

AN IMPROVED BONE AGE ASSESSMENT USING
ADVANCED IMAGE PROCESSING AND DEEP LEARNING
APPROACH

LIANG KIM MENG

FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR

2020

**AN IMPROVED BONE AGE ASSESSMENT USING
ADVANCED IMAGE PROCESSING AND DEEP
LEARNING APPROACH**

LIANG KIM MENG

**THESIS SUBMITTED IN FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF
PHILOSOPHY**

**FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2020

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Liang Kim Meng

Matric No: KHA140112

Name of Degree: Doctor of Philosophy

Title of Project Paper/Research Report/Dissertation/Thesis (“this Work”):

An Improved Bone Age Assessment Using Advanced Image Processing And Deep Learning Approach

Field of Study: Biomedical Engineering

I do solemnly and sincerely declare that:

- (1) I am the sole author/writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature

Date:

Subscribed and solemnly declared before,

Witness’s Signature

Date:

Name:

Designation:

AN IMPROVED BONE AGE ASSESSMENT USING ADVANCED IMAGE PROCESSING AND DEEP LEARNING APPROACH

ABSTRACT

Pediatricians often apply bone age assessment to measure the skeletal maturity of children and to predict the future height. These discrepancies are good indicators for diagnosing growth disorders. Normally, left hand skeletal is employed in this assessment. The low quality of ossification sites of carpals deteriorates the pediatrician's visibility in inspecting the pertinent radiographic manifestations. This in turn affects the bone age assessment. Therefore, we have to enhance the quality before assessing them. Histogram equalization is one of the contrast enhancement techniques that suit this type of enhancement. Existing histogram equalizations, however, are confronting with problems in preserving the brightness and details as well as preventing the contrast from being over-enhanced or under-enhanced simultaneously. The comprehensive histogram equalization was proposed by considering all criteria of the desired histogram-equalized image to produce moderately contrast enhanced carpals' ossification sites. Qualitative results show that the determining features of maturity stages have been emphasized in some of the Pareto optimized image. The improvement for Pareto optimized image by bi-histogram equalization is significant for five stages from stage D to stage H with improvement accuracy of 7.16%, 12.47%, 16.03%, 21.21% and 18.51%, respectively. Findings concluded that the Pareto optimized images able to improve the classifier accuracy that estimate the maturity stage of the carpal bones.

Keywords: visual enhancement, bone age assessment, histogram equalization

PENINGKATAN PENILAIAN UMUR TULANG MELALUI PEMROSESAN IMEJ TERMAJU DAN TEKNIK PEMBELAJARAN MENDALAM

ABSTRAK

Pediatrik sering memohon penilaian usia tulang untuk mengukur kematangan rangka kanak-kanak dan juga meramalkan ketinggian masa depan. Percanggahan ini adalah petunjuk yang baik untuk menguji ketidakseimbangan pertumbuhan. Biasanya, kerangka tangan kiri digunakan dalam penilaian ini. Kualiti rendah tapak-tapak ossifikasi carpals menyebabkan penurunan penglihatan pediatrik dalam memeriksa manifestasi radiografi yang berkaitan. Ini seterusnya memberi kesan kepada penilaian umur tulang. Oleh itu, kita perlu meningkatkan kualiti sebelum menilai mereka. Penyamaan histogram adalah salah satu teknik peningkatan kontras yang sesuai dengan peningkatan sedemikian. Walau bagaimanapun, penyamaan histogram sedia ada menghadapi masalah dalam memelihara kecerahan dan butiran serta menghindar kontras daripada terlalu tinggi atau rendah secara serentak. Penyamaan histogram yang komprehensif telah dicadangkan dimana semua kriteria imej histogram-yang disamakan telah dibentangkan untuk menghasilkan tapak ossifikasi yang berkontras sederhana. Keputusan kualitatif menunjukkan bahawa ciri-ciri yang menentukan peringkat kematangan telah ditekankan dalam beberapa imej yang dioptimumkan oleh Pareto. Peningkatan Pareto yang dioptimumkan untuk penyamaan dua histogram adalah ketara untuk lima peringkat, bermula dari peringkat D ke peringkat H dengan peratus peningkatan sebanyak 7.16%, 12.47%, 16.03%, 21.21% dan 18.51%, masing-masing. Pencarian menyimpulkan bahawa imej yang dioptimumkan oleh Pareto dapat meningkatkan ketepatan klasifikasi yang menganggarkan tahap kematangan tulang carpal.

Kata kunci: peningkatan visual, penilaian umur tulang, penyamaan histogram

ACKNOWLEDGEMENTS

Foremost, I would like to express my gratitude to my family members, in order for me to accomplish my thesis entitled “*An Improved Bone Age Assessment Using Advanced Image Processing and Deep Learning Approach*” as the requirement for the Degree of Doctor of Philosophy. This thesis is able to be completed with the kind help and support of many individuals. I would like to extend my sincere thanks to all of them.

Firstly, I would like to express my sincere gratitude to my supervisors, Ir. Dr. Lai Khin Wee, Associate Professor Dr. Belinda Murphy and Ir. Dr. Hum Yan Chai. Their continuous support, patience, motivation and knowledge contributed to the completion of my Ph.D study. They are my best advisors and mentors in my Ph.D study.

Secondly, I sincerely acknowledge Ministry of Higher Education (MOHE) of Malaysia for offering MyBrain scholarship to me, in terms of financially support me throughout my PhD study. Thank you notes also dedicated to Universiti Malaya (UM) for providing me sufficient resources and conducive research environment in pursuing knowledge. Lastly, my sincere gratitude is devoted to all my research colleagues, friends and all the academics staffs as well as the management staffs in my faculty.

TABLE OF CONTENTS

Acknowledgements	v
Table of Contents	vi
List of Figures	viii
List of Tables.....	xi
List of SYMBOLS AND ABBREVIATIONS	xii
List of Appendices	xv
CHAPTER 1: INTRODUCTION.....	1
1.1 Overview.....	4
1.2 Problem Statement.....	7
1.3 Aim and Objectives	9
1.4 Research Scope.....	10
1.5 Thesis Organization.....	12
CHAPTER 2: LITERATURE REVIEW.....	14
2.1 Bone Age Assessment	14
2.1.1 Greulich-Pyle Method	18
2.1.2 TW Method (TW2 and TW3)	20
2.2 Automated Bone Age Assessment Methods.....	23
2.3 Contrast Enhancement	27
2.3.1 Global Histogram Equalization	33
2.3.2 Local Histogram Equalization (LHE)	40
2.4 State-of-the-art Contrast Enhancement Algorithms	44
2.5 State-of-the-art Segmentation Algorithms.....	49
2.6 Summary of Literature Review of Contrast Enhancement and Segmentation	54

CHAPTER 3: METHODOLOGY	58
3.1 Overview of the Contrast Enhancement Methodology	58
3.1.1 Formulation of Multiple Objective Functions	59
3.1.1.1 The Measurement of the Enhanced Contrast	61
3.1.1.2 The Measurement of the Mean Brightness Deviation.....	67
3.1.1.3 The Measurement of Detail Loss	70
3.1.2 Modelling of Analytical Functions.....	73
3.1.3 Optimization of Analytical Function.....	84
3.1.4 Selection of Optimized Solutions.....	89
3.2 Denoising of the Hand Bone.....	91
3.3 Segmenting the Hand Bone	102
CHAPTER 4: RESULTS AND ANALYSIS.....	113
4.1 Qualitative Analysis on Visual Appearance of the Resultant Images	115
4.2 Improved Segmentation Result due to Enhanced Contrast	126
4.3 Quantitative Analysis using Classifier.....	148
CHAPTER 5: CONCLUSION AND FUTURE WORK	153
5.1 Thesis Contribution	153
5.2 Conclusion	153
5.3 Future Work.....	154
List of Publications and Papers Presented	157
References	158
APPENDIX A	170
APPENDIX B	171

LIST OF FIGURES

Figure 1.1: Sample of hand bone image.	2
Figure 1.2: Contrast enhancement on radius ossification site (a) Original Image (b) Histogram equalized image that exhibits an undesired background noise (c) Histogram equalized image that exhibits an undesired intensity saturation and mean brightness shift (d) Histogram equalized image that moderately enhanced the important ossification site without exhibiting any unwanted artifacts.	6
Figure 3.1: Work flow of holistic contrast enhancement framework	59
Figure 3.2: Identification of principal objectives	60
Figure 3.3: Quantitative measurement for contrast, brightness and detail considerations	61
Figure 3.4 The correspondence value between MCM and MWEME with α variation values.....	64
Figure 3.5 (a) Original Image of radius ossification sites (b) Amplified background noise found at the histogram equalized image (c) Unwanted intensity saturation and high mean brightness shift of the histogram equalized image (d) Well enhanced image without obvious unwanted artifacts.....	65
Figure 3.6: Formulation of analytical function model by objective functions, decision variables and constraints.	74
Figure 3.7: Formulation of histogram equalization model.....	75
Figure 3.8: Pre-determined decision variables search ranges of “Clipped Limit – L” (ranges from $AbL - \sigma L$ to $AbL + \sigma L$), “Clipped Limit – R” (ranges from $AbU - \sigma U$ to $AbU + \sigma U$) and “Separating Point”(ranges from AL to AU)......	79
Figure 3.9: Solution Space, X.	79
Figure 3.10: Computation of each objective value of the resultant image after histogram equalization process with the sampled decision variables values (a) Contrast value (b) Brightness Preservation value (c) Detail Preservation value.	82
Figure 3.11: Fitted curve plane for each objective (a) Contrast (b) Brightness Preservation (c) Detail Preservation (d) Combinations of all fitted curve planes in the same Cartesian space.....	82
Figure 3.12: Process flow of NSGA II in obtaining non-dominated solutions.....	89

