2 Reg.	0.0001	0.0001	0.0001

7	ResNet-50	Optimizer	SGDM	Adam	RMSPProp
		Epoch	30	30	30
		Mini batch size	64	64	64
		Init. Learn. R.	0.001	0.0001	0.0001
		Momentum	0.9000	-	-
		L2 Reg.	0.0001	0.0001	0.0001
8	VGG-16	Optimizer	SGDM	Adam	RMSPProp
		Epoch	30	30	30
		Mini batch size	128	128	128
		Init. Learn. R.	0.001	0.0001	0.0001
		Momentum	0.9000	-	-
		L2 Reg.	0.0001	0.0001	0.0001
9	ResNet-101	Optimizer	SGDM	Adam	RMSPProp
		Epoch	30	30	30
		Mini batch size	32	32	32
		Init. Learn. R.	0.001	0.0001	0.0001
		Momentum	0.9000	-	-
		L2 Reg.	0.0001	0.0001	0.0001
10	DenseNet-201	Optimizer	SGDM	Adam	RMSPProp
		Epoch	30	30	30
		Mini batch size	32	32	32
		Init. Learn. R.	0.001	0.0001	0.0001
		Momentum	0.9000	-	-
		L2 Reg.	0.0001	0.0001	0.0001
11	Inception-ResNet-v2	Optimizer	SGDM	Adam	RMSPProp
		Epoch	30	30	30
		Mini batch size	32	32	32
		Init. Learn. R.	0.001	0.0001	0.0001
		Momentum	0.9000	-	-
		L2 Reg.	0.0001	0.0001	0.0001

data set. Each condition is a combination of data labelled as Valid images and data labelled as Train images.

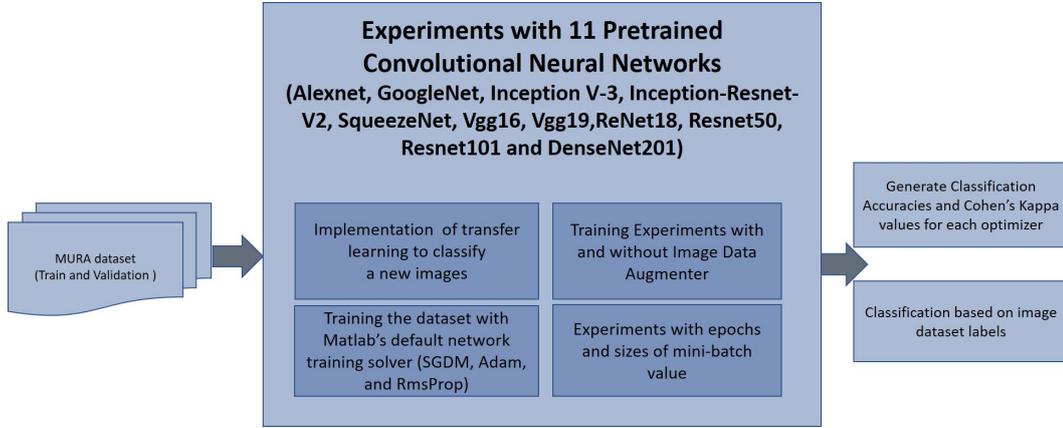


Figure 6.2: Block diagram which illustrates the classification of the wrist radiographs with 11 different Convolutional Neural Network (CNN) architectures. 9752 images from Musculoskeletal Radiographs (MURA) Wrist data set were used for training CNN architectures and 659 images were used for validation. Two different metrics, Accuracy (A_c) and Cohen's Kappa (κ) were computed to assess the performance of 11 pre-trained CNNs. Image data augmentation was used during training and a different number of epochs and mini batch sizes were tested.

A typical CNN combines a series of layers: convolutional layers followed by sub-sampling layers (pooling layer), then another convolutional layer followed by pooling layers, and can continue for a certain number of times after which fully-connected layers are added to produce a prediction (e.g. estimated class probabilities). This layer-wise arrangement allows CNNs to combine low-level features to form higher-level features, learn features and eliminate the need for hand-crafted feature extractors. In addition, the learned features are translation invariant and incorporate the two-dimensional (2D) spatial structure of images, contributing to CNN's achieving state-of-the-art results in image-related tasks [39].

The input to a CNN, i.e. an image to be classified, transits through the different layers to produce the end scores (one score per neuron in the last layer). In the case of image classification, these scores can be interpreted as the probability of the image belonging to each of the classes. The goal of the training process is to learn the weights of the filters at the various layers of the CNN. The output of one of the layers before the last layer, which is fully connected, can be used as a global descriptor for the input image. The descriptor can then be used for various image analysis tasks, including classification, recognition, and retrieval [71].

6.3 Results

Figures 6.3 and 6.4 illustrate some cases of the classification for the lateral and the postero-anterior views of wrist radiographs. In these two figures, images had labelled ‘Study Negative’ and ‘Study Positive’. Label ‘Study Negative’ correlates to a normal condition within the X-ray image, and label ‘Study Positive’ correlates to an abnormal wrist condition within the X-ray image.

Illustration of identified results of Postero-Anterior of the wrist in Figure 6.4 shows trained networks have been able to identify abnormality within the image. An unobvious appearance of fractures as we can see in Figure 6.4 (b) was correctly identified as well as Figure 6.4 (c) which was labelled as Study Negative means it is a normal condition image determined by experts. The other part of this graphic, Figure 6.4 (a) has an obvious fracture but is classified as Study Negative means it normal image by the trained network. Figure 6.4 (d) classified as abnormal despite labelled as a normal image by experts.

The lateral position has a challenging problem where the abnormality does not appear obvious in the whole picture because it was taken from the side position of the wrist. Some samples of results can see in Figure 6.3. Lay out in a similar fashion to Figure 6.4. The trained CNN network wrongly identified an obvious abnormality within the image in Figure 6.4 (a), which has been labelled as positive by the experts. Figure 6.4(b) and (c) images were correctly classified by trained networks but it was not for Figure (d).

General knowledge to enhance the performance of CNN architecture by putting more data into the training data set. The MURA data set has seven parts of a human anatomy data set. Most of this subset of the data set has a relation with wrists such as the hand, forearms, fingers, and elbow. Data are shown in Table 6.5 exposed and the effect of mixing anatomy to the training data set. It was measured by using Area under Curve (AUC) and recording the times of how long did it take to compute the gradually mixed data set.

The growing size of the data set gives a longer time for the machine to finish its training process. Wrist only data set which contains a 10,411 image takes 420 minutes to process and an AUC of 0.8564. Times of computation are getting longer when more sets of parts of the body are put in the train data set where the longest time was 1,800 minutes to process 7 parts of the body data set containing 40,005

X-ray images. Details can be seen in Table 6.5. The Area Under Curve for this multiple-domain knowledge added to the training data set was not hugely affected.

Table 6.3: Results of accuracy for eleven architecture of Convolutional Neural Networks used to classify the wrist images in MURA data set experiments. The best results for each row are highlighted in *italics* and the overall best results are highlighted in **bold**. Abbreviations: Epoch (Ep).

No.	CNNs	SGDM	Adam	RMS Prop	Mean	Ep.	Mini-batch Size
1	GoogLeNet	0.650	<i>0.671</i>	0.640	0.654	30	64
2	VGG-19	0.680	<i>0.681</i>	0.590	0.650	30	64
3	AlexNet	0.674	<i>0.690</i>	0.657	0.674	50	128
4	SqueezeNet	0.683	0.657	<i>0.690</i>	0.677	30	64
5	ResNet-18	0.704	<i>0.709</i>	0.668	0.693	30	64
6	Inception-v3	<i>0.710</i>	0.689	0.707	0.702	10	64
7	ResNet-50	0.686	<i>0.718</i>	0.716	0.707	30	64
8	VGG-16	0.692	0.713	<i>0.716</i>	0.707	30	128
9	ResNet-101	<i>0.715</i>	0.706	0.701	0.707	30	32
10	DenseNet-201	<i>0.733</i>	0.695	0.722	0.717	30	32
11	Inception-ResNet-v2	0.712	<i>0.747</i>	0.710	0.723	30	32

The effect of the number of epochs, size of mini-batch, and data augmentation was evaluated on the classification of wrist radiographs in eleven CNN architectures. Tables 6.3 and 6.4 present the aggregated best results for each architecture in prediction accuracy and Cohen’s Kappa score, respectively.

These eleven CNN architectures have been chosen because of those widely available to access. Details of CNNs used in this work can be referred to Table 4.4. Those architectures represented the range of progressive development of deep learning technology from classic AlexNet to complex and thick network layers of DenseNet-201.

Accuracy results from these eleven architectures show an increasing value. Classic CNN architectures such as AlexNet and GoogLeNet gave 60% accuracy with given training settings. Similar results were resulted by VGG-based CNN architectures such as VGG-19, VGG-16, and SqueezeNet as well.

For the eleven architectures without data augmentation, Inception-ResNet-v2 performs the best with mean accuracy ($Ac = 0.723$) and mean Cohen’s kappa ($\kappa = 0.506$). DenseNet-201 fares slightly lower ($Ac = 0.717$, $\kappa = 0.497$). The lowest results are by GoogLeNet ($Ac = 0.654$, $\kappa = 0.381$). ResNet-50, VGG-16, and ResNet-101 have a similar mean value of accuracy which is 0.707. The epoch value

Table 6.4: Cohen’s Kappa results from eleven architecture of Convolutional Neural Networks used to classify the wrist images in MURA data set experiments. The best results for each row are highlighted in *italics* and the overall best results are highlighted in **bold**. Abbreviations: Epoch (Ep).

No.	CNNs	SGDM	Adam	RMS Prop	Mean	Ep	Mini-batch Size
1	GoogLeNet	0.373	<i>0.412</i>	0.358	0.381	30	64
2	VGG-19	0.433	<i>0.446</i>	0.335	0.404	30	64
3	AlexNet	0.420	<i>0.450</i>	0.390	0.420	50	128
4	SqueezeNet	0.438	0.390	<i>0.448</i>	0.425	30	64
5	ResNet-18	0.474	<i>0.484</i>	0.408	0.455	30	64
6	Inception-v3	<i>0.487</i>	0.450	0.482	0.473	10	64
7	ResNet-50	0.441	<i>0.496</i>	0.494	0.477	30	64
8	VGG-16	0.453	0.491	<i>0.492</i>	0.479	30	128
9	ResNet-101	<i>0.495</i>	0.475	0.472	0.481	30	32
10	DenseNet-201	<i>0.524</i>	0.458	0.507	0.497	30	32
11	Inception-ResNet-v2	0.485	0.548	0.484	0.506	30	32

Table 6.5: The effect of mixed domain knowledge of anatomy data set to the area under curve and times of computation (ResNet-50 case study). Abbreviations: Times of computation (Time of comp), Area Under the Curve (AUC).

No	MURA data set	Total Images	AUC	Times of comp per fold (10 folds)
1	Wrist only data set	10,411	0.8564	420 minutes
2	Wrist+Hand	16,414	0.8244	649 minutes
3	Wrist+Hand+Forearms	18,540	0.8338	724 minutes
4	Wrist+Hand+Forearms+Finger	24,107	0.8333	1032 minutes
5	Wrist+Hand+Forearms+Finger+Elbow	29,503	0.8393	1344 minutes
6	Wrist+Hand+Forearms+Finger+Elbow+Humerus	31,063	0.8415	1406 minutes
7	Wrist+Hand+Forearms+Finger+Elbow+Humerus+Shoulder	40,005	0.8402	1800 minutes

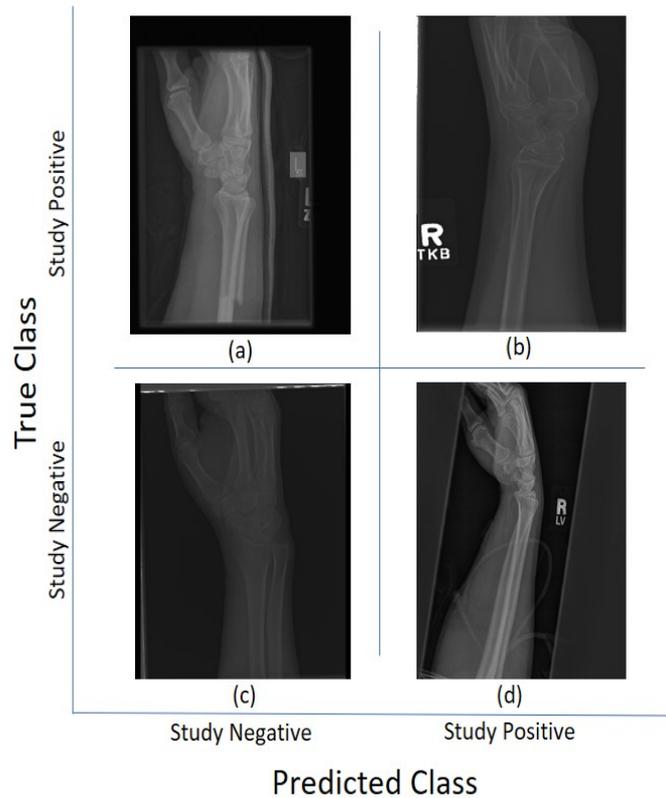


Figure 6.3: Illustration of classification results for Lateral (LA) views of wrist radiographs. (a) Corresponds to positive (abnormal) diagnosis image but predicted as negative (normal), (b) Abnormal diagnosis, and abnormal prediction. (c) Normal diagnosis image and normal prediction. (d) Normal diagnosis and abnormal prediction. Notice that the errors in classification may have been biased by artefactual elements on the images.

was static on 30 and only Inception-v3 used 10 as the value for an epoch. The value of the mini-batch was implemented in several different values which were 32, 64, and 128. Different values of Mini-Batch size were applied between VGG-19 which used 64 mini-batch size and VGG-16 which used 128 mini-batch size. VGG-19's mean accuracy is 0.650 and VGG-16's mean accuracy is 0.707. It results in an increased accuracy output of 0.057. CNN architectures with 32 as the value of mini-batch size were the top three in this experiment leaderboard. Among CNN architectures were used 64 as the value for the mini-batch size, ResNet-50 got the highest outcome of the accuracy of 0.707 but a note is taken that Inception-v3 has the second best result with only 10 epoch and the same value of mini-batch size. The highest value of mini-batch size in this experiment was 128. This was applied to AlexNet and VGG-16. There is an increased output of accuracy of 0.057 between those two architectures. There were three ResNet-based CNN architectures which are ResNet-18, ResNet-50, and ResNet-101. Each of them has reached accuracy values of 0.693, 0.707, and

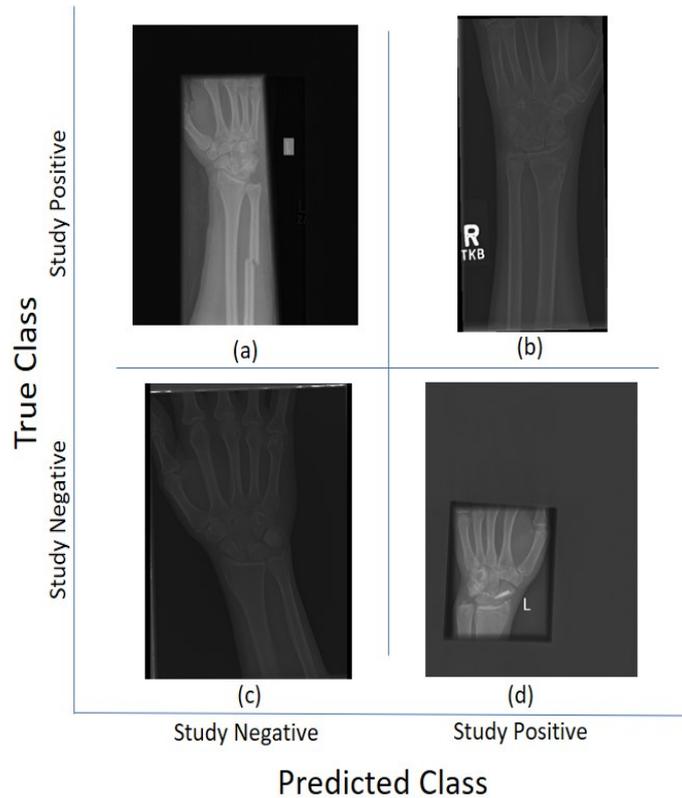


Figure 6.4: Illustration of classification results for Postero-Anterior (PA) views of wrist radiographs. (a) Corresponds to positive (abnormal) diagnosis image but predicted as negative (normal), (b) Abnormal diagnosis and abnormal prediction. (c) Normal diagnosis image and normal prediction. (d) Normal diagnosis and abnormal prediction. Notice again that the errors in classification may have been biased by artefactual elements on the images.