Figure 3.13: Illustration of the overall procedures including the search ranges determination, objective function modelling, multi-objective optimization solver searching for non-dominated solutions, and selection of representative solutions out of the non-dominated solutions.	90
Figure 3.14: Age and gender distribution in the selected training dataset.....	104
Figure 3.15: Original hand bone radiographic image (left image) and the labelled image (right image).....	104
Figure 3.16: Network architecture of FCN-8.....	106
Figure 3.17: Workflow for image segmentation of carpal bones.....	108
Figure 3.18: Training loss rate (blue), accuracy rate (orange) and validation loss rate (green) with respect to epochs.....	109
Figure 3.19: Learning Rate variation during the training of FCN-8 model.....	110
Figure 3.20: Segmented carpal bone using the trained FCN-8 model.....	110
Figure 3.21 Comparison of (a) DSC and (b) HD with respect to age.....	111
Figure 4.1: Output set of Pareto optimized bi-histogram equalized resultant images of the Hamate bone with its corresponding decision variables and each objective function values, in which $\mathbf{x1}$ represents the separating point, $\mathbf{x2}$ represents the clipped limit, C represents Contrast objective, B represents the mean brightness preservation and D represents the detail preservation, respectively.....	116
Figure 4.2: Output set of Pareto optimized bi-histogram equalized resultant images of the Lunate bone with its corresponding decision variables and each objective function values, in which $\mathbf{x1}$ represents the separating point, $\mathbf{x2}$ represents the clipped limit, C represents Contrast objective, B represents the mean brightness preservation and D represents the detail preservation, respectively.....	119
Figure 4.3: Output set of Pareto optimized bi-histogram equalized resultant images of the Capitate bone with its corresponding decision variables and each objective function values, in which $\mathbf{x1}$ represents the separating point, $\mathbf{x2}$ represents the clipped limit, C represents Contrast objective, B represents the mean brightness preservation and D represents the detail preservation, respectively.....	122
Figure 4.4: Output set of Pareto optimized bi-histogram equalized resultant images of the Scaphoid bone with its corresponding decision variables and each objective function values, in which $\mathbf{x1}$ represents the separating point, $\mathbf{x2}$ represents the clipped limit, C represents Contrast objective, B represents the mean brightness preservation and D represents the detail preservation, respectively.....	124

Figure 4.5: Overall mean for three experts for before and after enhancement	133
Figure 4.6: Boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1, 2 and 3 across all male hand bones with ages ranging from 0 to 18, before applying the contrast enhancement scheme.....	136
Figure 4.7 boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1, 2 and 3 across all male hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme.....	138
Figure 4.8 Boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1, 2 and 3 across all female hand bones with ages ranging from 0 to 18, before applying the contrast enhancement scheme.....	144
Figure 4.9 Boxplot of Dice Coefficient for hand Bone segmentation with ground truth delineated by expert 1, 2 and 3 across all female hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme.....	146
Figure 4.10 F1 score of classification accuracy without enhancement, enhancement using random value of separating point and clipped limit and enhancement using Pareto-optimized.....	152

LIST OF TABLES

Table 2.1: Weighted maturity scores of Radius, Ulna and short bones (RUS) for boy category.....	21
Table 2.2: Maturity scores of Carpal bones for boy category:.....	21
Table 2.3: Weighted maturity scores of Radius, Ulna and short bones (RUS) for girl category.....	22
Table 2.4: Weighted maturity scores of Carpal bones for girl category:.....	22
Table 3.1 Contrast measurement value with MWEME and MCM with different k_1 , k_2 and α values	67
Table 3.2 Validation of segmentation accuracy as measured using DSC and HD	111
Table 4.1: Experimental result for segmentation using Dice Coefficient across 19 age groups for male category. The value of each cell represents the mean dice coefficient of a particular age group using the segmentation scheme against the ground truth by a particular expert.	131
Table 4.2: T-test: Paired two sample for means for Expert 1 for male category.	134
Table 4.3: T-test: Paired two sample for means for Expert 2 for male category.	134
Table 4.4: T-test: Paired two sample for means for Expert 3 for male category.	135
Table 4.5: Experimental result for segmentation using Dice Coefficient across 19 age groups for female category. The value of each cell represents the mean dice coefficient of a particular age group using the segmentation scheme against the ground truth by a particular expert.	140
Table 4.6: T-test: Paired two sample for means for Expert 1 for female category.	142
Table 4.7: T-test: Paired two sample for means for Expert 2 for female category.	142
Table 4.8: T-test: Paired two sample for means for Expert 3 for female category.	143
Table 4.9: Classification Accuracy of Test Images without Enhancement	149
Table 4.10: Classification Accuracy of Test Images with Bi-Histogram Equalization using Random Values of Separating Points and Clipped Limit.....	150
Table 4.11: Classification Accuracy of Test Images with Pareto Optimized Bi-Histogram Equalization.....	150

LIST OF SYMBOLS AND ABBREVIATIONS

$h(r)$: Input Image Histogram
$p(r)$: Normalized Histogram Function
$c(n)$: Cumulative Histogram Function
$T(n)$: Scaled Cumulative Histogram Function
K	: Number Of Objectives
T	: Total Number Of Non-Dominated Solutions
X^S	: Non-Dominated Decision Variables Vector
I	: Input Image
A_I	: Average Intensity Of An Image
$I(i, j)$: Gray-scale Intensity OF A Two-Dimensional Image AT Coordinate (i,j)
H^I	: Histogram Distribution of Intensity for Image I
H^L	: Lower Sub-Histogram With Bin Index from 0 to A_I .
H^U	: Upper Sub-Histogram with Bin Index from A_I+1 to Maximum Gray Level
B	: Bin Index
n_B	: Total Number Of Pixels With Specific Gray Level Intensity Value
A^L	: Average Pixel Value Of Intensity in H^L
N^L	: Total Number Of Bin Values within H^L
A^U	: Average Pixel Value Of Intensity in H^U
N^U	: Total Bin Values Across H^U
$Lmax$: Maximum Gray Level Intensity
x_1^S	: First Element in the Non-dominated decision
x_2^S	: Second Element in the Non-dominated decision
x_3^S	: Third element in the non-dominated decision
σ^L	: Standard Deviation in Bin Value for H^L
σ^U	: Standard Deviation in Bin Value for H^U
AHE	: Adaptive Histogram Equalization
AMBE	: Absolute Mean Brightness Error
BAA	: Bone Age Assessment

BBHE	: Bi-Histogram Equalization
BHEPL	: Bi-histogram with Plateau Limit
BPDHE	: Brightness-Preserving Dynamic Histogram Equalization
BPHEME	: Brightness-Preserving Histogram Equalization With The Maximum Entropy
CASAS	: Computer-Aided Skeletal Age Scoring System
CASMAS	: Computer-aided Skeletal maturity system
CLAHE	: Contrast Limited Adaptive Histogram Equalization
CLAHEDWT	: Contrast Limited Adaptive Histogram Equalization Discrete Wavelet Transform
CHLA	: Children's Hospital Los Angeles
CT	: Computed Tomography
DSIHE	: Dualistic Sub-Image Histogram Equalization
DHE	: Dynamic Histogram Equalization
DQHEPL	: Dynamic Quadrants Histogram Equalization Plateau Limit
EME	: Measurement of Image Enhancement
GHE	: Global Histogram equalization
GP	: Greulich-Pyle
HE	: Histogram Equalization
HMF	: Histogram Modification Framework
LHE	: Local Histogram Equalization
LHS	: Left Hand Side
MCM	: Michelson Contrast Enhancement
MPHEBP	: Multi-Peak Histogram Equalization With Brightness Preserving
MMBEBHE	: Minimum Mean Brightness Error Bi-Histogram Equalization
MR	: Magnetic Resonance
NSGA	: Non-dominated Sorting Genetic Algorithm
PNSR	: Peak Noise Signal Ratio
PDF	: Probability Density Function
PLT	: Piecewise Linear Transformation
RUS	: Radius, Ulna and Short bones
RHS	: Right Hand Side

RMSHE	:	Recursive Mean-Separate Histogram Equalization
RSIHE	:	Recursive Sub-Image Histogram Equalization
RSWHE	:	Recursively Separated and Weighted Histogram Equalization
SCF	:	Structural Comparison Function
SSIM	:	Structural Similarity Index
TW	:	Tanner-Whitehouse
WCHE	:	Weighted Clustering Histogram Equalization

University of Malaya

LIST OF APPENDICES

Appendix A: Distribution number of hand bone image	170
Appendix B: Matlab codes.....	171

University of Malaya

CHAPTER 1: INTRODUCTION

This chapter describes an overview of the project. The objective of this project is to derive an accurate quantitative measurement for measuring physical maturity, where this measurement is not a trivial task. The common approach that utilizing standing height as a maturity measurement is also not a suitable measurement for physical maturity. For example, a child that is taller than other children of the same age may not indicate that this particular child is having a higher maturity level from others. Thus, standing height measurement and chronological age are not appropriate to measure maturity. However, there is specific biological events that is more appropriate for maturity measurement. The biological events for both genders may include presence of tooth eruption, first menstrual period, appearance of pubic hair, breast and penile development. With the occurrence of these biological events, the maturity level for each child can be compared. For example, an individual that had undergone these biological events is more matured as compared to an individual that is yet to undergo these events. Therefore, the sequence of the occurrence of the biological events may become the good indicator of maturity pathway of a child.

Although the occurrence of biological events may be used as an indicator to physical maturity measurement, the age development is considered as incomplete if the occurrence of these events that are not closely spaced (Tanner, 1994). With the incompleteness of the age development shown in occurrence of biological measurement as maturity indicator, bone measurement based on hand and wrist is gaining popularity as a better measurement technique (Tanner, Oshman et al., 1994). The suitability of this measurement is due to the nature of this type of bone has development sequence and its occurrence is closely spaced. In this bone development sequence, specific bones are exists during fetal life and other bones only appear after certain age of a child (Aicardi, Vignolo et al., 2000). Each specific

bone at each development stage is able to be recognized through radiograph image. Each development stage is differentiated by the invariant event sequences of the total of twenty bones in hand and wrist. The twenty bones includes radius, ulna, metacarpals, phalanges and carpals which are developed throughout the maturity pathway. Therefore, the maturity level is measurable by recognizing the evidences that are gathered from the skeletal development of hand and wrist bones. This type of inspection from hand bones development in order to measure the bone age of a children is known as Bone Age Assessment (BAA). **Figure 1.1** illustrates a sample of hand bone image with its bone structures.

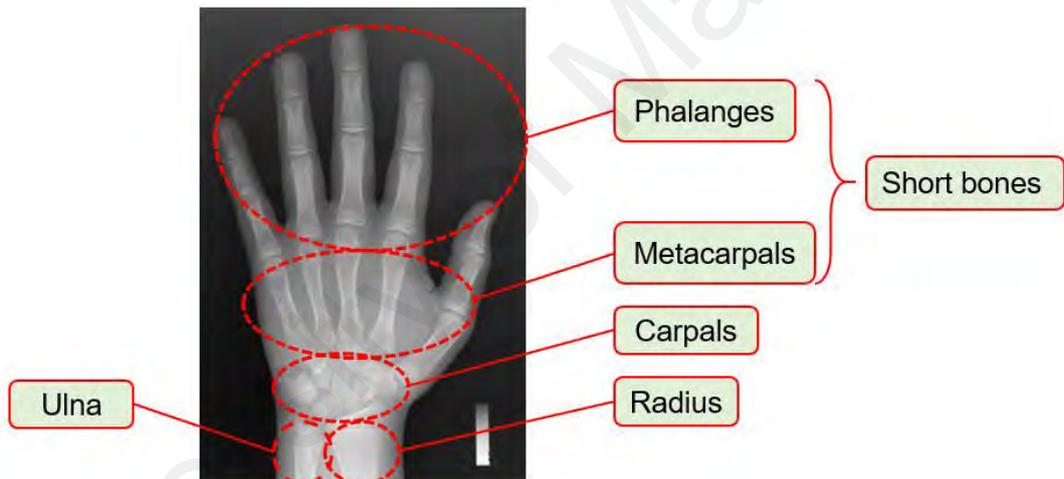


Figure 1.1: Sample of hand bone image.

BAA as a bone maturity assessment technique is applied to measure the skeletal development of a children and adolescents (Cao, Huang et al., 2000). The skeletal measure is used to gauge the skeletal maturity age as an indicator to any growth disorders of a children. This measurement also used as an indicator to predict the future height (Martin, Deusch et al., 2009). In addition, the discrepancy between the skeletal age and biological age is capable to detects endocrine disorder (Gertych, Piętko et al., 2007), chromosomal disorders (Hsieh, Liu et al., 2010), early sexual maturity and others abnormalities (Hsieh, Liu et al., 2010). With the availability of various skeletal bones, left

hand bone is chosen and proven to be highly suitable to be used in skeletal assessment. Thus, the inspection to the development of skeletal bones in left hand is assigned to measure the skeletal maturity. In this, the skeletal maturity is evaluated based on the manifestation that includes ossification area development and the position of calcification (Gertych, Piętka et al., 2007).

In skeletal assessment, pediatricians is embracing two common methods (Hsieh, Liu et al., 2010): Greulich-Pyle atlas method (William W Greulich, 1959) and Tanner-Whitehouse (TW2) methods (Landry, Raiside et al., 1979). In Greulich-Pyle method, pediatricians compare the left hand bone radiograph to the standard atlas in order to measure the skeletal age. On the other hand, TW2 method is an index system that measure the skeletal age through maturity point collection. Then, the TW2 method is extended to TW3 method. The differences between these two methods is depicted by the mechanism of the maturity points are collected. TW2 collects the maturity points from twenty bones score that consists of Radius, Ulna and Short bones (RUS) and Carpals. In contrast, TW3 collects the maturity points from RUS and Carpals, separately. It is claimed that the TW3 is more superior to TW2 due to the collectiveness of the maturity points from twenty bones that are calculated in TW3.

Although TW2 and TW3 is the current most promising skeletal measurement method, their efficiency and effectiveness are still limited to the physician background knowledge and excessive time consumption during measurement. This is due to the TW2 and TW3 methods are still been carried out as a visual inspection way. With this limitation, many computer-aided systems of BAA have been established, recently. For this aided system, TW2 is mainly chosen as a method to be used as this method is more appropriate for computerization purpose. However, most of the systems are still at its infancy due to its inability to measure skeletal age consistently and autonomously.

1.1 Overview

Bone age assessment is used as a clinical inspection by the pediatricians in order to detect any early occurrence of growth disorder in children. This is done by assessing the discrepancy between skeletal bone age and their chronological age (Giordano, Spampinato et al., 2010, Zhang, 2007). The most suitable visual inspection method is Tanner-White method (Babalola, Cootes et al., 2008, Hsieh, Jong et al., 2007). In TW method, pediatricians accumulate maturity points according to a development stages in specified ossification sites of left hand bone, by using human visual system. The total accumulated points is transformed to skeletal bone age based on a defined transformation table.

As the TW method is a manual visual inspection by human vision, the result obtained is subjective and time consuming. With this weaknesses, many research efforts is in place to measure the skeletal bone age, computationally. Computer-aided skeletal age scoring system (CASAS) as computer-aided systems (Thodberg, 2009) is created. CASAS is automatically gauge the skeletal bone age by interpreting the ossification sites in the radiograph image. However, there is very less works on enhancing the visibility of the ossification sites in radiograph image. The quality of the visibility of the ossification sites in radiograph image is important for manual and automated approach to measure the bone age. For automated system, the quality of the radiograph image is crucial for the internal processes in CASAS to work well. The internal process includes segmentation and classification functionalities. With inferior quality of the ossification sites, CASAS may make a wrong decision on the maturity stage of the ossification sites. On the other hand, the improved quality of the ossification sites in the radiograph image also will assist the pediatricians to perform manual visual inspection using GP or TW methods. The existing visual quality of ossification sites in radiograph image may deteriorates the perception of pediatricians in measuring accurate skeletal bone age through manual visual inspection.

Contrast enhancement as one of the approaches used in enhancing the quality of radiography image. This approach shows the ability to enhance images to an acceptable quality in medical image processing. In this thesis, the histogram equalization (HE) as a contrast enhancement technique is chosen as a focus approach. This technique increases the contrast of the analyzed image by stretching the dynamic range of histogram of the input image. HE is chosen from other heuristic approaches due to this technique is mathematically proven as a uniform distributed contrast enhancement technique that performs intensity remapping to entire image intensity range. However, it is known that the common limitation of this enhancement technique is the brightness shift, loss of detail, and under or over enhanced contrast. Thus, a significant number of literatures is focusing in improving limitations of HE.

From the literatures that is related to HE, the major problem of this enhancement technique is its improvement is too focusing on certain criterion. For example, Minimum Mean Brightness Error Bi-Histogram Equalization (MMBEBHE) (Soong-Der and Ramli, 2003) focuses on brightness shift by minimizing the absolute mean value of the brightness error. However, MMBEBHE has minimum consideration on the enhanced contrast and detail loss. Hence, the processed image may still poor quality as the over enhanced image contains undesired washed out artifacts that will obstruct important details in the image. The loss in detail has affected the quality of the skeletal age assessment as the important details at the ossification sites resembles the skeletal development stage. The second problem of HE is the metric definition that calculates the efficiency of HE is not resembling human visual perception. Although there is a number of proposed metric such as PNSR, AMBE and Entropy (Sundaram, Ramar et al., 2011, Soong-Der and Ramli, 2003, Rajavel, 2010), the alignment of the metric definition to human visual perception is still neglected by researchers and practitioners in contrast enhancement field.

It is a desire to design an optimized and holistic histogram equalization that is able to take comprehensive consideration on the detail loss, brightness and enhanced contrast simultaneously. It is also highly crucial for the histogram equalization to align to the human visual perception in order to emphasize the pertinent information of the input image and remove the unwanted noises (Chai, Swee et al., 2013). This thesis aims to deliver the optimum enhancement solution by deriving a comprehensive histogram equalization technique and at the same, focusing on the new direction in achieving a natural visual effect of the enhanced image that resembles human visual perception. In this thesis, the scope is restricted to histogram equalization enhancement on carpals ossification sites in bone age assessment. Figure 1.2 depicts the examples of resultant images from contrast enhancement on radius ossification site: (a) Original Image (b) Histogram equalized image that exhibits an undesired background noise (c) Histogram equalized image that exhibits an undesired intensity saturation and mean brightness shift (d) Histogram equalized image that moderately enhanced the important ossification site without exhibiting any unwanted artifacts.

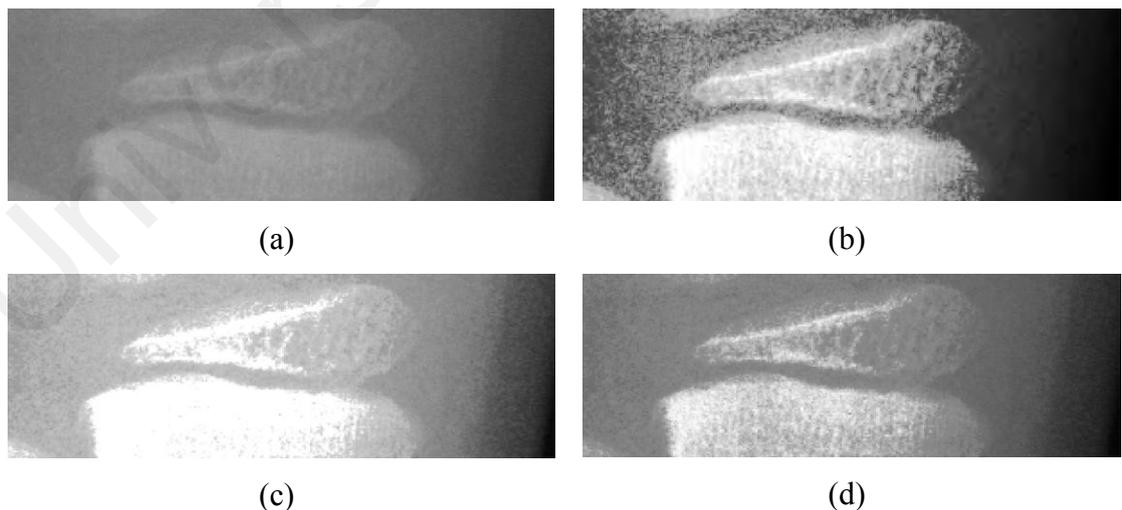


Figure 1.2: Contrast enhancement on radius ossification site (a) Original Image (b) Histogram equalized image that exhibits an undesired background noise (c) Histogram equalized image that exhibits an undesired intensity saturation and mean brightness shift (d) Histogram equalized image that moderately enhanced the important ossification site without exhibiting any unwanted artifacts.