0.707 respectively.

Cohen’s Kappa value for each CNN architecture was calculated for this experiment as well. A summary of Cohen’s Kappa for each CNN result is shown in Table 6.7. The highest score of Cohen’s Kappa among 13 CNN architectures in this table belongs to Inception-ResNet-v2 (augmentation) with Cohen’s Kappa mean of 0.703. The GoogLeNet has the lowest mean score of 0.301.

Data augmentation is implemented for the two best-performing architectures, namely ResNet-50 and Inception-ResNet-v2. These two networks had given results of accuracy and Cohen’s Kappa as well. The ResNet-50 has resulted in mean accuracy of 0.845 and Inception-ResNet-v2 has resulted in 0.857. Complete details are presented in Table 6.6 for accuracy and Table 6.7 for Cohen’s Kappa.

The results in Table 6.6 and Table 6.7 show a significant improvement when data augmentation is added to the training data set. Accuracy (Ac) increases by 0.314 and Cohen’s Kappa (k) by 0.197. In other words, accuracy (Ac) improved by

Table 6.6: Results of accuracy for eleven Convolutional Neural Networks used to classify the wrist images in the MURA data set. The best results for each row are highlighted in *italics* and the overall best results are highlighted in **bold**. Abbreviation: augmentation (aug).

No.	CNN	SGDM	Adam	Rms Prop	Mean	Epoch	Mini-batch Size
1	GoogLeNet	0.650	<i>0.671</i>	0.640	0.654	30	64
2	VGG-19	0.680	<i>0.681</i>	0.590	0.650	30	64
3	AlexNet	0.674	<i>0.690</i>	0.657	0.674	50	128
4	SqueezeNet	0.683	0.657	<i>0.690</i>	0.677	30	64
5	ResNet-18	0.704	<i>0.709</i>	0.668	0.693	30	64
6	Inception-v3	<i>0.710</i>	0.689	0.707	0.702	10	64
7	ResNet-50	0.686	<i>0.718</i>	0.716	0.707	30	64
8	VGG-16	0.692	0.713	<i>0.716</i>	0.707	30	128
9	ResNet-101	<i>0.715</i>	0.706	0.701	0.707	30	32
10	DenseNet-201	<i>0.733</i>	0.695	0.722	0.717	30	32
11	Inception-ResNet-v2	0.712	<i>0.747</i>	0.710	0.723	30	32
12	ResNet-50 (aug)	0.835	0.854	0.847	0.845	30	64
13	Inception-ResNet-v2 (aug)	0.842	0.869	0.860	<i>0.857</i>	30	32

Table 6.7: Cohen’s Kappa results from eleven Convolutional Neural Networks architectures used to classify the wrist images in the MURA data set. The best results for each row are highlighted in *italics* and the overall best results are highlighted in **bold**. Abbreviation: augmentation (aug).

No.	CNN	SGDM	Adam	Rms Prop	Mean	Epoch	Mini-batch Size
1	GoogLeNet	0.373	<i>0.412</i>	0.358	0.381	30	64
2	VGG-19	0.433	<i>0.446</i>	0.335	0.404	30	64
3	AlexNet	0.420	<i>0.450</i>	0.390	0.420	50	128
4	SqueezeNet	0.438	0.390	<i>0.448</i>	0.425	30	64
5	ResNet-18	0.474	<i>0.484</i>	0.408	0.455	30	64
6	Inception-v3	<i>0.487</i>	0.450	0.482	0.473	10	64
7	ResNet-50	0.441	<i>0.496</i>	0.494	0.477	30	64
8	VGG-16	0.453	0.491	<i>0.492</i>	0.479	30	128
9	ResNet-101	<i>0.495</i>	0.475	0.472	0.481	30	32
10	DenseNet-201	<i>0.524</i>	0.458	0.507	0.497	30	32
11	Inception-ResNet-v2	0.485	<i>0.548</i>	0.484	0.506	30	32
12	ResNet-50 (aug)	0.655	<i>0.696</i>	0.683	0.678	30	64
13	Inception-ResNet-v2 (aug)	0.670	0.728	0.711	0.703	30	32

19% and Cohen's Kappa (k) by 39%. About 10 of 13 architectures experimented has Cohen's Kappa below 0.5. Only Inception-ResNet-v2 has got better than 0.5. This was the result before data augmentation was added to the training data set. Afterward, the data set is augmented for the two architectures chosen. Cohen's Kappa score increased to above 0.6 which means better than 11's Cohen's Kappa without data set augmentation applied.

These improved accuracy and Cohen's Kappa also gives a sign that potentially indicates better feature extraction with deeper layer network architectures.

6.4 Discussion

The classification of wrist X-ray images toward properly identified abnormality within the wrist X-ray image. In this work, Deep learning is used to classify X-ray images into normal and abnormal conditions. A normal condition is a condition without any obvious strange visibility in the image. The abnormal condition is a condition of obvious fractures or deformity of the wrist within the X-ray image.

The main focus of this work is the wrist, then the wrist subset data set of MURA at the beginning of this work. We gradually added more images into the Train data set because the wrist part was often captured when another part of the hand, fingers, and forearm was acquired. The perspective of collecting more images will increase the performance network to detect abnormality in general applied in this work.

In this work, eleven publicly available CNN architectures to classify wrist radiographs were studied. The ability to identify abnormality within X-ray images based on learning of a given data set could expose its benefit for fracture image analysis. Various hyperparameters were attempted during the experiments. It was observed that Inception-ResNet-v2 provided the best results ($Ac = 0.747$, $\kappa = 0.548$), which were compared with leaders of the MURA challenge, which reports 70 entries. The top three places of the leaderboard were $\kappa = 0.843, 0.834, 0.833$, the lowest score was $\kappa = 0.518$ and the best performance for a radiologist was $\kappa = 0.778$. Thus, without data augmentation, the results of all the networks were close to the bottom of the table. Data augmentation significantly improved the results to achieve the 25th place on the leaderboard with ($Ac = 0.869$, $\kappa = 0.728$). Whilst this result was above the average of the table, the positive effect of data augmentation was confirmed to be close to the human-level performance.

To improve the results from the standard CNN architectures, such as those analysed in this work, the classification pipelines could include extra steps. Namely:

1. Pre-processing steps, which may consist of Low pass filtering to remove high-frequency noise, and cropping of images to remove excessive background region (notice that some of the incorrect classifications in Figure 6.3 have large background regions). More elaborate pre-processing approaches such as location and orientation of bones [114] could help detect the areas of real interest, and discard any region that may be biasing results, such as the labels for the right or left hand, which is always very bright might be confusing the architectures.
2. Post-processing steps may also be considered, for instance, the association between key features and the predicted classes [130, 98]. Furthermore, the visualisation of key features may be useful to stakeholders (e.g. clinicians or radiologists) who might be more interested in the attributes of the original data rather than the architectures themselves [91].
3. Ensembles or combinations of different configurations may also help increase the results of individual configurations.
4. Finally, adding domain knowledge in terms of knowledge of the anatomical region (i.e. elbow or hand) with the possible cases (i.e. fracture or implant) may allow the fine-tuning of the architectures to detect not only an abnormality but the type of abnormality and the location of this.

6.5 Summary

This section describes the classification of wrist fractures explored through eleven convolutional neural network architectures. Modifying the data set where mixed body parts were performed to better identify wrist fractures. Each part of the body data set was added. The time of computation and value of the area under the curve is then collected through this process. Eleven CNN architectures were used to train wrist-only data sets and similar training settings. It shows that results in accuracy and Cohen's Kappa of Inception-ResNet-v2 have the best result of all eleven other CNN architectures.

It is shown that the perspective of gradually adding more images into the train data set does not significantly increase the performance of the network even though

it still has a good outcome for the area under the curve. But this perspective cannot be obeyed because the wrist part is often captured when other parts related to the hand are captured, such as the hand, fingers, and forearm. This work has been published as a conference article [7].

The next chapter describes explorations of data augmentation to enhance CNN's performance and class activation mapping of trained deep learning architectures to visualise the area of suspect abnormality within an X-ray image.

Chapter 7

Class Activation Mapping for Wrist Fractures Image Analysis

7.1 Introduction

This chapter explores the convolutional neural networks' intention to decide abnormality within an image. The process of identifying abnormalities within the X-ray image through a training data set turned out to be possible. Despite the classic problem of a limited quantity of images, from the deep learning point of view, this problem could be optimised using the data augmentation technique. Class activation mapping was used to exploit visual prediction from the trained network when deciding the conditions of a wrist X-ray image. Part of the works in this chapter has been published as an article on [8].

Experiments in Chapter 7 have allowed further exploration of deep learning to identify abnormalities within a wrist X-ray image. Medical image data sets always have quantity limitations and often come with unstructured conditions. The learning process of this limited data set in this chapter has been explored by applying the data augmentation strategy. The aim of using the data augmentation strategy is to increase the accuracy as well as Cohen's Kappa resulting from previous experiments in chapter 7. The data augmentation technique was applied to the top two best-performance CNNs. These two are ResNet-50 and Inception-ResNet-v2.

Adam, SGDM, and RMSProp are used as optimizers for CNN. Summary details of hyperparameters for this work show in Table 6.2. Hyperparameters were applied to each CNN architecture. This setting was similar to previous experiments.

After training CNN using specific hyperparameters and data augmentation then accuracy and Cohen’s Kappa was calculated. ResNet-50 and Inception-ResNet-v2 were given a better performance than without data augmentation applied.

Class activation mapping (CAM) was used to observe what the CNN actually focus on in deciding which part within identified the image classified as abnormal or normal. The CAM-based visualisation depends on a specific layer to observe the location of abnormalities. The CAM-based visualisation was applied to the final convolution layer in this case. It was generated at layer ‘activation_49_relu’ for ResNet-50 and ‘conv_7_bac’ for Inception-ResNet-v2, respectively.

Visualisation using Class Activation Mapping gives us a graphical understanding of CNN’s performance. A comparison between before and after data augmentation modification could be used for further analysis. It gives us a window into what exactly this network learned through a given data set.

7.2 Materials

The data set for this experiment was obtained from the public data set MURA from Stanford University. This data set has been manually labelled by board-certified radiologists. The radiographs studies ($n = 14,656$) are divided into images for training ($n = 13,457$), and validation ($n = 1,199$). Furthermore, the radiographs belonging to group studies have been allocated in groups called abnormal (i.e., those radiographs that contained fractured bones, foreign bodies such as implants, wires or screws, etc.) ($n = 5,715$) or normal ($n = 8,941$).

We also used a second data set specifically acquired for wrist fracture research. It consists of 2 labels named Normals ($n = 122$) and Patients ($n = 882$). Image labels were changed where Normal to Negative and Patients to Positive. The wrist X-ray image labels were changed because the data set came in different names but in general, could be simplified into two categories of labels which are Negative and Positive. Negative labelled images mean images do not have any abnormality within them. Positive labelled images mean images have abnormality within.

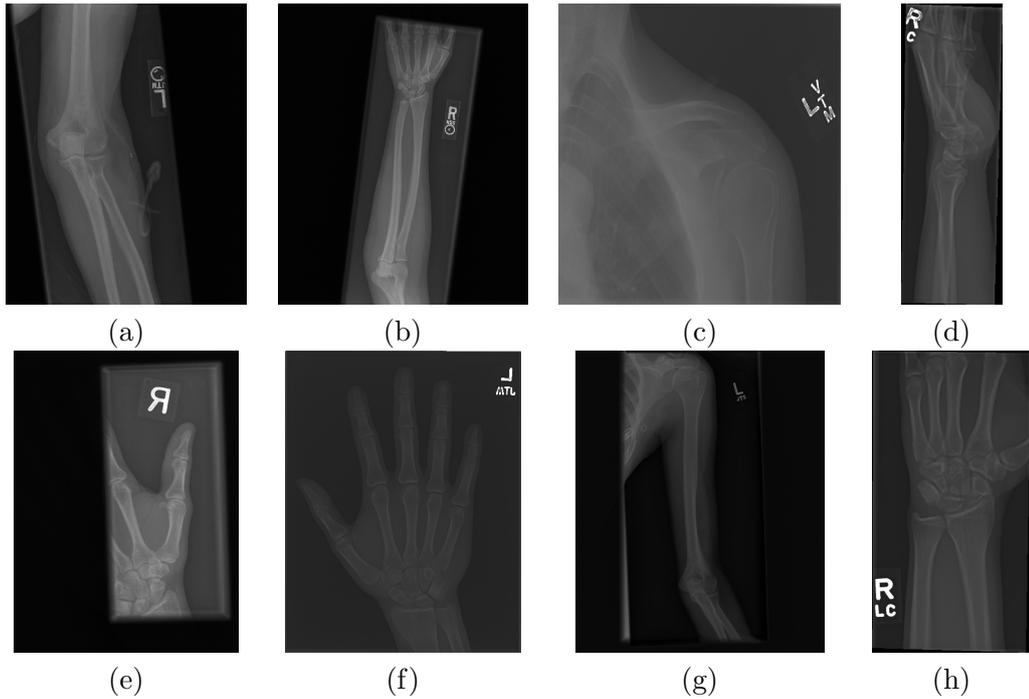


Figure 7.1: Eight examples of radiographs without abnormalities (considered negative) of the Musculoskeletal Radiographs (MURA) data set. (a) Elbow, (b) Forearm, (c) Shoulder, (d) Wrist (lateral view), (d) Lateral view of Wrist, (e) Finger, (f) Hand, (g) Humerus, (h) Wrist. It should be noted the variability of the images in terms of dimensions, quality, contrast, and the large number of labels (i.e., R for right and L for left), which appear in various locations.

7.3 Methods

In general, there are several strategies to increase the accuracy of CNN such as data augmentation, changing values of hyperparameters, and increasing the size of the data set means bringing more related images into the data set.

Data augmentation means making a variety of current data set we own. It includes changing the value range of rotation, reflection, and shear. Each image within the data set was rotated from -5 to 5 degrees, reflected through x axis and y axis, shear it through x -axis of -0.05 to 0.05 and y -axis of -0.05 to 0.05 .

We also experiment with hyperparameters values such as epoch values, batch size values, and initial learning rate. We also have made changes to the data set such as pre-processing it with Contrast Limited Histogram Equalisation (CLAHE) and its label structure. Image translation technique was not applied because images were captured with a good range of translation shifts.