1.2 Problem Statement

In bone age assessment, the low quality of the ossification sites in the radiograph image affects the visibility of the pediatrician in inspecting the pertinent manifestation that related to bone maturity and also the accuracy of the classification results from computer-assisted skeletal age assessment. Due to the low visibility, pediatrician may derives an incorrect maturity stage of the bone assessment. Therefore, it is highly important to enhance the quality of the ossification sites appearing in the radiographic image. From vast research studies, histogram equalization (HE) is known as the most suitable enhancement technique for this purpose. In this thesis, two histogram equalizations: bi-histogram equalization and clipped histogram equalization are examined. These two equalization techniques flat the bins height of the histogram that is computed from the standard radiograph image. In order to do the flattening, the decision variables that are related to these two equalization techniques are determined. However, the common practice is that these two decision variables are determined arbitrarily. The decision variables is determined by using brute-force method. However, this method utilizes high computation resources as it involves computation of dual objectives with dual decision variables. Therefore, feasible computation complexity technique is necessary to solve the problem of finding the optimal decision variables.

The hand bone radiograph image suffers from various noise artifacts that include low background brightness, low contrast and over-enhanced contrast. This variation impacts the rating of the bone age assessment in classifying maturity stage. The Global Histogram equalization (Bull, Edwards et al., 1997) is highly adopted image contrast enhancement technique for the purpose of noise removal. However, this technique does not deliver satisfying results. In the recent research studies, a histogram equalization that preserves the brightness, contrast and details are in high demands. Nevertheless, producing a well-

balanced histogram equalized radiograph in terms of its brightness preservation, detail preservation and contrast enhancement is a challenging task.

In this thesis, we hypothesize that optimal decision variables from the bi-histogram equalization and clipped histogram equalization will preserve the mean brightness, details and contrast without over or under enhancement. Both optimal decision variables are obtained within feasible computational complexity. These optimal decision variables are used to generate enhanced ossification sites of the hand bone radiographic image. This enhancement is significant to either computer-assisted skeletal age assessment system and radiologist clinical examination.

Another problem that arises subsequent from the abovementioned problem is the segmentation in order to determine the ossification sites. This is known as segmentation of sub-images. The sub-image contains the pertinent manifestation of ossification sites. Within this problem context, various research works in performing automated segmentation is studied so that this automated method is relatively more effective than fundamental segmentation techniques. The segmentation problem is a challenging task due to various factors and will be explained in details, subsequently. Firstly, the intensity variation in hand radiograph image reduces the precision consistency of the segmentation technique. The variation in hand radiograph image is caused by different input sources and different age groups of the hand bone provider. Secondly, the intensity distribution in the hand bone image is inconsistent due to the various anatomical density at different parts in the hand bone. This inconsistency intensity property has low inter-variance and high intra-variance of the intensity distribution. For low inter-variance intensity distribution, it is observed that there is an overlapping range of intensity that appears in some parts of the hand bone, such as the cancellous bones, soft tissues regions and the compact bones. For high intra-variance intensity distribution, even the same parts in the

hand bone may exhibit non-uniform intensity range. This non-uniformity of intensity is found in cancellous bone or the cortical bone. In addition, the brightness difference between edge border of compact bone and soft-tissue regions is small. The similar small brightness difference also appeared at the edges between soft-tissue regions and background. The small brightness difference at the edge further complicates the segmentation problem. With the above mentioned challenging factors, most of segmentation technique in the bone age assessment is done in manual way that is time consuming.

1.3 Aim and Objectives

In this thesis, the enhancement framework is designed and demonstrated in order to enhance the visual quality of the ossification sites in left hand bone in radiographic image. The aim of this visual quality enhancement is to obtain a resultant image that takes three objectives into considerations, which is the contrast enhancement, brightness preservation and detail preservation. In addition, this enhancement framework needs to be parallel to human perception that is capable of emphasizing the pertinent information and expelling the unwanted noises. The enhancement framework is specifically design in such a manner in order to derive an input image for the following procedure that highlights all the pertinent information, which is the segmentation procedure.

After the contrast enhancement procedure, the design framework requires the area of the ossification sites to be segmented (Michael and Nelson, 1989, Levitt, Hedgcock Jr et al., 1993, Manos, Cairns et al., 1993, Manos, Cairns et al., 1994, Morris and Walshaw, 1994). It is noted that the aim of the segmentation process is to remove background, noise and any regions in the image that do not contains any important information for the maturity classification purpose. As segmentation process is one of the significant initial

stages of the BAA system, the accuracy and effectiveness of the resultant from segmentation process is highly important as the segmentation output determines the accuracy of the maturity classification results.

The objectives of this thesis are further specified as following:

- 1) Derive a contrast enhancement framework that enhances the quality of the hand bone image that simultaneously considers brightness preservation, contrast enhancement and detail preservation.
- 2) Derive an effective and automated segmentation framework that segments the ossification sites with computationally feasible and less demanding large dataset for training.

1.4 Research Scope

In this thesis, the research scope includes major points that contributes to the outcome of the thesis. The major points include the used datasets for experimental analysis, deployed methodology and result analysis are discussed and clarified. The scope of this thesis project is defined as follows.

- 1) Research focuses on the design and implementation of multi-objective contrast enhancement and segmentation framework on specific bones such as Carpals of left hand bone radiographs for computer-aided skeletal age scoring system (TW). The left hand bone radiograph image is in gray scale mode.
- 2) Algorithms developments for the proposed contrast enhancement and segmentation framework are developed using the commercialized software

MATLAB version R2017b. The developed algorithm focuses on spatial domain image analysis.

- 3) In the algorithm design, development, testing and analysis of the multi-objective contrast enhancement framework, the image dataset comprises of 1390 hand bone radiograph images. This online hand bone images are collected and contributed from Children's Hospital Los Angeles (CHLA). They are downloaded from the available online hand bone database (Joseph, L., 2005). This online hand bone image includes two genders and four ethnicity categories, which is Caucasian, African American, Hispanic, and Asian. The images contain ages range between 0 to 18 years old. The hand bone images various from clear to blur quality and the available format is DICOM and JPEG. In this research work, only JPEG image is used. The resolution of the JPEG format is various from approximate 928 x 1218 to 1904 x 2653. The distribution of number of hand bone image for different ethnicity group and gender is depicted in Appendix A.
- 4) With regards to the testing result and analysis comparison, the efficiency of proposed algorithm is evaluated with two approaches:
 - (a) Analytical evaluations that compare the major objectives of the proposed algorithms within the scope of this thesis.
 - (b) Empirical evaluations that accesses the quality of the processed image from the proposed contrast enhancement and segmentation framework with the reference to standard image quality metrics.

1.5 Thesis Organization

The outline of the thesis is organized into five chapters as follows.

Chapter 1 describes an introduction overview to the background of the research problem, research motivation and problem statements in this thesis. This chapter also provides the objective statements to be achieved by this thesis in research mode. It defines the boundaries of the research and lastly signify the contribution of the research to the community that works in the similar research topic.

Chapter 2 provides an overview of the bone age assessment, contrast enhancement and segmentation. This chapter explains and reviews previous studies on automated skeletal bone age assessment as well as various type of contrast enhancement and segmentation techniques. This chapter covers the literature review of the fundamental background, mathematical formulation and proposition comparison of various related contrast enhancement and segmentation techniques. Critical appraisal on the usability and efficiency towards quality enhancement and segmentation output of the hand bone radiographs is given for a number of related techniques. The aim of this chapter is to delivers sufficient background in order to understand and appreciate the research problems.

Chapter 3 describes the novel algorithm method for multi objective contrast enhancement and segmentation for bone age assessment purpose. This chapter presents the design and implementation of the proposed algorithm framework. This framework consists of few modules. First module is devoted to the identification of the desired properties of desired contrast. The next module describes the method in creating the histogram equalization technique that take into consideration of several critical criteria. In the last module, the technique for setting the optimum decision variable value that influence the optimum quality of the bone image is illustrated. The purpose of this chapter

is to establish the methodology of the proposed multi-objective contrast enhancement that is crucial to the success of an automated skeletal age assessment system. This chapter also elaborates the proposed segmentation approach used in bone age assessment system.

Chapter 4 presents the results from various qualitative and quantitative analysis generated by the proposed algorithm that is described in Chapter 3. This chapter focuses on delivering the qualitative and quantitative analysis result in order to justify the achievement of the claimed objective. The aim of this chapter is to evaluate the performance of each proposed module and the overall automated contrast enhancement and segmentation framework. Each experimental result is analyzed and followed by discussion on the impact of the result towards the research outcome in this chapter.

Chapter 5 concludes the thesis and highlights the main conclusions that are delivered from this research. This chapter illustrates the contributions of the current research works and potential future direction works. The aim of Chapter 5 is to summarize the research findings and suggestions for future research exploration that is relevant and crucial for the development and extension of the proposed framework.

CHAPTER 2: LITERATURE REVIEW

This section is devoted to an overview of the standard histogram equalization method and segmentation techniques that are adopted in medical applications. The fundamental concept, advantages and limitations of each technique are discussed. This section also illustrates the unsuitability of the standard techniques of histogram equalization and segmentation technique in the context of computer-aided skeletal age scoring system. The standard technique is analyzed in the hand bone images for suitability evaluation. This evaluation and implementation of previous techniques in hand bone image enhancement and segmentation are crucial to motivate the objective of this thesis and subsequently, justify the contribution of this thesis. This section ends with the conclusion that there is a need to have better technique of hand bone image enhancement and segmentation for skeletal age assessment purpose instead of using the standard enhancement and segmentation techniques.

2.1 Bone Age Assessment

Bone age assessment (BAA) is a clinical application. The purpose of this application is to evaluate the skeletal development of a person especially in children and adolescents (Liu, Qi et al., 2008). As the chronological age is not suitable to determine the growth maturity of a person, skeletal age is gaining popularity as an indicator to measure the maturity stage of a person. With the known maturity stage, any growth disorder can be early determined. On the other hand, the skeletal age is used to predict the body height of a person. The left hand radiograph image is commonly used as a reliable indicator to measure the bone maturity. The development of ossification area and calcium position in the ossification area in the left hand determines the maturity stage. The discrepancy between the computed skeletal age and its corresponding biological age is useful to detect

diseases that includes endocrine disorders, chromosomal disorders and early sexual maturation (Gerstenfeld, 2003). In the following section, the skeletal development is further discussed prior to the explanation on maturity measurement.

In the context of measuring the stage of skeletal maturity, it is common to relate this type of maturity to bone characteristic and bone mineralization (Mahmoodi, Sharif et al., 2000, Pietka, Gertych et al., 2001, Sebastian, Tek et al., 2003, Mentzel, Vilser et al., 2005). Example of bone characteristic is size and shape of the bone. In measuring the skeletal development, it is highly important to consider the bone development progress. In this development, there is one type of the growth that known as longitudinal growth. This growth happened at the long bones where the process called “endochondral ossification” is occurred (Chai, Swee et al., 2013). The primary ossification is a process take place in the bone growth where the bone tissue is formed at the middle of the long bones. For the other long bones, the ossification is developed from the secondary center, which is located at the cartilage of bone extremities (Sato, Ashizawa et al., 1999). Another obvious bone development is the growth of bone width, which the growth is according to the increase of tissues from the fibrous membrane. The bone growth also can be noticed at flat bones of pelvis, ossification of the calvarium and scapulae (Verma, Peltomäki et al., 2009). It is known that the majority of the flat bone undergo ossification that is originated from the primary center.

In the bone development, there is two main parts in the long bone. Two main parts are known as diaphysis and epiphysis, respectively. Diaphysis refers to the bones where the ossification happens from the primary center, while epiphysis is from secondary center. The diaphysis and epiphysis are continue to grow until the epiphyseal plate is stop growing. Epiphyseal plate is a thin layer of cartilage at the end of the bone. At the end of the bone development, the bone cells will stop multiplying themselves and the ossification

of epiphyseal plate will starts. At the end of the growth of epiphyseal plate, diaphysis and epiphysis are combined and this process marks the end of the bone growth.

It is also an interest in the bone development study to further investigate the bone growth at the fetal stage. At this stage, the possibility of prematurity is detected by monitoring the skeletal growth during fetal stage. The fetal stage started when the embryonic period ended. At this fetal period, majority of the primary centers at the tubular bones are formed into diaphysis. This period takes about 13 weeks for completion. Within this period, the diaphysis is ossified, however epiphyses remains being cartilaginous. For secondary center, it is only complete in development at the infant stage. At the 40 week of fetal growth, the proximal epiphysis ossification center begins to develop. However, epiphysis at femur and tibia are yet to develop until the baby is born. After the baby is born, the epiphysis begins its growth with the ossification process and progressively combining with the main parts of the bone. The following section describes the explanation on skeletal maturity measurement.

In skeletal maturity measurement, the measurement is evaluated base on the growth level of the epiphysis to the standard reference growth. Although the determinant factor towards bone growth is yet to be clearly identified, it is related to the genetics and hormones. Studies highlighted that the gene of the estrogen receptor plays a crucial role towards the growth of epiphysis. However, estrogen may not be the only factor that determines the bone growth. This type of maturity measurement is known as bone age or skeletal age.

There are two established evaluation system in the skeletal maturity measurement. The two systems are Greulich-Pyle and Tanner-Whitehouse atlas (TW2). In Greulich-Pyle system, the patient's hand bone radiograph is manually compared with the reference atlas. The outcome of the comparison determine the skeletal age. TW2 is different from

Greulich-Pyle system, as TW2 determines skeletal maturity base on point collection index. Both system utilizes the left hand and wrist radiographs. This type of radiograph image is used due to few reasons. The first reason is due to the hand and wrist comprises of high number and type of bones. The next obvious reason comes from its easiness to acquire the radiograph as compared to others body parts of the body. The left hand and wrist is not chosen from the right side due to the possibility of injury at the right side as it is common used in accomplishing daily tasks. Left side is preferred due to the legacy selection done in the early 1900s where the physical anthropologists preferred the left side as the element to be used in the physical measurement purpose. However, the reliability and efficiency of both the skeletal evaluation system are still in questions as both system is carried out base on manual visual inspection. This type of inspection is subjective, high time consumption and the accuracy is highly dependent to the knowledge background of the physician. Thus, a number of automated bone age assessment (BAA) system are developed for this purpose. The BAA system is mainly built base on the TW2 method, which is more appropriate in terms of index value computation. However, the automated system is still at its infancy stage due to its insufficient stability in the system.

Apart of the efficiency in evaluation the maturity index from the BAA system, it is crucial to ensure the appropriate and correct steps are undertaken during the radiograph images are acquired. There exists a proper guideline for taking hand radiographs, where the position of the hand is very important. Incorrect hand position during acquisition stage will impacts the visibility of some of the important features of the bone in the resultant radiograph. In the guideline, the correct side of the hand to be taken for radiograph is the left side of the hand. During the acquisition process, the palm face of the left hand has to be in downward position and it has to be in contact with the cassette. The right contact is determine by the axis of the middle finger has to be aligned with the axis of the forearm. Meanwhile, the upper arm and forearm are aligned in the same horizontal plane. In terms

of finger positions, all fingers are separated evenly and the thumb finger is positioned at the comfortable rotation angle, which is about 30 degree with the first index finger. Additional aids such as bandage or tape can be applied to the children during acquisition of hand radiograph images. In the following sections, the details of the two popular skeletal age assessment system is elaborated.

2.1.1 Greulich-Pyle Method

Bone age is one of the main indicators to reflect the skeletal development of children. The bone age or also known as biological age is a good indicator for growth disorder in children. For example, it is a strong indicator of abnormal growth of a children if the chronological age of this children is 8 years old and the bone age is only 5 years old. This kind of assessment is called bone age assessment. Bone age is also useful in sorting out legal issues for children criminal case and estimating the age of children refugees, which the birth date of this children is not available for reference. For legal cases, bone age is important to determine the correct age as the legal punishment is sentenced base on the age of the criminal. For refugee cases, school misplacement can be happened if the bone age is wrongly measured.

The standard way of measuring the skeletal age is using Greulich-Plye (GP) method. In this method, the measurement is conducted base on the reference atlas. This reference atlas was originated by Dr. Sarah Idell Pyle together with Dr. William Walter Greulich in 1959. The atlas is created with the data that is obtained from Professor Wingate Todd, which lead to a project called as the “Brush Foundation Study of Human Growth and Development”. This project focuses on Caucasian children in the United States of America. The atlas comprises of x-ray left hand images of children with both genders and their age is ranged from zero to 19 years old. With the atlas, the left hand radiographic

image is compared with the reference atlas. The bone age is assessed based on the nearest similarity found between the inspected radiograph and reference atlas. In this method, the comparison can be done due to the concept that the ossification sites in the hand and wrist bones are formed in a specific sequential order. Therefore, the bone age is determined by comparing the most similar matched of the ossification sites in the hand and wrist bones radiographic image with the reference atlas.

GP method classifies radiographic images to bone age group, where each group corresponds to specific age. This is done by referring to the reference atlas, which is developed by Greulich and Pyle in 1959. The characteristic of this atlas includes a series of standard radiographs of children that is labelled with specific maturity scale. In the implementation, the inspected radiograph is similar match with the series of standard radiographs in the atlas. The best match standard radiograph represents the corresponding bone age in the atlas. Although the GP is the potential method to be used in the skeletal age assessment, this method has three limitations. The first limitation is the best match radiographs in the atlas may also contain major differences in the inspected radiograph. The limited guidance to perform the matching is the second limitation in this method. Guidance is crucial in this method as manual matching is prone to errors. The third limitation is due to the validity of the series of standard radiographs in the atlas. The validity is questionable as these radiographs are taken between 1931 and 1942. The old radiographs do not represent the current changes of bone structure in today's children. Due to the limitations, GP method are seldom in use.

2.1.2 TW Method (TW2 and TW3)

TW method measures maturity score of each bone in the hand and wrist structure. Thus, the understanding about the anatomical structure in the hand and wrist is highly correlated with the maturity score of each bone. The anatomical structure of human hand and wrist consist of radius, ulna, 19 short bones (5 metacarpals and 14 phalanges) and 7 carpals. The main substance in radius, ulna and short bones is the endochondral ossification, while intramembranous ossification forms the carpal bones. The rate of skeletal maturation at carpal bones has high variability from one individual to another individual. The carpal bones reach the full maturity rate faster than the long and short bones.

TW2 method comprises of three type of measurement method that is involves with different set of bones. The first set of TW2 method involves the 13 bones that consists of radius, ulna and short bones (RUS). The second set of TW2 method comprises of the 7 carpal bones. The third set of TW2 method is the combination of the first and second set, which contains 20 bones of radius, ulna, short bones and carpal bones. In general, TW2 method measures bone age base on scoring technique. In this scoring technique, each bone is classified to different maturity stage where each stage consist of a score. After the score for each bone is determined and accumulated, the final score is mapped to a reference graph and the bone age is calculated from the mapping. In TW2 method, the matching graph is created base on the radiographs that are obtained from United Kingdom children from 1950 to 1960. In 2001, the maturity score and bone age in the matching graph are updated. The update is known as Tanner-Whitehouse 3 (TW3) method. The standard deviation between TW2 and TW3 is approximately about 12 months.