During experiments, combining data sets is a strategy to increase image variety within the data set. In this case, the university’s owned acquired data set combined with MURA’s subset data set of the wrist. The university’s data set has not had

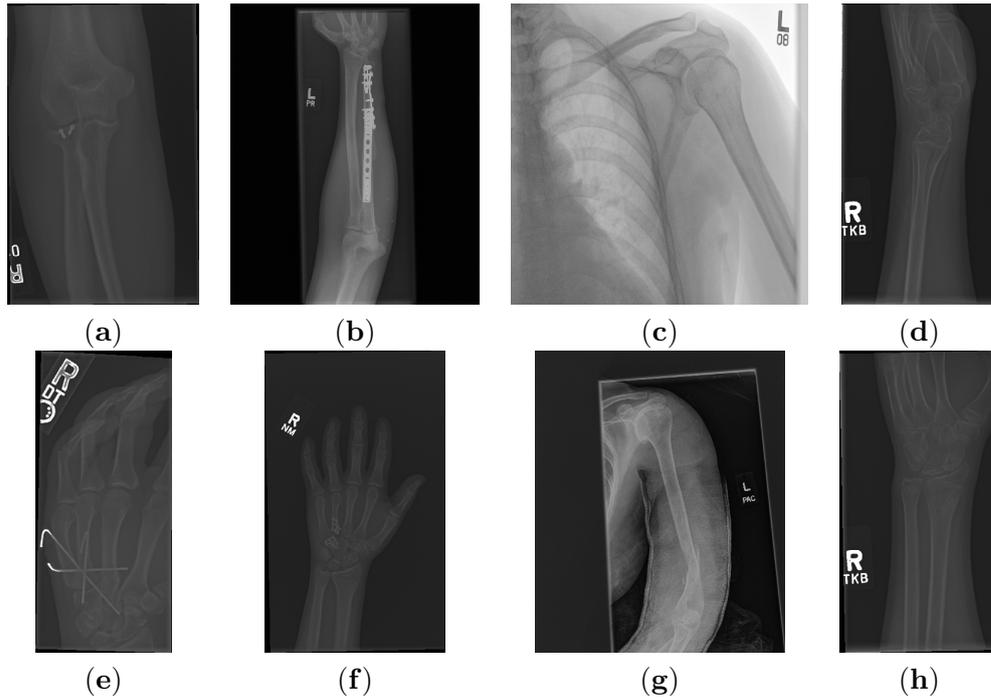


Figure 7.2: Eight examples of radiographs with abnormalities (considered positive) of the Musculoskeletal Radiographs (MURA) data set. (a) Elbow, (b) Forearm, (c) Shoulder, (d) Wrist (lateral view), (e) Finger, (f) Hand, (g) Humerus, (h) Wrist (PA view). As for the cases without abnormalities, it should be noted the variability of the images and in addition the abnormalities themselves. There are cases of metallic implants some of which are smaller (a) than others (b), as well as fractures.

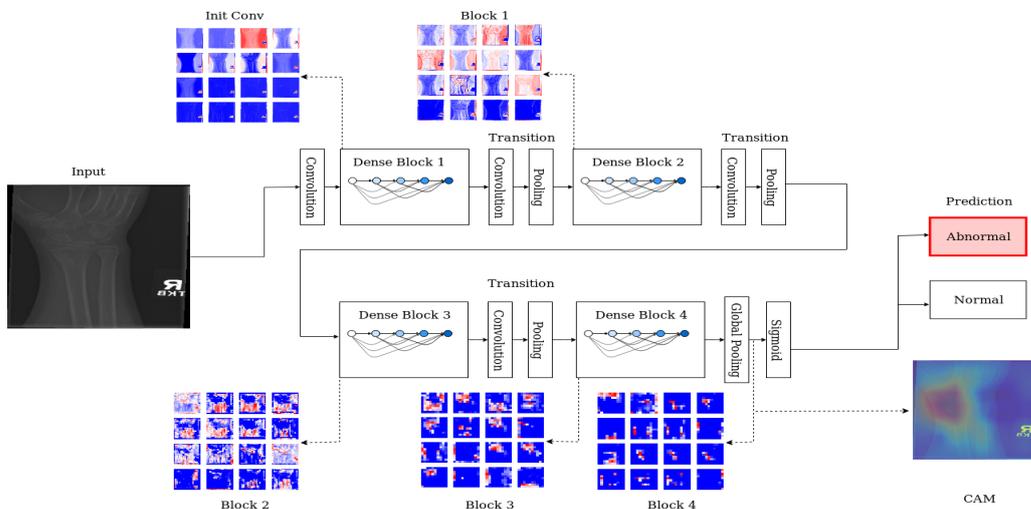


Figure 7.3: Illustration of Class Activation Mapping which generates the highlight of the class intention regions.

many images but important to be added to the experiment because this small-size data set focuses on research we carrying on which is wrist fractures.

The data set was also modified by applying a data augmentation technique. This technique was performed for both data sets. The training process for the mixed data

set strategy had performed in all three optimizers (SGDM, Adam, and RMSProp).

Class activation mapping is applied to visualise the actual intention decision of the network of images identified. The targeted class is based-on labels from the training data set. The visualisation provides an indication of what the network actually focuses on. It generates from the last layer of the convolutional network. The CAM representations were generated at layer ‘activation\49\relu’ for ResNet-50 and ‘conv\7\bac’ for Inception-ResNet-v2, respectively. The CAM’s maps were up-scaled to the input resolution and overlaid on top of the original radiography for the location of the abnormalities.

Schematic illustration of Figure 7.3 shows the X-ray classification process and class activation mapping through layer-wise activation maps across different dense blocks. At each level, a series of feature maps are generated, and the resolution decreases progress through the blocks. Colours indicate the range of activation: blue corresponds to low activation, and red for highly activated features. The final output, visualised here using Class Activation Mapping, which highlights the area(s) where abnormalities can be located.

Accuracy was calculated as the prediction among the total number of cases examined. Cohen’s Kappa was also calculated, as it is the metric used to rank the MURA challenge [111] and because it also takes an account of random agreements possibilities makes it considered more robust.

Experiments were implemented with several scenarios of training options. Each training option was used to observe its ability to produce higher accuracy and Cohen’s Kappa. Each training solver was implemented with a similar exact setting for consistency reasons.

Data augmentation strategy to enhance network performance was evaluated as well. The following augmentations have been performed to each of the training images: (1) rotations of (-5° to 5°), (2) vertical and horizontal reflections, (3) shear deformations of (-0.05° to 0.05°) in horizontal and vertical directions, and (4) contrast-limited adaptive histogram equalisation (CLAHE) [105]. Translations were not applied as the training images were captured with a good range of translation shifts. Settings of hyperparameters were referred to Table 6.2. Adam and RMSProp have similar hyperparameter settings and SGDM has a slightly different ‘Initial Learning’ value which is 0.001 rather than 0.0001 of the other two. This is the default value of the ‘Initial Learning’ value from these optimisers. Complete settings

of hyperparameters are shown in Table 6.2. Epoch values and mini-batch sizes were referred to Table 4.4.

Due to the hardware limitations, the number of folds increases little by little. In the beginning, it started with only 2 folds then increase to 5 and 10 folds. Each fold will transfer its learned network to a new one to increase the final network's accuracy and Cohen's Kappa we assumed. In the end, we hope this experiment will answer our question about the size of the data set and will enhance the network's ability to identify abnormalities within the X-ray image.

Several images were tested to understand visually how the network determined abnormality within the identified wrist X-ray image. The images tested by the CAM technique are valid labelled images from both the dataset and random images taken from the Internet. Then, the performance results between optimizers, combined dataset strategies, and variation in hyperparameter values are analysed.

7.4 Results

The results of class activation mapping (CAM) were visualised and overlaid on top of the representative images. CAM is applied to those images in Figures 7.1 and 7.2. The focus of the CAM's heat map on images on both figures is on the 'suspected' abnormality within each image. The cases in Figure 7.2 reflected positive abnormality within images to trigger red-coloured activation region to get the most attention by the networks. On the other side, Figure 7.1 is used to show how networks pay attention to the region where the CAM is being activated to determine whether the X-ray image is without abnormality within.

The result summary of overlaid CAMs is shown in Figure 7.4 and Figure 7.5. Visualisations of CAM for ResNet-50 could be seen on Figure 7.4 (a) and and Figure 7.5 (a). Inception-ResNet-50's result of CAM-based visualisations are on Figure 7.4 (b) and Figure 7.5 (b). All images have been able to detect based on what should be decided. CAM visualisation is focused on suspected abnormality within the image and then gives a prediction of the detected image. All images give results of prediction over 90%.

Inception-Resnet-v2 presented a more focused and smaller area of activation maps. It should also be noted that whilst for correct classifications, the highlighted regions are similar, for some incorrect classifications Figure 7.5(a,b, top left and

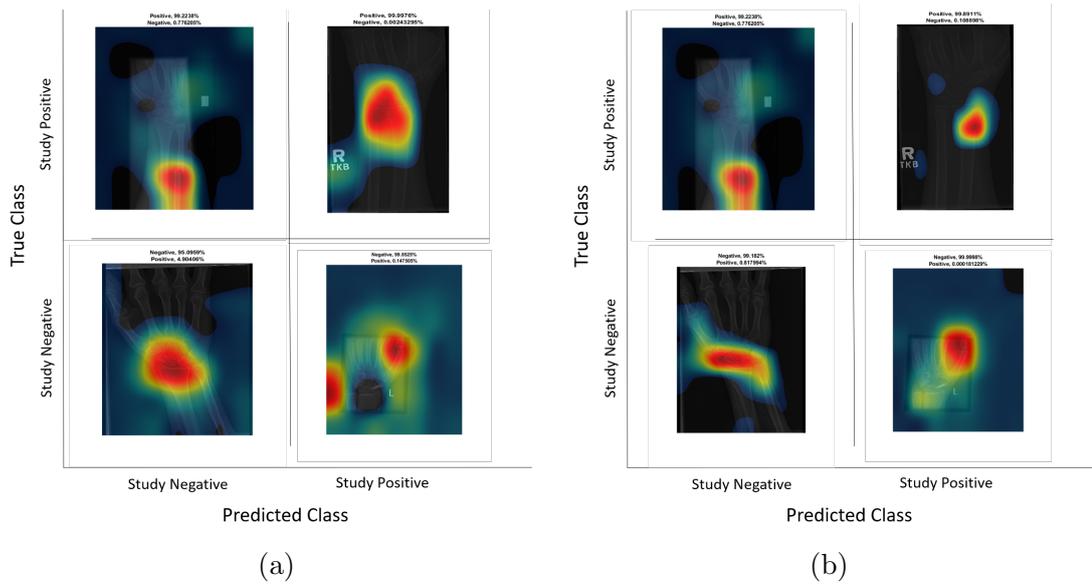


Figure 7.4: Illustration of the Class Activation Mapping overlaid on the four classification results for (a,b) Postero-Anterior views shown in Figure 6.4.

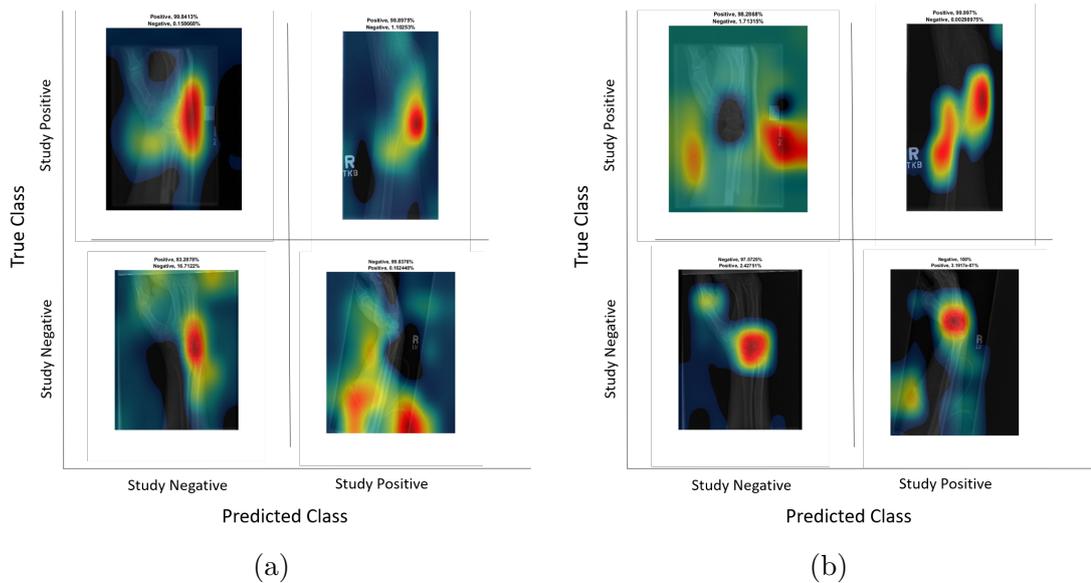


Figure 7.5: Illustration of the Class Activation Mapping (CAM) overlaid on the four classification results for (a,b) Lateral views shown in Figure 6.3

bottom right), the activation is quite different, which suggest that the architectures may not be confusing salient regions that are not related with the condition of normal or abnormal.

The CAMs obtained for ResNet-50 are shown in Figures 7.9 and 7.6, whilst those for Inception-ResNet-V2 are shown in Figures 7.7 and 7.8. In all cases, the CAMs were capable of indicating the region of attention used in the two architectures applied. This is especially valuable for identifying where the abnormalities are in

Figures 7.6 and 7.8.

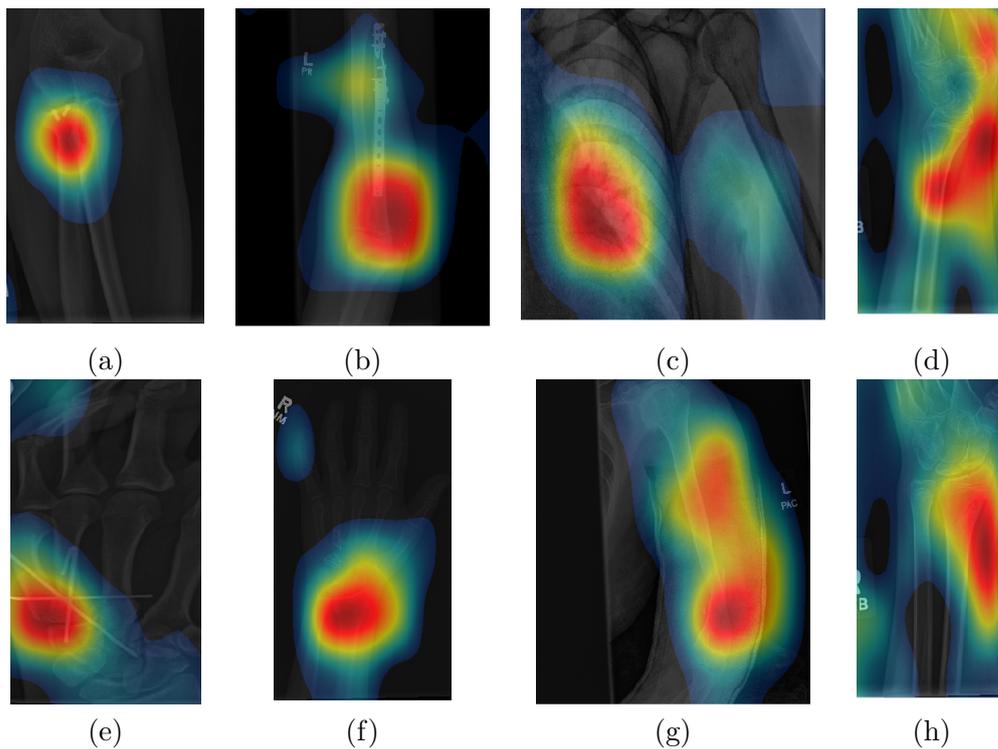


Figure 7.6: Illustration of activation maps overlaid over the eight radiographs with abnormalities of Figure 7.2 to indicate the regions of the image that activated a ResNet 50 architecture. (a) Elbow, (b) Forearm, (c) Shoulder, (d) Wrist (lateral view), (e) Finger, (f) Hand, (g) Humerus, (h) Wrist. The activation maps illustrate the location of the abnormalities, e.g., (a,e), but appears spread in other cases (b,g) where the abnormality is detected together with a neighbouring region. In other cases (c), the abnormality is not detected.