The Table 2.1 illustrates the weighted maturity scores of Radius, Ulna and short bones (RUS) for boy category (Tanner et al., 2001):

Table 2.1: Weighted maturity scores of Radius, Ulna and short bones (RUS) for boy category

	A	B	C	D	E	F	G	H	I
Radius	0	16	21	30	39	59	87	138	213
Ulna	0	27	30	32	40	58	107	181	
Metacarpal									
I	0	6	9	14	21	26	36	49	67
III	0	4	5	9	12	19	31	43	52
V	0	4	6	9	14	18	29	43	52
Proximal Phalanx									
I	0	7	8	11	17	26	38	52	67
III	0	4	4	9	15	23	31	40	53
V	0	4	5	9	15	31	30	39	51
Middle Phalanx									
III	0	4	6	9	15	22	32	43	52
V	0	6	7	9	15	23	32	42	49
Distal Phalanx									
I	0	5	6	11	17	26	38	46	66
III	0	4	6	8	13	18	28	34	49
V	0	5	6	9	13	18	27	34	48

The Table 2.2 illustrates the maturity scores of Carpal bones for boy category (Tanner et al., 2001):

Table 2.2: Maturity scores of Carpal bones for boy category:

	A	B	C	D	E	F	G	H	I
Capitate	0	100	104	106	113	133	160	214	
Hamate	0	73	75	79	100	128	159	181	194
Triquetral	0	10	13	28	57	84	102	124	
Lunate	0	14	22	39	58	84	101	120	
Scaphoid	0	26	36	52	71	85	100	116	
Trapezium	0	23	31	46	66	83	95	108	117
Trapezoid	0	27	32	42	51	77	93	115	

The Table 2.3 illustrates the weighted maturity scores of Radius, Ulna and short bones (RUS) for girl category (Tanner et al., 2001):

Table 2.3: Weighted maturity scores of Radius, Ulna and short bones (RUS) for girl category

	A	B	C	D	E	F	G	H	I
Radius	0	23	30	44	56	78	114	160	218
Ulna	0	30	33	37	45	74	118	173	
Metacarpal									
I	0	8	18	18	24	31	43	53	67
III	0	5	8	12	16	23	37	47	53
V	0	6	9	12	17	23	35	48	52
Proximal Phalanx									
I	0	9	11	14	20	31	44	56	67
III	0	5	7	12	19	27	37	44	54
V	0	6	7		18	26	35	42	51
Middle Phalanx									
III	0	6	8	12	18	27	36	45	52
V	0	7	8	12	18	28	35	43	49
Distal Phalanx									
I	0	7	9	15	22	33	48	51	68
III	0	7	8	11	15	22	33	37	49
V	0	7	8	11	15	22	32	36	47

The Table 2.4 illustrates the weighted maturity scores of Carpal bones for girl category (Tanner et al., 2001):

Table 2.4: Weighted maturity scores of Carpal bones for girl category:

	A	B	C	D	E	F	G	H	I
Capitate	0	84	88	91	99	121	149	203	
Hamate	0	72	74	78	102	131	161	183	194
Triquetral	0	11	16	31	56	80	104	126	
Lunate	0	16	24	40	59	84	106	122	
Scaphoid	0	24	35	51	71	88	104	118	
Trapezium	0	20	27	42	60	80	95	111	119
Trapezoid	0	21	30	43	53	77	97	118	

2.2 Automated Bone Age Assessment Methods

The result from bone age scoring as observed in TW method is subjective, time consuming, tedious, domain knowledge dependency and inconsistent evaluation from different physicians. These limitations from manual inspection of bone age scoring approach motivate an automatic method to measure bone age, precisely. The automatic bone age assessment method is also motivated by the nature of the TW2 measurement method, which the score calculation is executed at higher speed and more accurate by computer as compared to human judgement. The motivations inspire the invention of digital CASAS in which the digital radiograph is used. The digital radiograph is produced from medical equipment as an input to the CASAS for bone age assessment.

As compared to GP method and TW method, CASAS is performed better in terms of minimizing the gaps of intra-observer variability and inter-observer variability. From one of the observation studies done by Tanner and others in 1994, three observers are involved in the evaluation of 57 radiographs of normal children with bone age between 4.5 years old to adult year. From the evaluation result, the discrepancy rate for manual evaluation is 17% within the same observers and 5% from the CASAS. The same result pattern also is found between observers, where the discrepancy rate for manual evaluation is 27% as compared to 9% if the evaluation is done using CASAS. From this study, it is a clear indication that CASAS is better and more reliable in assessing bone age as compared to manual evaluation as the perceived bone age in the latter evaluation is inconsistency although repetition is made.

The digital CASAS at the early stage is still operating in the semi-automatic mode. In this mode, the human operator needs to select the radial epiphyseal area on the digitized inspected radiograph that is displayed on the viewing screen. After the area selection, each of the bone template from the reference template set is overlaid on the inspected radiograph. For each overlay, the human operator needs to manually fit the template onto

the inspected radiograph by resizing and rotating steps. Once the matching template is found, the available grading button will be pressed in order to obtain the corresponding bone age with the chosen template, automatically. Therefore, the infancy stage of CASAS still demands human operator to intervene at the initial process and the system will perform the final alignment of the template to the inspected radiograph and grading score is calculated, automatically. Apart of tedious manual intervention, the initial system may deviates from the actual bone maturity score if the inspected radiographs has more than one most similar template to be matched. For this semi-automatic CASAS, there is no intermediate bone maturity score for two most similar template. In addition, this CASAS demands high definition quality of the radiograph image where the hand is properly positioned during acquisition process. Due to the current limitation of CASAS, various improvement efforts are continue to be proposed and explored with the intention to build a reliable CASAS for bone age assessment purpose.

Other than CASAS, exists other bone age assessment methods that utilizing different image modalities. The image modalities include ultrasonography and computerized magnetic resonance imaging. In the article entitled “Bone age: assessment methods and clinical applications”, Mari Satoh summarizes the details of other bone age assessment methods (Satoh, 2015).

For ultrasonography method, there is an instrument known as “BoneAge®” from the Sunlight Medical Ltd, Tel Aviv, Israel to calculate bone age. In this method, the hand is placed at the arm rest between two transducers in order to allow ultrasonic waves to permits through the ulnar epiphysis and distal radius of the hand bone. With this, the bone age is calculated by adopting algorithm in analyzing the sound velocity and distance between the transducers.

In article by Mari Satoh, there is a study between BoneAge® and standard methods includes GP and TW (Mentzel, Vilser et al., 2005). The study indicates there is a high correlations exist between the BoneAge® and standard methods. Besides, report by (Khan, Miller et al., 2009) stated that the accuracy of the BoneAge® in measuring bone age is inaccurate for samples with delayed or advanced bone age. BoneAge® has the tendency to compute higher age value for delayed bone age, meanwhile lower age value for advanced bone age. This report indicates that the BoneAge® may not be the most appropriate and reliable technique in measuring bone age. However, continuous research efforts are conducted using ultrasonography method are seen in the recent years.

Aside from CASAS, other computerized methods in assessing bone age exists (Rucci, Coppini et al., 1995). Although CASAS had the automation procedures in grading the bone age, the time taken to calculating the age may be longer than manual grading by GP or TW methods (Przelaskowski, 2008). Therefore, other computerized method to automate the grading process of skeletal age does exists. For example, (Sato, Ashizawa et al., 1999) introduces a new approach for this purpose and the system is tested on Japanese children. This computerized method is known as Computer-Aided Skeletal Maturity System (CASMAS). Grading process in CASMAS does not rely on TW method. In CASMAS, the third phalanges is segmented from the digital radiograph, automatically. Pertinent features are extracted from the segmented area of third phalanges. The features are epiphyses's width, metaphases and the overlapping areas of the epiphysis or metaphysis of the third phalanges. The final skeletal age is computed by using multiple regression analysis (Tanner, Oshman et al., 1994).

In 2009, a software known as BoneXpert are developed by a group of researcher from Denmark (Thodberg, 2009, Thodberg, Kreiborg et al., 2009, Thodberg and Sävendahl, 2010). The software is further developed and maintained by Visiana. BoneXpert

performed enhancement on the border of 13 RUS bones by reconstructing the bone borders. The enhanced border of RUS bones are used as an input to intrinsic parameter extraction process, where the parameters include shape, density and texture scores. These parameters are used to compute the bone age. The BoneXpert has superior enhancement technique that has the ability to filter the poor quality images or images with abnormal bone morphology. 1559 radiographic images from 2 years old to 17 years old are used to develop BoneXpert and these images are validated by using GP atlas. As compared to GP atlas, the standard deviation of the result reported in BoneXpert is 0.42 years with 95% confidence level. BoneXpert also been tested with TW method. The test is conducted with 84 radiographs, which are rated with TW method, initially. With TW method, BoneXpert gives standard deviation of 0.80 years with 95% confidence level. The efficiency of BoneXpert is further validated using data from Caucasian children, which the children suffers from short statures (Thodberg and Sävendahl, 2010) and precocious puberty (Thomas, Flores-Tapia et al., 2011). This system is further tested with variety of ethnic categories (Thodberg, Böttcher et al., 2016, Thodberg and Martin, 2016, Thodberg, Sørensen et al., 2016, Thodberg, Sørensen et al., 2016, Thodberg, van Rijn et al., 2017).

Researchers claim five contributions from BoneXpert as a computerized skeletal age assessment system. The first contribution is devoted to the novelty of bone reconstruction by using active appearance model. Second contribution is due to the new extracted intrinsic features in predicting bone age with principal component analysis. Thirdly, the bone age for each bone is correlated together in order to determines the final age value. The fourth contribution is the introduction of one general model for bone age calculation regardless the gender of the perceived skeletal. The last contribution is the unified assessment model for both GP and TW models. The following section is devoted to the background review for contrast enhancement.

2.3 Contrast Enhancement

Image enhancement is one of the most crucial and active topic in the digital image processing field. Image enhancement is categorized to two categories where either the processing process is taken place in spatial domain or frequency domain. Enhancement in spatial domains means pixels in the two dimensional space in the image are manipulated. Meanwhile, frequency domain processing manipulates the intensity values in the Fourier Transform space in the image (Thung & Raveendran, 2009).

From the past decades, researchers are searching the exact definition of contrast in image as different researcher has its own description and understanding on this research topic. It is common understanding that contrast of an image is considered good if the intensity values in the image spans in the full intensity range, which is from zero to 255. However, image contrast is known as visibility as it changed by scale while maintaining the average luminance to be constant in Michelson definition (Peli, 1990). Nevertheless, Michelson definition on contrast may not be appropriate and it is sensitive to extreme change of intensity value in the image. For example, a small number of pixels with either extreme brightness or darkness is able to change the total contrast of an image. The perceived contrast in image may be low although the contrast value in Michelson definition is high with some extreme intensity value changes happen in the image.

Contrast enhancement has become a necessity as a preliminary pre-processing process for medical images prior to the following analysis processes in the advancement of medical imaging field (Gonzalez and Woods, 2002, Gonzalez and Woods, 2007). For contrast enhancement in medical imaging, histogram equalization (HE) is the most established contrast enhancement method to be applied (Qing and Ward, 2007). Histogram equalization is an image manipulation technique that remap the intensity value to a new intensity value for each pixel in the image according to a specific function. This

specific function is commonly known as cumulative distribution function. The cumulative distribution function represents the statistical information of the intensity distribution of an image. The purpose of intensity remapping is to ensure the intensity range of the resultant image covers the entire intensity range. With the histogram equalization, the intensity difference between neighboring pixels is higher and this will makes the object of interest in the image is obviously highlighted. This type of visual enhancement is known as contrast enhancement.

Histogram equalization demands a low computation resources due to its simplicity in its algorithm structure. However, the resultant image from the histogram equalization comes with visual artifacts due to its simplicity. The visual artifacts include washed out effects. This effect is due to the incorrect brightness shift, detail loss and over-enhancement of contrast. The following paragraphs illustrates the computation of histogram equalization in detail descriptions (Gonzalez and Woods, 2002).

Histogram equalization determines the new intensity value for the resultant image based on a single-indexed mapping function, $T[n]$. This mapping function is computed from the new generated histogram (Arıcı, Dikbas et al., 2009). The new generated histogram represents the uniform distribution of the resultant image, which captures the optimum dynamic range of intensity. Let $h(r)$ denotes the histogram prior to the enhancement of the input image, which describes the pixel intensity value in terms of discrete probability mass function. Let $p(r)$ denotes the normalized histogram of $h(r)$ that is normalized by the total number of pixels in the image. In this context, $p(r)$ is the approximation of probability density function (PDF) of the intensity distribution in the image. With the probability density function, the approximated cumulative distribution function (CDF) is computed and denoted as $c(r)$. The final mapping function is a scaled CDF denoted by $c(n)$.

In the continuous form, the scaled $c(n)$ is denoted as $T(n)$ in continuous form when $r = n_x$ is illustrated in Equation (2.1).

$$T(n_x) = (L - 1) \int_0^{n_x} p_X(r) dr \quad (2.1)$$

In the discrete form, the scaled $c(n)$ is denoted as $T(n)$ where $r = n_x$ is illustrated in Equation (2.2).

$$T(n_x) = \text{floor}((L - 1) \sum_{n=0}^{n_x} p_X(r)) \quad (2.2)$$

The mapping transformation are derived, respectively by representing the intensities of input image and transformed image as continuous random variables, X and Y on $[0, L - 1]$ with n_y defined by Equation () and Equation (2.4) where $n_y = T(n_x)$, respectively:

$$n_y = T(n_x) = (L - 1) \int_0^{n_x} p_X(r) dr \quad (2.3)$$

$$n_y = T(n_x) = \text{floor}((L - 1) \sum_{n=0}^{n_x} p_X(r)) \quad (2.4)$$

Mathematically, n_y that is defined by $T(n_x)$ is uniformly distributed ($n_y = \frac{1}{L-1}$) on $[0, L-1]$ range of the input image PDF, independently as described in Equation (2.5):

$$\begin{aligned} \int_0^{n_y} p_Y(s) ds &= p(0 \leq Y \leq n_y) \\ &= p(0 \leq X \leq T^{-1}(n_y)) \\ &= \int_0^{T^{-1}(n_y)} p_X(r) dr \end{aligned} \quad (2.5)$$

From Equation (2.5), it is defined as $\int_0^{n_y} p_Y(s) ds = \int_0^{T^{-1}(n_y)} p_X(r) dr$

For left-hand side (LHS) of Equation (2.5), $\int_0^{n_y} p_Y(s) ds$, derivative of integral is applied to $\int_0^{n_y} p_Y(s) ds$ with respect to n_y as defined in Equation (2.6).

$$\frac{d(\int_0^{n_y} p_Y(s) ds)}{dn_y} = p_Y(n_y) \quad (2.6)$$

For right-hand side (RHS) of Equation (), derivative of integral is applied to $\int_0^{T^{-1}(r)} p_X(r) dr$ with respect to r . Equation (2.7) is derived based on the theorem of chain rule of derivatives.

$$\frac{d(\int_0^{T^{-1}(y)} p_X(r) dr)}{dr} = \frac{dT^{-1}(n_y)}{dr} \cdot \frac{d(\int_0^{T^{-1}(n_y)} p_X(r) dr)}{dT^{-1}(n_y)} \quad (2.7)$$

Let $F(r) = \int_0^{T^{-1}(n_y)} p_X(r) dr$, derivative with respect to n_y is applied to get the following equation.

$$F'(r) = p_X(r)$$

$$\int_0^{T^{-1}(n_y)} p_X(r) dr = F[T^{-1}(n_y)] - F[0] = F[T^{-1}(n_y)]$$

$$\frac{d(\int_0^{T^{-1}(y)} p_X(r) dr)}{dn_y} = \frac{dT^{-1}(n_y)}{dn_y} \cdot \frac{d(F[T^{-1}(n_y)])}{dT^{-1}(n_y)}$$

If $F'(r) = p_X(r)$, Equation (2.8) is defined as follows:

$$\frac{d(F[T^{-1}(n_y)])}{dT^{-1}(n_y)} = p_X(T^{-1}(n_y))$$

$$\frac{d(\int_0^{T^{-1}(y)} p_X(r) dr)}{dn_y} = \frac{dT^{-1}(n_y)}{dn_y} \cdot p_X(T^{-1}(n_y)) \quad (2.8)$$

By integrating LHS and RHS results, Equation (2.9) is achieved.

$$p_Y(n_y) = \frac{dT^{-1}(n_y)}{dn_y} \cdot p_X(T^{-1}(n_y)) \quad (2.9)$$

From Equation (2.3), the following equations are derived:

$$n_y = T(n_x) = (L - 1) \int_0^{n_x} p_X(r) dr$$

$$n_x = T^{-1}(n_y)$$

$$T(T^{-1}(n_y)) = (L - 1) \int_0^{T^{-1}(n_y)} p_X(r) dr$$

LHS and RHS on the above equation is differentiated with respect with n_y , the following equation is obtained:

$$LHS = \frac{d(T(T^{-1}(n_y)))}{dn_y} = \frac{d(n_y)}{dn_y} = 1$$

$$RHS = (L - 1) \frac{d \int_0^{T^{-1}(n_y)} p_X(r) dr}{dn_y}$$

The above RHS equation is updated with Equation (2.8) in order to obtain the following equation:

$$RHS = (L - 1) \frac{dT^{-1}(n_y)}{dn_y} \cdot p_X(T^{-1}(n_y))$$

By matching the above RHS equation with LHS equation, the following equation is achieved:

$$1 = (L - 1) \frac{dT^{-1}(n_y)}{dn_y} \cdot p_X(T^{-1}(n_y))$$

From Equation (2.9), the above equation is updated as follows:

$$1 = (L - 1)p_Y(n_y)$$

Therefore, the mapping function in the histogram equalization is uniformly distributed as illustrated in Equation (2.10).

$$p_Y(n_y) = \frac{1}{L - 1} \tag{2.10}$$

In the research study, there is another variation of histogram equalization. This variation is known as histogram specification. In histogram specification, the input image is transformed to the resultant image by matching the histogram of the resultant image with the specified histogram (Alemán-Flores, Álvarez et al., 2007, Thomas, Flores-Tapia et al., 2010, Thomas, Flores-Tapia et al., 2011). Histogram specification consists three standard steps in obtaining enhanced image. The first step involves equalization process onto the histogram of the input image. The second step equalizes the specified histogram. In the third step, the two equalized histograms are matched in order to generate a lookup table for mapping purpose. By performing intensity mapping from the generated lookup table, the enhanced image is computed. Although the histogram specification allows adjustment on the shape of the histogram for different type of application, this approach suffers from major limitation that it needs user intervention in determining the specified

histogram. Therefore, HE is still considered as the most convenient approach to be used in order to generate enhanced image as no parameters needed to be set. In this thesis, histogram equalization is chosen to be the focus as the contrast enhancement approach.

2.3.1 Global Histogram Equalization

In general, histogram equalization is classified into two major approaches. The approaches are Global Histogram Equalization (GHE) (Bull, Edwards et al., 1999) and Local Histogram Equalization (LHE). This section describes the GHE and the latter will be elaborated in the following section. In GHE, all the pixels in the images are considered in generating the cumulative distribution function, which is the remapping function.

Although GHE demands a very low computation resources, research works shown that the GHE is not able to preserve the visual perception of the enhanced image. Therefore, this over-simplified approach is not widely used in major application area, especially in medical field (Chen Hee and Isa, 2010, Chen Hee and Isa, 2010). The resultant image that is generated from GHE contains unwanted artifacts that include incorrect mean brightness shift, loss of detail and over contrast enhancement. Either partial or full unwanted artifacts are commonly seen in the resultant image by GHE. With these artifacts, the resultant image does not contain the similar visual properties as the original input image. The negative effect of this enhancement will be more critical if prominent features are washed out in the resultant image. The appearance of unwanted artifacts is due to the nature of the GHE in transforming the image, where the equalized histogram is not balance as there is a possibility that the high intensity level dominates the low intensity level in the histogram. In addition, the flattening effect of GHE on the histogram of the original image may trigger incorrect mean brightness shift. This artifact carries unnatural visual quality on the enhanced image (Chen and Rahman Ramli, 2004).

As the research works on histogram equalization is keep progressing, researchers enhance the histogram equalization with reference to the three specific limitation of this approach that includes incorrect mean brightness shift, loss of detail and over or under contrast enhancement. Many variation type of histogram equalization are recommended in the literature review (Sang-Yeon, Dongil et al., 1999). In general, there is three characteristic type of the variation of histogram equalization. The first type of histogram equalization considers plateau limit in the histogram equalization approach. This approach introduces a specific limit of the total number of a pixel intensity value in the mapping operation (Kim and Min, 2008, Kim and Paik, 2008, Sengee and Heung, 2008, Sengee, Bazarragchaa et al., 2009, Se-Hwan, Jin Heon et al., 2010, Sengee, Sengee et al., 2010, Se-Hwan, Jin Heon et al., 2011). The first enhancement type is known as clipped histogram equalization approach. The second type of histogram equalization divides the histogram of the input image into few sub-histograms. Each sub-histogram undergoes a standard equalization histogram process and all the equalized sub-histogram is combined as a single resultant histogram (Yeong-Taeg, 1997, Yeong-Taeg, 1997, Yeong-Taeg, 1997, Soong-Der and Ramli, 2003). The second type of enhancement is referred as sub-histogram equalization approach. The third type of histogram equalization is the output of the mixture of different variation of histogram equalization approaches such as plateau limit histogram equalization approach (Chen Hee, Kong et al., 2009, Kong, Ibrahim et al., 2009, Chen Hee and Isa, 2010, Chen Hee and Isa, 2010), sub-histogram equalization approach (Kim and Min, 2008), weighted histogram equalization approach (Se-Hwan, Jin Heon et al., 2010) and dynamic range reallocation approach (Abdullah-Al-Wadud, Kabir et al., 2007, Abdullah-Al-Wadud, Kabir et al., 2008). This thesis is devoted to the mixture type by combining first and second type of histogram equalization variation, which is the bi-histogram equalization approaches and clipped histogram equalization.