Whilst both architectures indicate similar regions of attention, Inception-ResNet-v2 appears to have smaller attention regions (i.e., more focused) than those in ResNet-50. This may indicate a better extraction of features in the Inception-ResNet-v2 leading to a better prediction. The class activation mapping corresponding to Figures 6.3, and 6.4 are presented in Figures 7.4 for PA view and 7.5 for Lateral view.

A small set of X-ray images consisting of 30 wrist radiographs was selected and analysed using CAM. Figure B.1 shows images with class activation maps triggered from ResNet-50. The same 30 images were also evaluated in Inception-ResNet-v2 and the results are shown in Figure B.2.

7.5 Discussion

The class activation mapping technique provides a way to visualise what the convolutional neural network has learned to classify an X-ray wrist image. Thirty

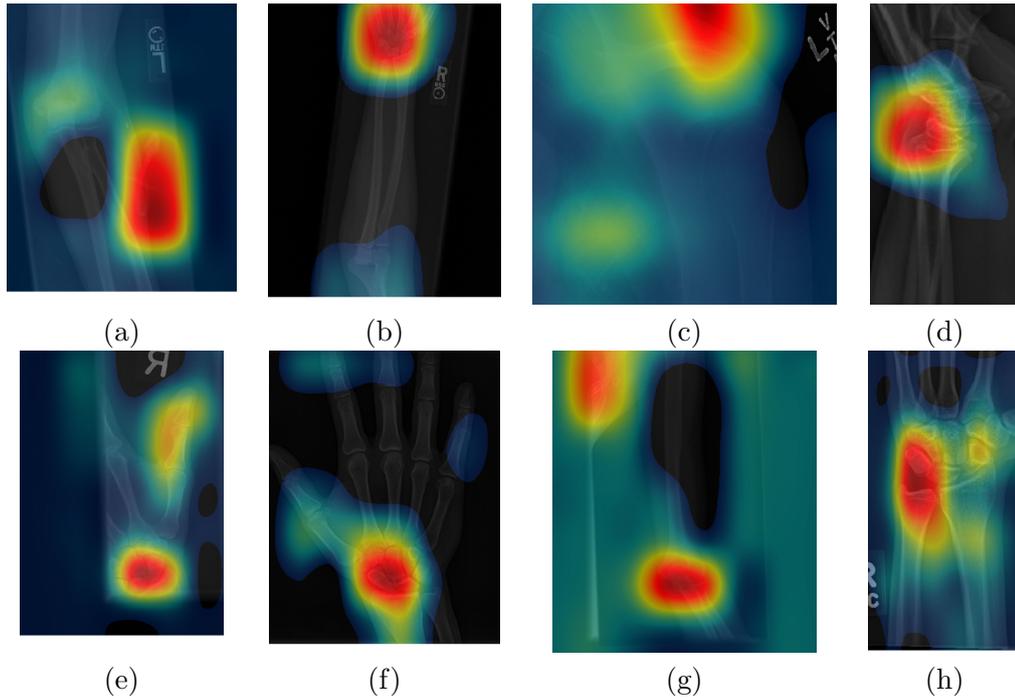


Figure 7.7: Illustration of activation maps overlaid over the eight radiographs without abnormalities of Figure 7.1 to indicate the regions of the image that activated an Inception-ResNet-V2 architecture. (a) Elbow, (b) Forearm, (c) Shoulder, (d) Wrist (lateral view), (e) Finger, (f) Hand, (g) Humerus, (h) Wrist. It should be noted that the activation regions are more localised than those of the ResNet-50.

wrist X-ray images were evaluated using the CAM technique. Its aim is to show how those two architectures work with randomly picked wrist X-ray images. Those thirty images were tested to ResNet-50 and Inception-ResNet-V2 architectures. As mentioned in Chapter 6, those two architectures were trained to learn about normal and abnormal wrist X-ray images from the MURA data set. In this chapter where the focus is on Class Activation Mapping for wrist X-ray images then those two architectures are evaluated. The representation of Class Activation Mapping was generated at layer ‘activation\49_relu’ for ResNet-50 and ‘conv\7_bac’ for Inception-ResNet-v2, respectively. The CAM map is then adjusted to the input resolution and overlaid over the top of the wrist X-ray image to indicate the location of the network intention to classify the images. The CAM’s visualisation also might give an indication to highlight the area(s) where abnormalities are possibly located.

The results are shown in Figure B.1 and Figure B.2. Those results showed the visualisation of this through the activation layer within the networks. Visually, both architectures could point out the location of suspected area(s) that lead the network to decide whether the wrist X-ray image is normal or abnormal. Those two architectures were tested with the same thirty wrist X-ray images but in this

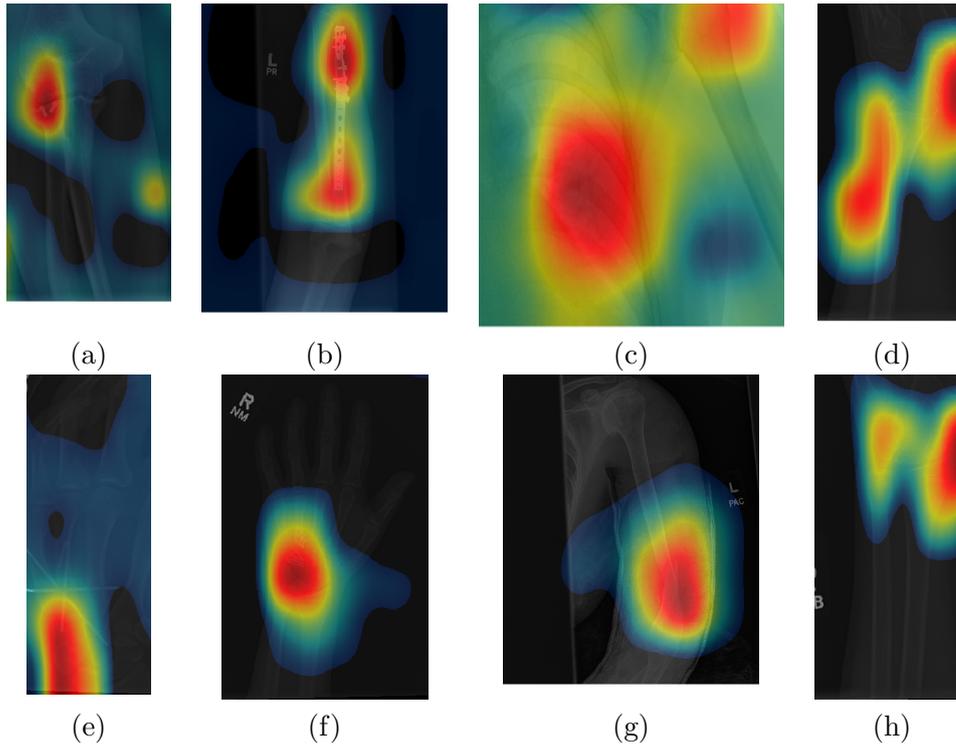


Figure 7.8: Illustration of activation maps overlaid over the eight radiographs with abnormalities of Figure 7.2 to indicate the regions of the image that activated an Inception-ResNet-v2 architecture. As for the cases without abnormalities, the activation regions are more located, e.g., ((c) Shoulder, (d) Lateral view of Wrist, and (h) Posterior-Anterior view of Wrist)) and in addition, the abnormalities are better located, e.g., ((a) Elbow, (b) Forearm, (e) Finger, (f) Hand, and (g) Humerus).

particular case, ResNet-50 shows more focus location of the suspected location of abnormal or normal image. Inception-ResNet-v2 has a more complex architecture than ResNet-50 but for some images such as image no (1), (2), (9), it shows a broad-covered heatmap style visualisation compared to the same images tested on ResNet-50.

The depth of the network also becomes a factor in its ability to detect better. For instance, the result of identification from Inception-ResNet-v2 has more precise and focuses on the abnormality suspected area compared to the same image identified by using ResNet-50. If we look furthermore, the activated regions in ResNet-50 appeared more *broad-brushed* than those of the Inception-ResNet-v2. This applied to the cases without abnormalities (Figures 7.9 and 7.7) and those with abnormalities (Figures 7.6 and 7.8); second, the localisation of regions of attention by Inception-ResNet-v2 also appeared more precise than the ResNet-50. When it comes to a non-fractures abnormality such as an image containing a metallic implant, the activation provides a consistent focus in areas where abnormalities are expected to appear.

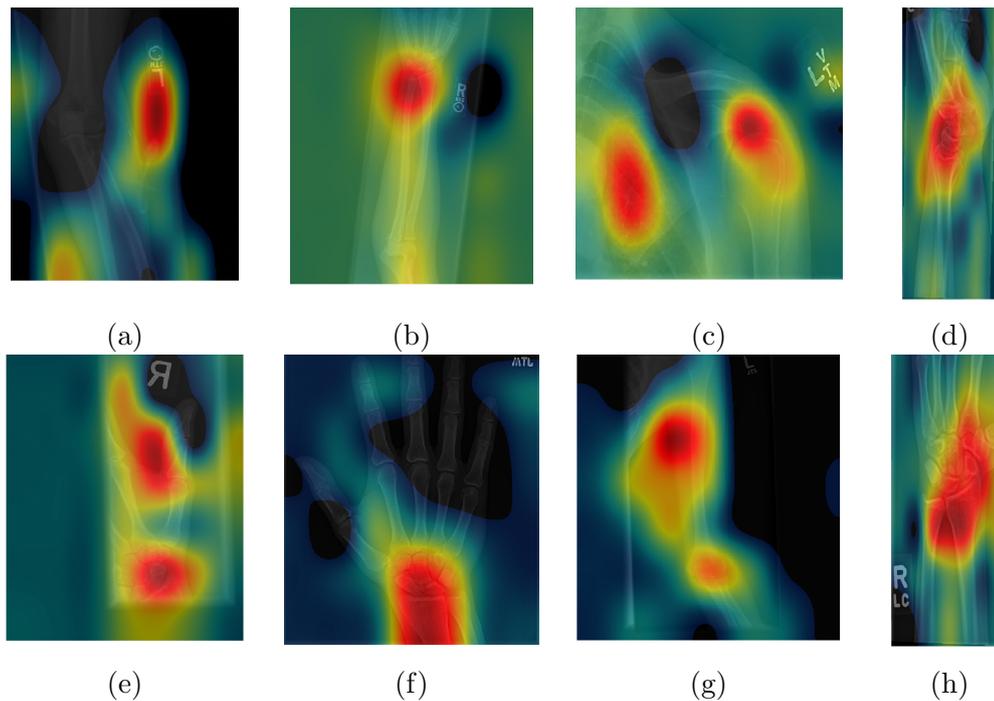


Figure 7.9: Illustration of class activation mapping overlaid over the eight radiographs without abnormalities of Figure 7.1 to indicate the regions of the image that activated a ResNet-50 architecture. (a) Elbow, (b) Forearm, (c) Shoulder, (d) Wrist (lateral view), (d) Wrist (lateral view), (e) Finger, (f) Hand, (g) Humerus, (h) Wrist. As these cases are positive (no abnormality), the regions of activation are not as critical as those with abnormalities.

The correct class prediction suggests that the network has learned the features from regions properly.

This result could be optimised by modifying training strategies such as a similar body of anatomy data set. In this work, as we have read, the combination of the data sets provided by the University of Exeter and MURA has given a promising outcome.

Several strategies to increase network performance, such as additional steps within pre-processing data sets like features irrelevant removing, appropriate data augmentations turn to provide higher accuracy of prediction. Not just modifying images within the data set before putting it up in the deep network training process but also after post-training. As the base of visualisation for class activation mapping, the activation layer allows architectures to be re-trained with additional data to avoid bias and distinction of the CNN architecture.

The comparison between SGDM, Adam, and RMSProp shows no indicative superiority implying that each of these optimisers was capable of achieving the optimal solution. Incremental change to the number of epochs beyond step 30 yields no improvement in accuracy indicating that the architectures have converged. The choice

of the attempted mini-batches shows no difference in results.

Ensemble architectures have shown better results than individual settings. A combination of architectures means data being learned further within the network. Which combination of architectures gave better results specifically for wrist analysis must be studied further. The wrist is part of human anatomy that directly connects with the hand, elbow, and forearm. The features such as the hand, elbow, and forearm bone structure can be added to the training data set to provide robust fine-tuning in anomaly detection. This strategy also gives guidance to the network for better outcomes.

7.6 Summary

This work investigated the analysis of fractures within detected wrist X-ray images using CNN. Eleven convolutional neural networks were performed to classify wrist radiographs into two groups which are abnormal and normal. Wrist X-ray categorised by experts as abnormal contains fractures or metallic plates. Normal images are categorised as healthy and nothing suspicious the image by experts. Accuracy and Cohen's Kappa value were obtained to measure the performance of each CNN.

Data augmentation and mixed data sets were used to enhance its performance. Data augmentation was applied to give additional knowledge variations to the networks during the network training phase. The augmented data set has shown a positive contribution to predicting an abnormality within the image-predicted classification outcome. As shown in this chapter, the tweak strategy with CNN hyper-parameters can be used further by paying attention to the specific settings of the epoch, initial learning rate, and mini-batch size to avoid the misunderstood result.

Adam optimiser has performed well compared to SGDM and RMSProp. This optimiser can further train CNN-based networks in medical image analysis cases. The optimiser's choice is important in enhancing trained-based image-predicted classification outcomes.

The highest score of accuracy and Cohen's Kappa were achieved by Inception-ResNet-v2, which has the thickest depth among the eleven architectures evaluated. Figure 7.10 shows how each architecture is represented by a circle, except those with augmentation represented by an asterisk.

Numbers in Figure 7.10 were added and these corresponded to the order of

Table 7.1: Summary of CNNs Activation Size and Parameters

No	CNN	Total of Activation Size	Total Parameters
1	GoogLeNet	10,062,768	6,998,552
2	VGG-19	31,411,664	143,667,240
3	AlexNet	2,080,811	58,639,715
4	SqueezeNet	7,990,891	1,235,496
5	ResNet-18	8,532,432	11,694,248
6	Inception-v3	32,657,126	24,310,664
7	ResNet-50	35,597,968	25,562,762
8	VGG-16	28,802,512	138,357,544
9	ResNet-101	54,896,592	44,618,216
10	DenseNet-201	76,359,248	20,034,304
11	Inception-ResNet-v2	93,359,072	57,134,728
Total		381,751,084	532,253,469

Table 6.7 for visualisation purpose where 1 correlated to GoogLeNet, 2 for VGG-19, 3 for AlexNet, 4 for SqueezeNet, 5 for ResNet-18, 6 for Inception-v3, 7 for ResNet-50, 8 for VGG-16, 9 for ResNet-101, 10 for DenseNet-201, 11 for Inception-ResNet-v2, 12 for ResNet-50 (augmentation), 13 for Inception-ResNet-v2 (augmentation)). Notice the slight improvement provided by deeper networks and the significant improvement that corresponds to data augmentation.

The top two best-performance architectures (ResNet-50 and Inception-ResNet-v2) are considered for further experiments using these extra techniques. Inception-ResNet-v2 could deliver a better prediction of visualisation outcome even though we have seen Resnet-50 has been applied and showed good performance to identify abnormalities within the wrist X-ray image in many studies. The result of the classification of wrist X-ray images between normal and abnormal conditions has been improved.

A convolutional neural network architecture contains large numbers of activation sizes and parameters depending on how many layers the architecture has. The thicker layers architecture has translated to more activation and parameter it has. Table 7.1 shows the summary of CNNs activation size and parameters each of the eleven CNNs has.

The Class Activation Mapping technique has been adopted to exploit visual prediction from the trained network by overlaying it onto the predicted X-ray image. This layer-wise activation predicted area of abnormality within an image shows that extra steps of augmentation to the training process of the network have contributed

Table 7.2: Summary of accuracy within chosen activation layer of CNN.

No	CNN	Activation Layer	Accuracy
1	AlexNet	fc7	0.6591
2	GoogLeNet	inception_5b-output	0.6956
3	SqueezeNet	relu_conv10	0.6991
4	ResNet-50	fc1000	0.6997
5	VGG-19	fc8	0.7036
6	VGG-16	fc8	0.7039
7	ResNet-101	fc1000	0.7074
8	Inception-v3	predictions	0.7218
9	ResNet-18	fc1000	0.7279
10	DenseNet-201	fc1000	0.7286
11	Inception-ResNet-v2	predictions	0.7442

to determining the position of abnormality. A better visualisation of the predicted area of abnormality within the image using the CAM technique to bridge the gap between what deep learning-based machines has been learned from the image and the human perspectives such as radiologists.

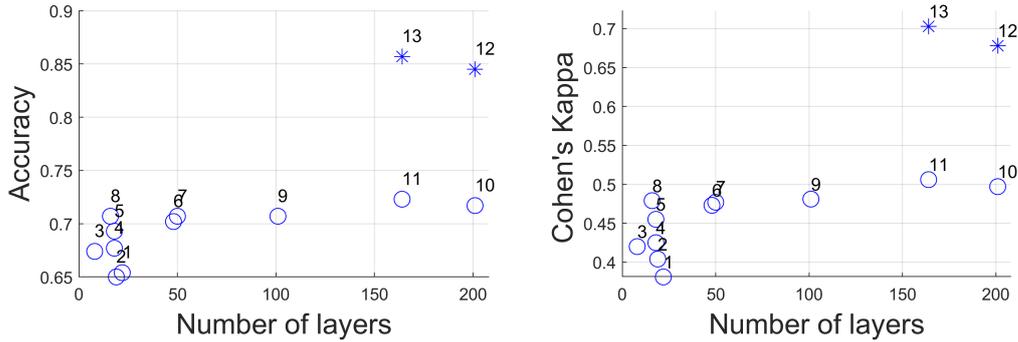


Figure 7.10: Illustration of the effect of the number of layers of architectures against the two metrics used in this work accuracy and Cohen's Kappa.

Visual prediction from the trained network using Class Activation Mapping depends on which activation layer is used. Table 7.2 shows a sample of the accuracy of chosen activation layer of CNN. The hidden activation layer will control how well the network architecture learns the training data set. In this work, the activation layer of each of the eleven CNN architectures was chosen and took the learning accuracy from the training process. The activation layer's accuracy determined which architecture of CNN architectures worked well to detect abnormality within the wrist X-ray image. The activation layer named 'predictions' shown in Table 7.2 has the highest accuracy compared to other CNN architectures experimented with in this work.

This study concludes that the parameters and combinations of data sets are necessary to improve the classification results. The problem of the small size of the dataset can be solved by combining it with a similar publicly available dataset. Visualizing suspicious areas of abnormalities in an established network-based image can facilitate the work of stakeholders, e.g. radiologists.

Chapter 8

Discussion and Conclusion

8.1 Discussion

The analysis of wrist X-ray images was started by exploring DICOM's metadata attributes. Metadata was used to classify the wrist X-ray image data set into the PA and LA positions and 'Photometric Interpretation' as a metadata parameter. Several image processing techniques such as edge detection and refining the image's position were applied to classified images. An image profile was also used to characterize the wrist position inside the X-ray image.

Texture features of the wrist within the X-ray image were also explored. Local Binary Patterns (LBP) is a visual descriptor to classify the texture features of an image. The aim to use the texture features is to determine if there are geometric differences in wrist X-ray images between the successful and unsuccessful cases of Manipulation under Anaesthesia (MuA) of wrist X-ray images. The semi-automatic comparisons extracted a series of measurements, e.g. widths of forearm and metacarpal, based on three manually-placed landmarks. This study has given an encouraging statistical difference in measurements.

The next part of the work focused on identifying abnormalities within an X-ray of the wrist. Eleven models of CNN were explored. Training strategies include modification of the data set, data augmentation, hyperparameters settings, and applying the geometric transformation to the augmented data set. The depth of the architecture is also related to the increase of network's performance to classify the wrist X-ray image. In this case explained in Chapter 6, Inception-ResNet-V2 has more network layers performed better than the ResNet-50.

Transfer learning was also applied to the CNNs training. This strategy has been applied to the eleven convolutional neural networks that have been used in this research. Those eleven deep learning architectures were pre-trained in non-medical images, i.e. ImageNet. This thesis used those eleven deep learning architectures to recognise medical images, specifically wrist X-ray images. The transfer learning technique was adopted to train those eleven networks with the wrist X-ray images. The strategy of CNN's network training to detect abnormalities has been taken further by using a more complex data augmentation strategy and testing separate from each of the three optimizations (SGDM, ADAM, RMSProp) available.

Although the prediction results were not straightly comparable because of different strategies and details in some hyperparameters, the changing value of hyperparameters has affected the classification results. To enhance its performance, a data augmentation strategy was also applied to the data set. It was applied to ResNet-50 and Inception-ResNet-v2. Better results were obtained by Inception-ResNet-v2 (mean accuracy = 0.723, mean Cohen's Kappa = 0.506). These were significantly improved with augmentation to Inception-ResNet-v2 (mean accuracy = 0.857, mean Cohen's Kappa = 0.703). Furthermore, by using the strategies mentioned above, this research using wrist X-ray image classification has to perform slightly better than [68]'s wrist X-ray image classification results.

The identified result is then supported by the visualisation of the area of abnormality located in the wrist X-ray image. Class Activation Mapping is a technique to exploit visual prediction from the trained network by overlaying it onto the predicted X-ray image. This visualisation supports the predicted result by CNN. The CAM is mapped to the suspect abnormal area predicted in the wrist X-ray image by the CNN. It can be a better understand of what deep learning-based machine has learned from the wrist X-ray image.

The experiments and the results described in this thesis are available through the author's GitHub repository on <https://github.com/amno/WristMatlab>.

8.2 Contributions

This work in the wrist fractures analysis as observed with X-ray imaging has several contributions to knowledge

1. The geometric analysis contributes to a new perspective of using several body

landmarks as tools to extract the texture of bone, which can be used to learn about bone-based illnesses like *osteoporosis*.

2. Through the use of Local Binary Patterns (LBP) as a visual descriptor, the image texture has come out as the most distinct among extracted measurements to analyse X-ray image-based wrist fractures.
3. Dataset augmentation can be used to enhance the performance of the convolutional neural network model to classify a wrist X-ray image. This thesis implemented data augmentation strategies such as using geometric transformation like rotation, reflection, translation, shear operation of the wrist X-ray images, Hough's transform, Contrast Limited Adaptive Histogram Equalisation and combining data sets with other anatomical images that still contain the wrist as the area of interest.
4. In this thesis, transfer learning is adapted into eleven CNN architectures. It has been proven that those eleven CNN architectures can be used to train collections of wrist X-rays without having to build networks from scratch and still benefit of pre-trained model.
5. It has been demonstrated that the Class Activation Mapping (CAM) technique can be used for medical image analysis, especially to highlight abnormal areas. It visually displays the network's intention whilst deciding the area of interest to classify the identifiable objects based on the learning network input.

8.3 Conclusions

Several conclusions to the knowledge of the area wrist fractures analysis based on X-ray image are

1. Image texture is a geometric feature of an image that comes up as the most distinct among the measurements extracted for analysing wrist fractures based on radiographic images. Measurements were extracted using Local Binary Patterns (LBP). The LBP technique can also be used to identify abnormalities in radiographs of the wrist and other areas associated with bone disease.
2. The deep learning model can be used to classify normal and abnormal wrist X-ray images through the use of pre-trained CNN architectures. Classifica-

tion performance can be improved by Transfer Learning techniques and data limitations can be enlarged by implementing data augmentation strategies.

3. The Class activation mapping (CAM) techniques can be used to determine which parts of the wrist X-ray images might trigger a decision to be abnormal like fractures using a Convolutional Neural Network model. The CAM is visualised and superimposed on the predicted image.

8.4 Future work

This section is dedicated to some ideas for future works related to this research. Future work could focus on a different anatomical part other than the wrist, such as the forearm, hand, and fingers. Training data sets of combinations of anatomical parts have been carried out during the research period but need further analysis to obtain better information.

Medical problems such as wrist fractures are caused by accidents and can also be triggered by other things, such as reduced bone mass, which puts patients at risk for *osteoporosis* [69]. This situation could be explored further as it requires a unique multi-modal analysis of medical data related to the patient's case to provide appropriate treatment.

The performance of deep learning to identify abnormality within the wrist X-ray images in this thesis is also promising for further explorations. Predicting post-surgical results before their understanding of the problem spreads widely can be explored further.

The Class Activation Mapping can be further investigated to support the post-surgical results so that clinicians can get not just visualisation prediction area of abnormality within the image but also the possibility of post-treatment.

Appendix A

The MIUA 2019 Poster

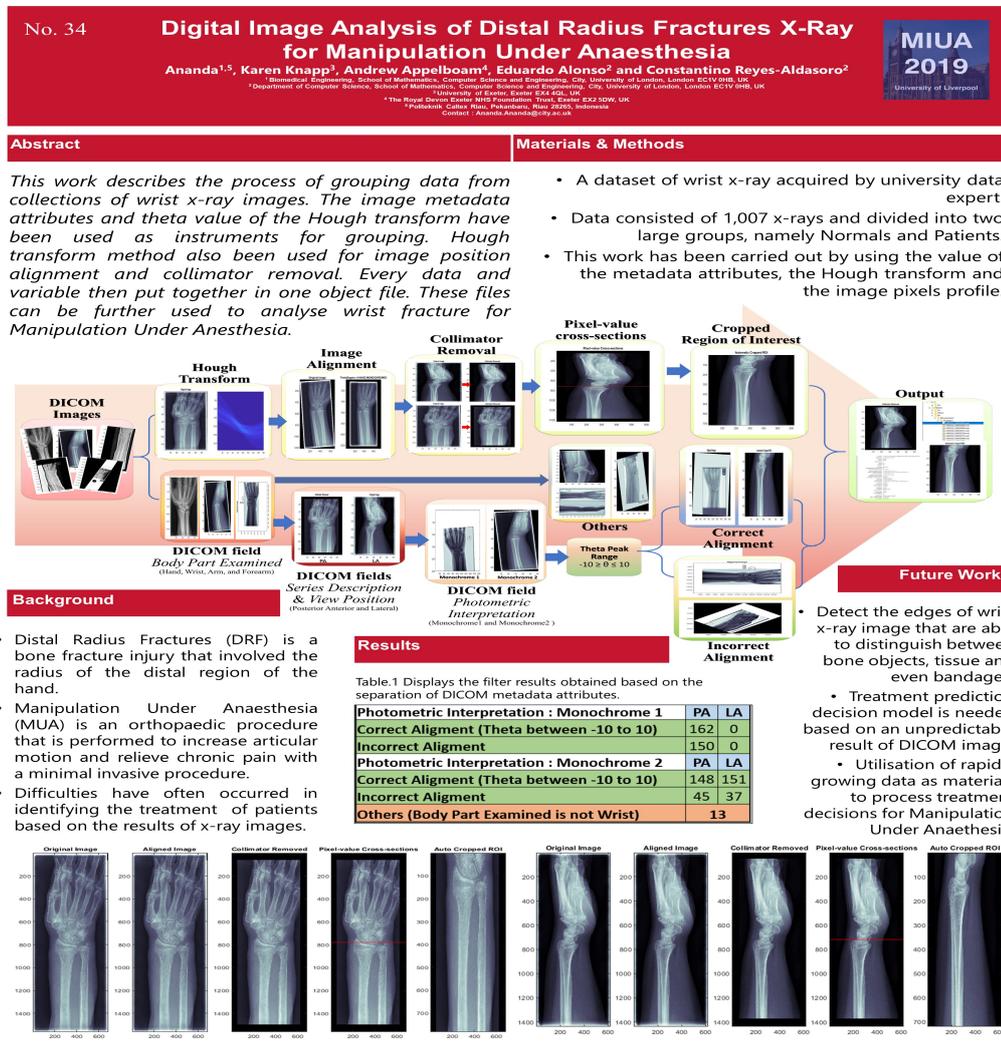
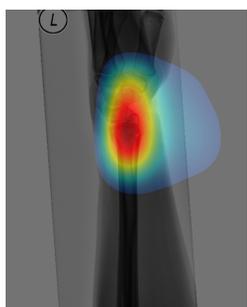


Figure A.1: Poster presented at the MIUA 2019 conference

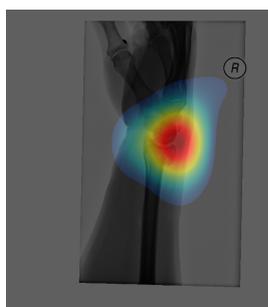
Appendix B

Samples of CAM on Wrist

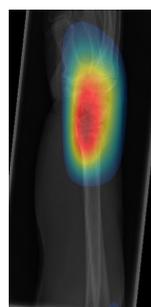
X-Ray



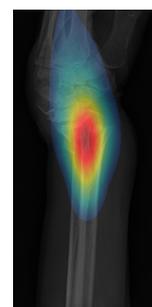
(1)



(2)



(3)



(4)



(5)



(6)



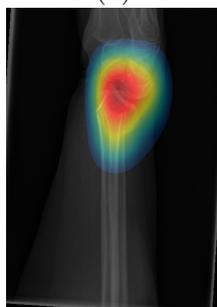
(7)



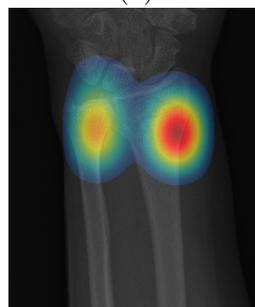
(8)



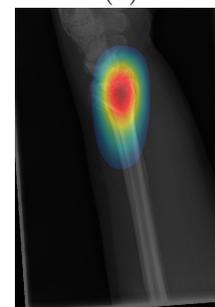
(9)



(10)



(11)



(12)

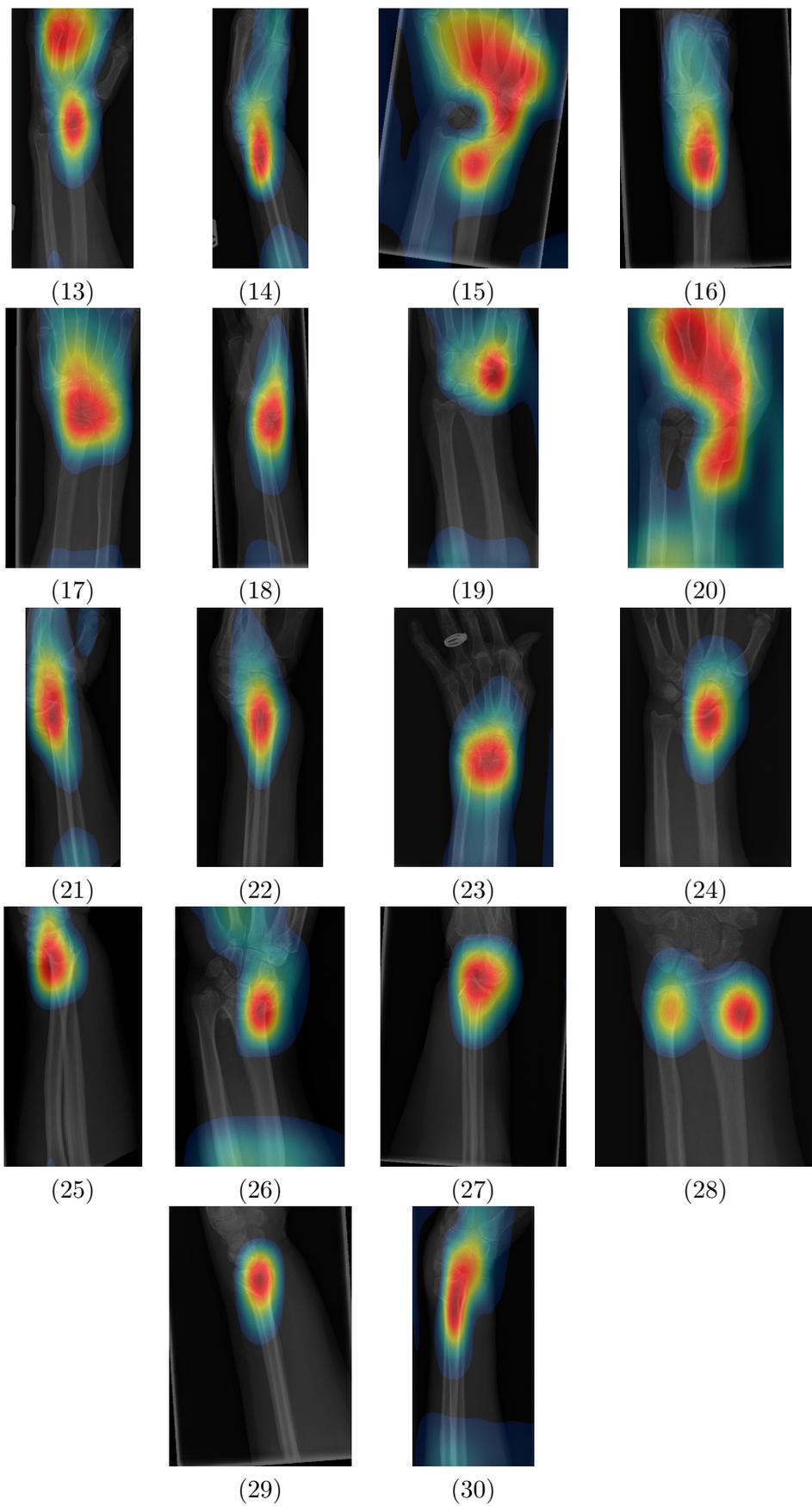
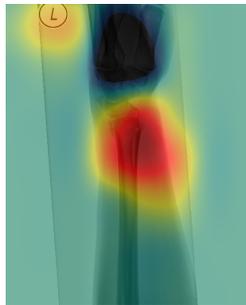
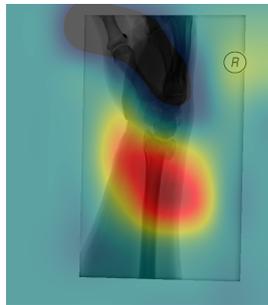


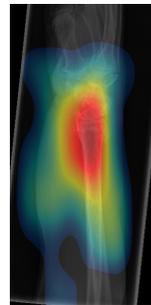
Figure B.1: Thirty sample images of Wrist X-ray overlaid with CAM based on ResNet-50 architecture.



(1)



(2)



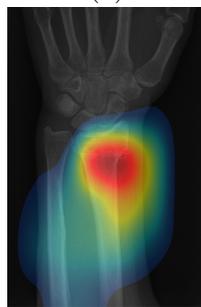
(3)



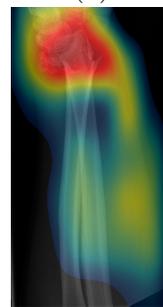
(4)



(5)



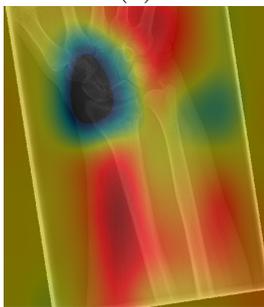
(6)



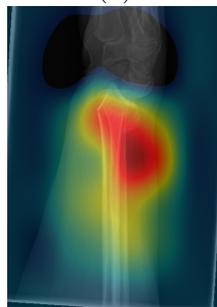
(7)



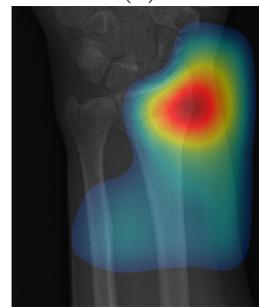
(8)



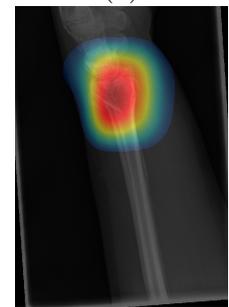
(9)



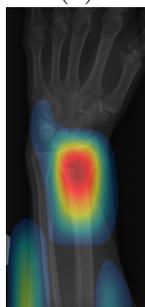
(10)



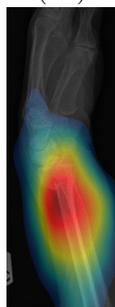
(11)



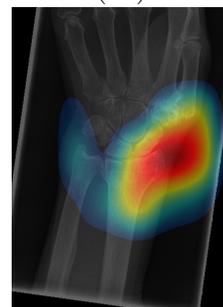
(12)



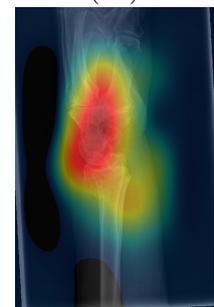
(13)



(14)



(15)



(16)

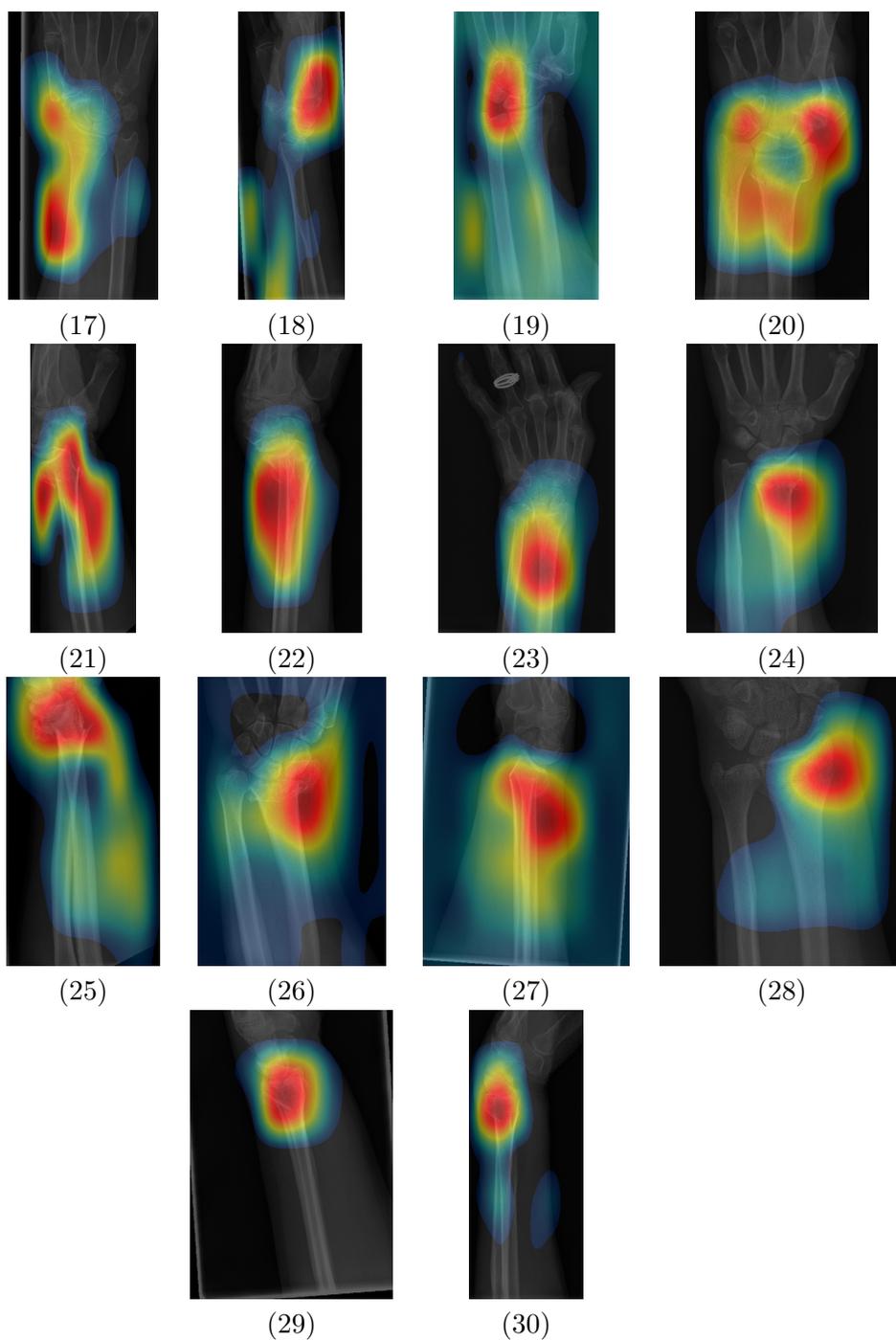


Figure B.2: Thirty sample images of Wrist X-ray overlaid with CAM based on Inception-ResNet-V2 architecture.

Bibliography

- [1] B. D. Adams, B. P. Kleinhenz, and J. J. Guan, “Wrist arthrodesis for failed total wrist arthroplasty,” *The Journal of Hand Surgery*, vol. 41, no. 6, pp. 673–679, Jun 2016.
- [2] S. M. M. R. Al Arif, K. Knapp, and G. Slabaugh, “Fully automatic cervical vertebrae segmentation framework for x-ray images,” *Computer Methods and Programs in Biomedicine*, vol. 157, pp. 95–111, Apr 2018.
- [3] O. M. Al Okashi, H. Du, and H. Al-Assam, “Automatic pelvis segmentation from x-ray images of a mouse model,” in *Proceedings Volume 10221, Mobile Multimedia/Image Processing, Security, and Applications 2017*, S. S. Agaian and S. A. Jassim, Eds., May 2017, p. 1022108. [Online]. Available: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2264803>
- [4] J. E. Aldrich, “Basic physics of ultrasound imaging,” *Critical Care Medicine*, vol. 35, no. 5, p. S131, May 2007.
- [5] L. Alzubaidi, M. Al-Amidie, A. Al-Asadi, A. J. Humaidi, O. Al-Shamma, M. A. Fadhel, J. Zhang, J. Santamaría, and Y. Duan, “Novel transfer learning approach for medical imaging with limited labeled data,” *Cancers*, vol. 13, no. 7, p. 1590, Mar 2021.
- [6] A. F. Ambrose, L. Cruz, and G. Paul, “Falls and fractures: A systematic approach to screening and prevention,” *Maturitas*, vol. 82, no. 1, pp. 85–93, Sep 2015.
- [7] Ananda, C. Karabag, A. Ter-Sarkisov, E. Alonso, and C. C. Reyes-Aldasoro, “Radiography classification: A comparison between eleven convolutional neural networks,” in *2020 Fourth International Conference on Multimedia Computing, Networking and Applications (MCNA)*, Oct 2020, p. 119–125.

- [8] A. Ananda, K. H. Ngan, C. Karabağ, A. Ter-Sarkisov, E. Alonso, and C. C. Reyes-Aldasoro, “Classification and visualisation of normal and abnormal radiographs; a comparison between eleven convolutional neural network architectures,” *Sensors*, vol. 21, no. 1616, p. 5381, Jan 2021.
- [9] L. Ardouin, A. Durand, A. Gay, and M. Leroy, “Why do we use arthroscopy for distal radius fractures?” *European Journal of Orthopaedic Surgery & Traumatology*, vol. 28, no. 8, pp. 1505–1514, Dec 2018.
- [10] J. N. Aronowitz, “An early x-ray tube,” *International Journal of Radiation Oncology*Biography*Physics*, vol. 88, no. 1, p. 5–6, Jan 2014.
- [11] R. Arora, M. Gabl, M. Gschwentner, C. Deml, D. Krappinger, and M. Lutz, “A Comparative Study of Clinical and Radiologic Outcomes of Unstable Colles Type Distal Radius Fractures in Patients Older Than 70 Years: Nonoperative Treatment Versus Volar Locking Plating,” *Journal of Orthopaedic Trauma*, vol. 23, no. 4, pp. 237–242, Apr. 2009.
- [12] A. Barai, B. Lambie, C. Cosgrave, and J. Baxter, “Management of distal radius fractures in the emergency department: A long-term functional outcome measure study with the disabilities of arm, shoulder and hand (dash) scores,” *Emergency Medicine Australasia*, vol. 30, no. 4, pp. 530–537, 2018.
- [13] T. Bashore, “Fundamentals of x-ray imaging and radiation safety,” *Catheterization and Cardiovascular Interventions*, vol. 54, no. 1, pp. 126–135, Sep 2001.
- [14] F. Bianconi, A. Fernández, F. Smeraldi, and G. Pascoletti, “Colour and texture descriptors for visual recognition: A historical overview,” *Journal of Imaging*, vol. 7, no. 1111, p. 245, Nov 2021.
- [15] Bioengineer, “World’s first human x-ray 1895 röntgen’s wife hand,” Dec 2014. [Online]. Available: <https://bioengineer.org/worlds-first-human-x-ray-1895-rontgens-wife-hand/>
- [16] B. Blaus, “English: X-ray. see a full animation of this medical topic.” Nov 2015. [Online]. Available: <https://commons.wikimedia.org/wiki/File:X-Ray.png>
- [17] R. A. Bloom and J. W. Laws, “Humeral cortical thickness as an index of osteoporosis in women,” *The British Journal of Radiology*,

- vol. 43, no. 512, pp. 522–527, Aug. 1970. [Online]. Available: <https://www.birpublications.org/doi/10.1259/0007-1285-43-512-522>
- [18] S. Bolander, “A systematic approach to describing fractures,” *Journal of the American Academy of Physician Assistants*, vol. 32, no. 5, p. 23–29, May 2019.
- [19] A. P. Brady, “Error and discrepancy in radiology: inevitable or avoidable?” *Insights into Imaging*, vol. 8, no. 1, pp. 171–182, Dec 2016.
- [20] L. Cai, S. Zhu, S. Du, W. Lin, T. Wang, D. Lu, and H. Chen, “The relationship between radiographic parameters and clinical outcome of distal radius fractures in elderly patients,” *Orthopaedics & traumatology, surgery & research: OTSR*, vol. 101, no. 7, p. 827–831, Nov 2015.
- [21] J. Canny, “A computational approach to edge detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, Nov 1986.
- [22] J. Canny, “A Computational Approach to Edge Detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
- [23] M. E. Celebi, N. Codella, and A. Halpern, “Dermoscopy image analysis: Overview and future directions,” *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 2, p. 474–478, Mar 2019.
- [24] J. Cohen, “A coefficient of agreement for nominal scales,” *Educational and Psychological Measurement*, vol. 20, no. 1, p. 37–46, Apr 1960.
- [25] A. Colles, “On the fracture of the carpal extremity of the radius [1814],” *Injury*, vol. 2, no. 1, pp. 48–50, Jul 1970.
- [26] W. Cooney, J. Manuel, J. Froelich, and M. Rizzo, “Total wrist replacement: A retrospective comparative study,” *Journal of Wrist Surgery*, vol. 1, no. 2, pp. 165–172, Nov 2012.
- [27] E. Cremata, S. Collins, W. Clauson, A. B. Solinger, and E. S. Roberts, “Manipulation under anesthesia: A report of four cases,” *Journal of Manipulative and Physiological Therapeutics*, vol. 28, no. 7, p. 526–533, Sep 2005.

- [28] C. S. Crowe, B. B. Massenburg, S. D. Morrison, J. Chang, J. B. Friedrich, G. G. Abady, F. Alahdab, V. Alipour, J. Arabloo, M. Asaad, M. Banach, A. Bijani, A. M. Borzì, N. I. Briko, C. D. Castle, D. Y. Cho, M. T. Chung, A. Daryani, G. T. Demoz, Z. V. Dingels, H. T. Do, F. Fischer, J. T. Fox, T. Fukumoto, A. K. Gebre, B. Gebremichael, J. A. Haagsma, A. Haj-Mirzaian, D. W. Handiso, S. I. Hay, C. L. Hoang, S. S. N. Irvani, J. J. Jozwiak, R. Kalhor, A. Kasaeian, Y. S. Khader, R. Khalilov, E. A. Khan, R. Khundkar, S. Kisa, A. Kisa, Z. Liu, M. Majdan, N. Manafi, A. Manafi, A.-L. Manda, T. J. Mere-toja, T. R. Miller, A. Mohammadian-Hafshejani, R. Mohammadpourhodki, M. A. M. Bandpei, A. H. Mokdad, M. D. Naimzada, D. E. Ndwandwe, C. T. Nguyen, H. L. T. Nguyen, A. T. Olagunju, T. O. Olagunju, H. Q. Pham, D. R. A. Pribadi, N. Rabiee, K. Ramezanzadeh, K. Ranganathan, N. L. S. Roberts, L. Roever, S. Safari, A. M. Samy, L. S. Riera, S. Shahabi, C.-G. Smarandache, D. O. Sylte, B. E. Tesfay, B. X. Tran, I. Ullah, P. Vahedi, A. Vahedian-Azimi, T. Vos, D. H. Woldeyes, A. B. Wondmieneh, Z.-J. Zhang, and S. L. James, “Global trends of hand and wrist trauma: a systematic analysis of fracture and digit amputation using the global burden of disease 2017 study,” *Injury Prevention*, vol. 26, no. Suppl 2, p. i115–i124, Oct 2020.
- [29] E. M. Curtis, R. van der Velde, R. J. Moon, J. P. W. van den Bergh, P. Geusens, F. de Vries, T. P. van Staa, C. Cooper, and N. C. Harvey, “Epidemiology of fractures in the united kingdom 1988-2012: Variation with age, sex, geography, ethnicity and socioeconomic status,” *Bone*, vol. 87, pp. 19–26, Jun 2016.
- [30] S. Dagenais, J. Mayer, J. R. Wooley, and S. Haldeman, “Evidence-informed management of chronic low back pain with medicine-assisted manipulation,” *The Spine Journal*, vol. 8, no. 1, pp. 142–149, 2008.
- [31] M. de Bruijne, “Machine learning approaches in medical image analysis: From detection to diagnosis,” *Medical Image Analysis*, vol. 33, pp. 94–97, Oct 2016.
- [32] C. E. de Putter, R. W. Selles, S. Polinder, K. A. Hartholt, C. W. Looman, M. J. M. Panneman, J. A. N. Verhaar, S. E. R. Hovius, and E. F. van Beek, “Epidemiology and health-care utilisation of wrist fractures in older adults in the Netherlands, 1997-2009,” *Injury*, vol. 44, no. 4, pp. 421–426, Apr 2013.

- [33] P. K. Desai, “Colles fracture — radiology reference article — radiopaedia.org.” [Online]. Available: <https://radiopaedia.org/articles/colles-fracture?lang=us>
- [34] R. J. Diaz-Garcia and K. C. Chung, “The evolution of distal radius fracture management – a historical treatise,” *Hand Clinics*, vol. 28, no. 2, pp. 105–111, May 2012.
- [35] D. DiGiorgi, “Spinal manipulation under anesthesia: a narrative review of the literature and commentary,” *Chiropractic and Manual Therapies*, vol. 21, p. 14, May 2013.
- [36] P. Divya, R. Geetha, and P. S. Kavitha, “Semi-automatic detection of fracture in wrist bones using graph-based grammar approach,” in *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*, Jul 2012, pp. 1–7.
- [37] R. O. Duda and P. E. Hart, “Use of the Hough transformation to detect lines and curves in pictures,” *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972. [Online]. Available: <https://doi.org/10.1145/361237.361242>
- [38] Y. Durani, “X-ray exam: Wrist (for parents) - nemours kidshealth.” [Online]. Available: <https://kidshealth.org/en/parents/xray-exam-wrist.html>
- [39] R. Ebsim, J. Naqvi, and T. F. Cootes, “Automatic detection of wrist fractures from posteroanterior and lateral radiographs: A deep learning-based approach,” in *International Workshop on Computational Methods and Clinical Applications in Musculoskeletal Imaging*. Springer, 2018, pp. 114–125.
- [40] M. Edwards, K. Jameson, H. Denison, N. Harvey, S. A. Aihie, E. Dennison, and C. Cooper, “Clinical risk factors, bone density and fall history in the prediction of incident fracture among men and women,” *Bone*, vol. 52, no. 2, pp. 541–547, Feb 2013.
- [41] T. E. Encyclopaedia Britannica, “diagnostic imaging — definition and types — britannica.” [Online]. Available: <https://www.britannica.com/science/diagnostic-imaging>
- [42] —, “radiology — medicine — britannica.” [Online]. Available: <https://www.britannica.com/science/radiology>

- [43] —, “wrist — anatomy — britannica.” [Online]. Available: <https://www.britannica.com/science/wrist-anatomy>
- [44] K. E. Ensrud, “Epidemiology of fracture risk with advancing age,” *The Journals of Gerontology: Series A*, vol. 68, no. 10, pp. 1236–1242, Oct 2013.
- [45] T. Fawcett, “An introduction to roc analysis,” *Pattern Recognition Letters*, vol. 27, no. 8, p. 861–874, Jun 2006.
- [46] H. Fonseca, D. Moreira-Gonçalves, H.-J. A. Coriolano, and J. A. Duarte, “Bone quality: The determinants of bone strength and fragility,” *Sports Medicine*, vol. 44, no. 1, p. 37–53, Jan 2014.
- [47] F. Gaillard, “Barton fracture — radiology reference article — radiopaedia.org.” [Online]. Available: <https://radiopaedia.org/articles/barton-fracture>
- [48] E. D. H. Gates, J. S. Lin, J. S. Weinberg, S. S. Prabhu, J. Hamilton, J. D. Hazle, G. N. Fuller, V. Baladandayuthapani, D. T. Fuentes, and D. Schellingerhout, “Imaging-based algorithm for the local grading of glioma,” *American Journal of Neuroradiology*, vol. 41, no. 3, p. 400–407, Mar 2020.
- [49] A. Goel, “Colles fracture — radiology case — radiopaedia.org.” [Online]. Available: <https://radiopaedia.org/cases/colles-fracture-10?lang=us>
- [50] C. A. Goldfarb, Y. Yin, L. A. Gilula, A. J. Fisher, and M. I. Boyer, “Wrist fractures: What the clinician wants to know,” *Radiology*, vol. 219, no. 1, p. 11–28, Apr 2001.
- [51] W. Gong, J. Wang, L. Huang, X. Yang, D. Chen, and M. Zheng, “Diagnosis of cervical plexus tumours by high-frequency ultrasonography,” *BMC Medical Imaging*, vol. 21, p. 148, Oct 2021.
- [52] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. Pearson, 2018.
- [53] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. The MIT Press, 2016.
- [54] R. Gordon, E. Cremata, and C. Hawk, “Guidelines for the practice and performance of manipulation under anesthesia,” *Chiropractic and Manual Therapies*, vol. 22, p. 7, Feb 2014.

- [55] J. B. Green, C. Deveikas, H. E. Ranger, J. G. Draghetti, L. C. Groat, E. D. Schumer, and B. M. Leslie, *Chapter 10 - Hand, Wrist, and Digit Injuries*. W.B. Saunders, Jan 2016, pp. 344–435.
- [56] P. E. Greenman, “Manipulation with the patient under anesthesia,” *The Journal of the American Osteopathic Association*, vol. 92, no. 9, p. 1159–1159, Sep 1992.
- [57] A. Halim and A.-P. C. Weiss, “Total wrist arthroplasty,” *The Journal of Hand Surgery*, vol. 42, no. 3, pp. 198–209, Mar 2017.
- [58] P. Hamet and J. Tremblay, “Artificial intelligence in medicine,” *Metabolism*, vol. 69, pp. S36–S40, Apr 2017.
- [59] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2016, pp. 770–778.
- [60] P. V. Hough, “Method and means for recognizing complex patterns,” Dec. 18 1962, uS Patent 3,069,654.
- [61] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” *arXiv:1608.06993 [cs]*, Aug 2016, arXiv: 1608.06993. [Online]. Available: <http://arxiv.org/abs/1608.06993>
- [62] W. Huda and R. B. Abrahams, “X-ray-based medical imaging and resolution,” *American Journal of Roentgenology*, vol. 204, no. 4, p. W393–W397, Apr 2015.
- [63] R. Hummel, “Image enhancement by histogram transformation,” *Computer Graphics and Image Processing*, vol. 6, no. 2, p. 184–195, Apr 1977.
- [64] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, “Squeezenet: Alexnet-level accuracy with 50x fewer parameters and ̄0.5mb model size,” *arXiv:1602.07360 [cs]*, Feb 2016, arXiv: 1602.07360. [Online]. Available: <http://arxiv.org/abs/1602.07360>
- [65] Innolitics. Photometric interpretation attribute - dicom standard browser. [Online]. Available: <https://dicom.innolitics.com/ciods/ct-image/image-pixel/00280004>

- [66] C. Jantzen, L. K. Cieslak, A. F. Barzanji, P. B. Johansen, S. W. Rasmussen, and T. A. Schmidt, “Colles’ fractures and osteoporosis—A new role for the Emergency Department,” *Injury*, vol. 47, no. 4, pp. 930–933, Apr. 2016.
- [67] J. Jones, “Fall onto an outstretched hand — radiology reference article — radiopaedia.org.” [Online]. Available: <https://radiopaedia.org/articles/fall-onto-an-outstretched-hand?lang=us>
- [68] I. Kandel, M. Castelli, and A. Popovič, “Musculoskeletal images classification for detection of fractures using transfer learning,” *Journal of Imaging*, vol. 6, no. 1111, p. 127, Nov 2020.
- [69] K. M. Knapp, R. M. Meertens, and R. Seymour, “Imaging and opportunistic identification of fractures,” *Pavilion Publishing*, vol. Vol.48(11), pp. 10–12, Nov 2018.
- [70] A. Krizhevsky, I. Sutskever, and G. E. Hinton, *ImageNet Classification with Deep Convolutional Neural Networks*. Curran Associates, Inc., 2012, p. 1097–1105. [Online]. Available: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>
- [71] H. Laga, Y. Guo, H. Tabia, R. Fisher, and M. Bennamoun, *3D Shape Analysis: Fundamentals, Theory, and Applications*. United States: Wiley-Blackwell, 2019.
- [72] K. Lameka, M. D. Farwell, and M. Ichise, *Chapter 11 - Positron Emission Tomography*, ser. Neuroimaging Part I. Elsevier, Jan 2016, vol. 135, p. 209–227. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780444534859000118>
- [73] C. F. Larsen, V. Brøndum, G. Wienholtz, J. Abrahamsen, and J. Beyer, “An algorithm for acute wrist trauma: A systematic approach to diagnosis,” *Journal of Hand Surgery*, vol. 18, no. 2, p. 207–212, Apr 1993.
- [74] C. H. Lee and H.-J. Yoon, “Medical big data: promise and challenges,” *Kidney Research and Clinical Practice*, vol. 36, no. 1, p. 3–11, Mar 2017.
- [75] H. Lee, S. Tajmir, J. Lee, M. Zissen, B. A. Yeshiwas, T. K. Alkasab, G. Choy, and S. Do, “Fully automated deep learning system for bone age assessment,” *Journal of Digital Imaging*, vol. 30, no. 4, pp. 427–441, Aug 2017.

- [76] P. R. Lokhande, S. Balaguru, G. Deenadayalan, and R. R. Ghorpade, "A review of contemporary researches on biomedical image analysis," in *Recent Trends in Image Processing and Pattern Recognition*, ser. Communications in Computer and Information Science, K. C. Santosh and R. S. Hegadi, Eds. Springer, 2019, p. 84–96.
- [77] J. Luo, M. Wu, D. Gopukumar, and Y. Zhao, "Big data application in biomedical research and health care: A literature review," *Biomedical Informatics Insights*, vol. 8, p. BII.S31559, Jan 2016.
- [78] P. MacPherson, E. L. Webb, W. Kamchedzera, E. Joeke, G. Mjoli, D. G. Lalloo, T. H. Divala, A. T. Choko, R. M. Burke, H. Maheswaran, M. Pai, S. B. Squire, M. Nliwasa, and E. L. Corbett, "Computer-aided x-ray screening for tuberculosis and hiv testing among adults with cough in malawi (the prospect study): A randomised trial and cost-effectiveness analysis," *PLOS Medicine*, vol. 18, no. 9, p. e1003752, Sep 2021.
- [79] H. Malik, A. Appelboam, and G. Taylor, "Colles' type distal radial fractures undergoing manipulation in the ed: a multicentre observational cohort study," *Emergency medicine journal: EMJ*, vol. 37, no. 8, p. 498–501, Aug 2020.
- [80] M. K. Manhard, J. S. Nyman, and M. D. Does, "Advances in imaging approaches to fracture risk evaluation," *Translational Research*, vol. 181, pp. 1–14, Mar 2017.
- [81] M. L. McHugh, "Interrater reliability: the kappa statistic," *Biochemia Medica*, vol. 22, no. 3, p. 276–282, Oct 2012.
- [82] W. H. Meadowcroft, *The ABC of the X Rays*. American Technical Book Company, 1896, no. 2.
- [83] S. Meena, P. Sharma, A. K. Sambharia, and A. Dawar, "Fractures of distal radius: An overview," *Journal of Family Medicine and Primary Care*, vol. 3, no. 4, pp. 325–332, 2014.
- [84] L. Monfils, *X-ray of a collesfracture of the left wrist accompanied by an ulnar styloid fracture*. [Online]. Available: <https://commons.wikimedia.org/wiki/File:Collesfracture.jpg>

- [85] R. J. Moon, N. C. Harvey, E. M. Curtis, F. de Vries, T. van Staa, and C. Cooper, “Ethnic and geographic variations in the epidemiology of childhood fractures in the united kingdom,” *Bone*, vol. 85, pp. 9–14, Apr 2016.
- [86] N. T. Morrell and A.-P. C. Weiss, *Total Wrist Arthroplasty for Treatment of Distal Radius Fractures*. Springer International Publishing, 2018, pp. 81–90. [Online]. Available: https://doi.org/10.1007/978-3-319-94202-5_6
- [87] W. J. Morton, *The X-ray; or, Photography of the invisible and its value in surgery*. American Technical Book Company, 1896.
- [88] L. Nanni, A. Lumini, and S. Brahmam, “Local binary patterns variants as texture descriptors for medical image analysis,” *Artificial Intelligence in Medicine*, vol. 49, no. 2, p. 117–125, Jun 2010.
- [89] Y. Nasser, M. E. Hassouni, A. Brahim, H. Toumi, E. Lespessailles, and R. Jenane, “Diagnosis of osteoporosis disease from bone x-ray images with stacked sparse autoencoder and svm classifier,” in *2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, May 2017, pp. 1–5.
- [90] N. National Electrical Manufacturers Association. Dicom standard. [Online]. Available: <https://www.dicomstandard.org/>
- [91] K. H. Ngan, A. d. Garcez, K. M. Knapp, A. Appelboam, and C. C. Reyes-Aldasoro, “Making densenet interpretable, a case study in clinical radiology,” *medRxiv*, p. 19013730, Dec 2019.
- [92] Q. Nguyen, S. Chaudhry, R. Sloan, I. Bhoora, and C. Willard, “The clinical scaphoid fracture: Early computed tomography as a practical approach,” *Annals of The Royal College of Surgeons of England*, vol. 90, no. 6, p. 488–491, Sep 2008.
- [93] L. Obert, F. Loisel, E. Jardin, N. Gasse, and D. Lepage, “High-energy injuries of the wrist,” *Orthopaedics & Traumatology : Surgery & Research*, vol. 102, no. 1, Supplement, pp. S81–S93, Feb 2016.
- [94] L. Obert, P. B. Rey, J. Uhring, N. Gasse, S. Rochet, D. Lepage, A. Serre, and P. Garbuio, “Fixation of distal radius fractures in adults: A review,”

- Orthopaedics & Traumatology: Surgery & Research*, vol. 99, no. 2, p. 216–234, Apr 2013.
- [95] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, p. 971–987, Jul 2002.
- [96] T. Ojala, M. Pietikäinen, and D. Harwood, “A comparative study of texture measures with classification based on featured distributions,” *Pattern Recognition*, vol. 29, no. 1, p. 51–59, Jan 1996.
- [97] O. A. Okashi, H. Du, and H. Al-Assam, “Automatic spine curvature estimation from x-ray images of a mouse model,” *Computer Methods and Programs in Biomedicine*, vol. 140, p. 175–184, Mar 2017.
- [98] J. Oramas, K. Wang, and T. Tuytelaars, “Visual explanation by interpretation: Improving visual feedback capabilities of deep neural networks,” in *International Conference on Learning Representations*, 2019. [Online]. Available: <https://openreview.net/forum?id=H1ziPjC5Fm>
- [99] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, p. 62–66, Jan 1979.
- [100] S. Oucharqui, H. Adil, E. Benaissa, F. Bssaibis, I. En-nafaa, A. Maleb, J. El fenni, and M. Elouennass, “Breast tuberculosis: A forgotten diagnosis,” *IDCases*, vol. 26, p. e01341, Jan 2021.
- [101] J. Oyen, C. Brudvik, C. G. Gjesdal, G. S. Tell, S. A. Lie, and L. M. Hove, “Osteoporosis as a risk factor for distal radial fractures: a case-control study,” *The Journal of Bone and Joint Surgery. American Volume*, vol. 93, no. 4, pp. 348–356, Feb 2011.
- [102] M. P. O’Malley, C. Rodner, A. Ritting, M. P. Cote, R. Leger, H. Stock, and J. M. Wolf, “Radiographic interpretation of distal radius fractures: visual estimations versus digital measuring techniques,” *Hand (New York, N.Y.)*, vol. 9, no. 4, p. 488–493, Dec 2014.

- [103] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, p. 1345–1359, Oct 2010.
- [104] B. Z. Phillips, “Wrist joint anatomy: Overview, gross anatomy, natural variants,” Jun 2020, publication: Medscape - eMedicine. [Online]. Available: <https://emedicine.medscape.com/article/1899456-overview>
- [105] S. Pizer, R. Johnston, J. Ericksen, B. Yankaskas, and K. Muller, “Contrast-limited adaptive histogram equalization: speed and effectiveness,” in *[1990] Proceedings of the First Conference on Visualization in Biomedical Computing*. IEEE Comput. Soc. Press, 1990, p. 337–345. [Online]. Available: <http://ieeexplore.ieee.org/document/109340/>
- [106] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, “Adaptive histogram equalization and its variations,” *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, p. 355–368, Sep 1987.
- [107] M. O. Premaor, J. E. Compston, F. F. Avilés, A. Pagès-Castellà, X. Nogués, A. Díez-Pérez, and D. Prieto-Alhambra, “The association between fracture site and obesity in men: A population-based cohort study,” *Journal of Bone and Mineral Research*, vol. 28, no. 8, pp. 1771–1777, Aug 2013.
- [108] C. U. Press.(n.d.), “diagnosis.” [Online]. Available: <https://dictionary.cambridge.org/dictionary/english/diagnosis>
- [109] N. Raby, L. Berman, S. Morley, and G. De Lacey, *Accident and Emergency Radiology: A Survival Guide (Third Edition)*. Saunders Elsevier, 2015.
- [110] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, “Transfusion: Understanding transfer learning for medical imaging,” *arXiv:1902.07208 [cs, stat]*, Oct 2019, arXiv: 1902.07208. [Online]. Available: <http://arxiv.org/abs/1902.07208>
- [111] P. Rajpurkar, J. Irvin, A. Bagul, D. Ding, T. Duan, H. Mehta, B. Yang, K. Zhu, D. Laird, R. L. Ball, and et al., “MURA: Large dataset for abnormality detection in musculoskeletal radiographs,” *arXiv:1712.06957*, Dec 2017. [Online]. Available: <http://arxiv.org/abs/1712.06957>

- [112] T. Randen and J. Husoy, “Filtering for texture classification: a comparative study,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, p. 291–310, Apr 1999.
- [113] R. Ravindrarajah, N. C. Hazra, J. Charlton, S. H. D. Jackson, A. Dregan, and M. C. Gulliford, “Incidence and mortality of fractures by frailty level over 80 years of age: cohort study using uk electronic health records,” *BMJ Open*, vol. 8, no. 1, Jan 2018.
- [114] C. C. Reyes-Aldasoro, K. H. Ngan, A. Ananda, A. d. Garcez, A. Appelboam, and K. M. Knapp, “Geometric semi-automatic analysis of radiographs of colles’ fractures,” *PLOS ONE*, vol. 15, no. 9, p. e0238926, Sep 2020.
- [115] J. Ruskin, D. Seigerman, M. Sirkin, M. Reilly, and M. Adams, “Propionibacterium acnes infection of the shoulder after a manipulation under anesthesia for stiffness status post open reduction and internal fixation proximal humerus: A case report,” *Journal of Orthopaedic Case Reports*, vol. 8, pp. 19–22, 03 2018.
- [116] P. J. Ryan, “Bone densitometry in the management of Colles’ fractures: which site to measure?” *The British Journal of Radiology*, vol. 74, no. 888, pp. 1137–1141, Dec. 2001.
- [117] F. S. Schleich, M. Schürch, M. W. Huellner, U. Hug, U. von Wartburg, K. Strobel, and P. Veit-Haibach, “Diagnostic and therapeutic impact of spect/ct in patients with unspecific pain of the hand and wrist,” *EJNMMI Research*, vol. 2, no. 1, p. 53, Sep 2012.
- [118] D. Shen, G. Wu, and H.-I. Suk, “Deep learning in medical image analysis,” *Annual review of biomedical engineering*, vol. 19, pp. 221–248, Jun 2017.
- [119] M. Shetty, “Imaging and differential diagnosis of ovarian cancer,” *Seminars in Ultrasound, CT and MRI*, vol. 40, no. 4, p. 302–318, Aug 2019.
- [120] C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, no. 1, p. 60, Jul 2019.
- [121] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv:1409.1556 [cs]*, Apr 2015, arXiv: 1409.1556. [Online]. Available: <http://arxiv.org/abs/1409.1556>

- [122] S. Siuly and Y. Zhang, “Medical big data: Neurological diseases diagnosis through medical data analysis,” *Data Science and Engineering*, vol. 1, no. 2, pp. 54–64, Jun 2016.
- [123] A. Stanislavsky, “Chauffeurs fracture.” [Online]. Available: <https://radiopaedia.org/cases/12360>
- [124] C. Sun, A. Shrivastava, S. Singh, and A. Gupta, “Revisiting unreasonable effectiveness of data in deep learning era,” in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017, p. 843–852. [Online]. Available: https://openaccess.thecvf.com/content_iccv_2017/html/Sun_Revisiting_Unreasonable_Effectiveness_ICCV_2017_paper.html
- [125] A. Sureshkumar, B. Hansen, and D. Ersahin, “Role of nuclear medicine in imaging,” *Seminars in Ultrasound, CT and MRI*, vol. 41, no. 1, p. 10–19, Feb 2020.
- [126] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2015, p. 1–9.
- [127] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, “Inception-v4, inception-resnet and the impact of residual connections on learning,” *arXiv:1602.07261 [cs]*, Aug 2016, arXiv: 1602.07261. [Online]. Available: <http://arxiv.org/abs/1602.07261>
- [128] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” *arXiv:1512.00567 [cs]*, Dec 2015, arXiv: 1512.00567. [Online]. Available: <http://arxiv.org/abs/1512.00567>
- [129] A. Tafti and D. W. Byerly, *X-ray Radiographic Patient Positioning*. Treasure Island (FL): StatPearls Publishing, 2022. [Online]. Available: <http://www.ncbi.nlm.nih.gov/books/NBK565865/>
- [130] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, “A survey on deep transfer learning,” *ArXiv*, 2018. [Online]. Available: <http://arxiv.org/abs/1808.01974>

- [131] E. P. Thompson and E. P. Thompson, *Roentgen Rays and Phenomena of the Anode and Cathode: Principles, Applications and Theories*. D. Van Nostrand Company, 1896.
- [132] R. van der Velde, C. E. Wyers, E. Teesselink, P. P. M. M. Geusens, J. P. W. van den Bergh, F. de Vries, C. Cooper, N. C. Harvey, and T. P. van Staa, “Trends in oral anti-osteoporosis drug prescription in the united kingdom between 1990 and 2012: variation by age, sex, geographic location and ethnicity,” *Bone*, vol. 94, pp. 50–55, Jan 2017.
- [133] M. I. Vargas, J. Boto, and T. R. Meling, “Imaging of the spine and spinal cord: An overview of magnetic resonance imaging (mri) techniques,” *Revue Neurologique*, vol. 177, no. 5, p. 451–458, May 2021.
- [134] I. Vergara, K. Vrotsou, M. Orive, S. Garcia-Gutierrez, N. Gonzalez, C. Las Hayas, and J. M. Quintana, “Wrist fractures and their impact in daily living functionality on elderly people: a prospective cohort study,” *BMC geriatrics*, vol. 16, p. 11, Jan 2016.
- [135] S. Waite, A. Grigorian, R. G. Alexander, S. L. Macknik, M. Carrasco, D. J. Heeger, and S. Martinez-Conde, “Analysis of perceptual expertise in radiology – current knowledge and a new perspective,” *Frontiers in Human Neuroscience*, vol. 13, 2019. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fnhum.2019.00213>
- [136] L. Wang and D.-C. He, “A new statistical approach for texture analysis,” *Photogrammetric Engineering & Remote Sensing*, vol. 56, no. 1, p. 61–66, Jan 1990.
- [137] L. Wang and D. C. He, “Texture classification using texture spectrum,” *Pattern Recognition*, vol. 23, no. 8, p. 905–910, Jan 1990.
- [138] H. S. Ward, *Practical Radiography: a Hand-book of the Applications of the X-rays*. Pub. for the Photogram, Limited, Dawbarn & Ward, Limited, 1896.
- [139] T. Webber, S. P. Patel, M. Pensak, O. Fajolu, T. D. Rozental, and J. M. Wolf, “Correlation Between Distal Radial Cortical Thickness and Bone Mineral Density,” *Journal of Hand Surgery*, vol. 40, no. 3, pp. 493–499, Mar.

2015. [Online]. Available: [https://www.jhandsurg.org/article/S0363-5023\(14\)01720-1/abstract](https://www.jhandsurg.org/article/S0363-5023(14)01720-1/abstract)
- [140] J. C. Weinreb and P. A. Wilcox, “How do training, education, and experience affect quality in radiology?” *Journal of the American College of Radiology*, vol. 1, no. 7, p. 510–515, Jul 2004.
- [141] W. M. Wells, “Medical image analysis – past, present, and future,” *Medical Image Analysis*, vol. 33, p. 4–6, Oct 2016.
- [142] M. L. Wilczek, J. Kälvesten, J. Algulin, O. Beiki, and T. B. Brismar, “Digital x-ray radiogrammetry of hand or wrist radiographs can predict hip fracture risk—a study in 5,420 women and 2,837 men,” *European Radiology*, vol. 23, no. 5, pp. 1383–1391, May 2013.
- [143] R. M. D. Zebaze and E. Seeman, “Epidemiology of hip and wrist fractures in Cameroon, Africa,” *Osteoporosis International*, vol. 14, no. 4, pp. 301–305, Jun 2003.
- [144] L. Zhang, H. Wang, Q. Li, M.-H. Zhao, and Q.-M. Zhan, “Big data and medical research in China,” *BMJ*, vol. 360, p. j5910, Feb 2018.
- [145] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, “Random erasing data augmentation,” *arXiv:1708.04896 [cs]*, Nov 2017, arXiv: 1708.04896. [Online]. Available: <http://arxiv.org/abs/1708.04896>
- [146] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, “Learning deep features for discriminative localization,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2016, p. 2921–2929.
- [147] T. Zimmermann, *Galeazzi fracture consist of fracture of the distal radius with dislocation of distal radioulnar joint.*, Feb 2017. [Online]. Available: https://commons.wikimedia.org/wiki/File:Galeazzi_Fracture_of_Distal_Radius.jpg