The initial research works on sub-histogram equalization approach is known as Bi-Histogram Equalization (BBHE) (Yeong-Taeg, 1997). In this research work, this particular equalization approach determines the mean value of the intensity distribution of the image, statistically. This mean value is used as a cutting point value to separate the histogram of the original image into two sub-histograms. Then, histogram equalization is implemented into each of the sub-histogram. The intention of histogram equalization implementation on the individual separated histogram is to minimize the visual artifact that is caused by incorrect mean brightness shift. The minimization is possible due to the intensity distribution is separated into two groups, where each group contains the distribution of lower intensity and the other group remains with the higher intensity. Another similar sub-histogram equalization approach to BBHE is introduced as Dualistic Sub-Image Histogram Equalization (DSIHE) (Yu, Qian et al., 1999). Both approaches shared high similarity in terms of histogram equalization implementation on separated sub-histogram. The only difference occurred in the value used as the cutting point. In this context, DSIHE utilizes the median value as compared to mean value in BBHE as a cutting point. DSIHE claimed to be better than BBHE in terms of addressing the effects of incorrect mean brightness shift of the resultant image. Chen and Ramli proposes an improvement version of BBHE and this improvement approach is termed as Minimum Mean Brightness Error Bi-Histogram Equalization (MMBEBHE) (Soong-Der and Ramli, 2003, Soong-Der and Ramli, 2003). The effectiveness of BBHE approach is determined by the cutting point value that separates histogram into two sub-histograms. The cutting point value is optimum if the error value of the mean brightness shift of the resultant image from BBHE approach is minimum. In MMBEBHE, the optimum separating point value is determined with the error value of the mean brightness shift in the resultant image is minimum. The mean brightness shift is measured by using an image quality metric known as absolute mean brightness error (AMBE).

The extension works on sub-histogram equalization involves more than one cutting points of histogram from the input image into multiple sub-images. The variation also depends on the cutting point value whether it is computed based on the mean or median of the statistical intensity information of the input image. Chen and Ramli (Soong-Der and Ramli, 2003) proposed to divide the histogram into multiple histograms by using mean as the cutting point, continuously. The stated approach is known as Recursive Mean-Separate Histogram Equalization (RMSHE). It can be seen that RMSHE is the generalized and recursive solution for the BBHE where the number of sub-histogram is in power of two. The valid range of the number of sub-histogram is from one to less than the maximum number of intensity level in the image. The separation of histogram of input image recursively based on mean value is known as RMSHE. On the other hand, the approach is known as Recursive Sub-Image Histogram Equalization (RSIHE) (Sim, Tso et al., 2007) if the separation of histogram of input image recursively based on median value. It is claimed that the recursive solution for BBHE is not only addressing the mean brightness shift issue, but also the outcome of this variation of BBHE gives a more natural resultant image.

The development of recursive histogram equalization in bi-histogram is continue to progress, where Kim and Min (Kim and Min, 2008) introduced a weighted version of recursive histogram equalization, which this version is termed as the Recursively Separated and Weighted Histogram Equalization (RSWHE). In RSWHE approach, different weight value is assigned to each separated sub-histogram. The weight is assigned with a normalized power law function. The final step involves implementation of histogram equalization into each weighted sub-histogram, accordingly. The variation of RWSHE approach is based on the type of cutting point value used. If the cutting point value is computed using mean, the weighted recursive HE is known as RSWHE-M. Otherwise, it is refers as RSWHE-D if the cutting point value is derived from median.

The result comparison shown that the RSWHE-M is superior to RSWHE-D in having the lowest error value of the mean brightness shift.

With relation to the brightness preservation, another histogram equalization technique is explored towards this direction and this technique is known as the Brightness-Preserving Dynamic Histogram Equalization (BPDHE) (Ibrahim and Kong, 2007). BPDHE is an extension works to two previous sub-histogram equalization variation techniques, which is the variation approach to Multi-Peak Histogram Equalization With Brightness Preserving (MPHEBP) (Wongsritong, Kittayaruasiriwat et al., 1998) and Dynamic Histogram Equalization (DHE) (Abdullah-Al-Wadud, Kabir et al., 2007, Abdullah-Al-Wadud, Kabir et al., 2008). The main difference that distinguish between MPHEBP and DHE approaches is the technique used to separates the histogram into sub-histograms. MPHEBP performs the separation base on the identified high peaks in the histogram. This high peaks are known as local maxima. On the other hand, DHE separates the histogram base on the local minima points found in the histogram. In BPDHE, this approach determines the sub-histograms using local maxima points as similar to MPHEBP approach. In order to preserve the mean brightness of the resultant image closely similar to the input image, BPDHE maps each sub-histogram into a new dynamic range prior to the standard equalization technique. This initial mapping procedure is resembles with the DHE approach.

Another approach in preserving mean brightness of the input image is proposed by (Abdullah-Al-Wadud, Kabir et al., 2007). This approach is refers to A Dynamic Quadrants Histogram Equalization Plateau Limit (DQHEPL). DQHEPL separates the histogram of the input image into two sub-histogram based on mean intensity value. The two sub-histogram is continue to be separated into another four sub-histograms. The number of separation of sub-histogram can be continued with power of 2. During the

continuous separation, the plateau limit is calculated. With the calculated plateau limit, each of the separated sub-histograms are clipped. Prior to the equalization process, a new dynamic range is assigned to each sub-histogram. The resultant image based on DQHEPL approach is promising as it able to maintain the mean brightness of the input image. Another histogram equalization approach that devoted to mean brightness preservation is proposed by (Thomas, Flores-Tapia et al., 2011). This approach combines two approaches, which is Brightness-Preserving Histogram Equalization with the Maximum entropy (BPHEME) (Chao and Zhongfu, 2005) and Piecewise linear transformation (PLT) (Abdullah-Al-Wadud, Kabir et al., 2008). BPHEME includes maximum entropy concept into the mean brightness preservation technique. In BPHEME, a histogram with maximum entropy is specified. BPHEME is a variation approach from histogram specification that incorporates maximum entropy knowledge into the histogram. On the other hand, PLT is useful in ensuring small difference value between the mean brightness of the input image and resultant image. With combination of these two approaches, the resultant image has high potential in preserving the mean brightness of the input image, precisely.

The above described histogram equalization approaches are mainly focus in minimizing visual artifacts caused by mean brightness shift. It is also exist other histogram equalization that consider minimization of other visual artifacts too. Weighted Clustering Histogram Equalization (WCHE) (Sengee and Heung, 2008) not only consider the mean brightness shift but also the artifacts caused by over-enhancement. WCHE incorporates clustering concept in its technique in determining the sub-histogram. WCHE groups each pixel in the input image into different cluster group. As the number of cluster group is very high, these clusters are re-grouped together into more prominent group. The grouping criteria is based on the grouping weight for each cluster, distance between the vicinity clusters and ratio information of the cluster. The resultant image from this

approach is obtained by transforming the cluster information into image form. Although WCHE is able to preserve the overall brightness and contrast of the original image, this approach still suffers from loss of detail. The pertinent information in the image is lost due to the equalization process on the clusters that contains high number of intensity bins into narrow intensity range.

It is a non-trivial task in the visual enhancement study with the objective in preserving the details of the pertinent information in the image. This task is even difficult if the pertinent information only appears in the small areas in the image. With regards to this challenge, Bi-Histogram with Plateau Limit (BHEPL) is introduced by (Chen Hee, Kong et al., 2009). In addition, BHEPL is known as a fast sub-histogram equalization approach by reducing the computational burden in each computed sub-histograms. BHEPL separates the histogram of the input image into two sub-histogram by using mean intensity value in the image. For each separated sub-histogram, a “clipping” process is carried out. The purpose of clipping is to limit the total number of pixels at each bins in the sub-histogram. After the clipping process is completed, general histogram equalization process is implemented.

From the intensive reviews of GHE as described above, the objective of all the histogram equalization approaches is focus in minimizing issue of visual artifacts caused by mean brightness shift, loss of details and over-enhancement. Majority of the proposed histogram equalization technique mainly focuses on addressing only single issue, while neglecting the others two issues. For example, standard histogram equalization enhances the contrast quality of the image but neglecting the mean brightness shift and detail loss issues. The well-known sub-histogram equalization approach intended to focus into issue of mean brightness shift but neglecting the issues of detail loss and contrast enhancement. Meanwhile, histogram equalization that utilizes plateau limit concept only focus on detail

loss issue but not consider the issues of brightness mean shift and contrast enhancement. To address the limitation of the common histogram equalization approaches, (Chen Hee, Kong et al., 2009) and (Chen Hee and Isa, 2010) attempt to tackle both the mean brightness shift and detail loss issues. Although the attempts shown improvement on reducing the artifacts, but the minimization of the three visual artifact issues simultaneously is still remain as a high challenging problem.

Therefore, there is a need to establish a holistic histogram equalization algorithm that able to consider the three visual artifact issues, simultaneously. It is a predicted behaviour that an introduced solution to improve one issue, will affect the improvement criteria on the other issue. This behaviour is known as trade-offs. Hence, to address this challenges, a novel model of histogram equalization is proposed. This proposed model is an extension variation to clipped histogram equalization and bi-histogram equalization approaches by utilizing optimization technique in multiple criteria. In this model, the histogram is separated using an optimal separating point value and clipped limit is introduced to each sub-histogram.

2.3.2 Local Histogram Equalization (LHE)

LHE is also known as Adaptive Histogram Equalization (AHE). LHE performs equalization process from a group of pixels that are located in the small region in the input image. Equalization process is performed through windowing technique where the center of this window will move from one pixels to another adjacent pixels for processing purpose. The advantage of LHE is its ability to preserve the details of the specific area in the input image as the enhancement is done locally and not on the whole image as compared to GHE approach. Due to the nature of local processing mechanism, the

resultant image is less suffering from over enhancement effects although the input image had non-uniform intensity distribution.

LHE is also a contrast enhancement technique that use remapping technique base on a transformation function as similar to GHE approach (Chen Hee, 2010). As compared to GHE, LHE applies histogram equalization process at different area of the input image, continuously. In GHE approach, there is only one transformation function throughout the equalization approach. However, LHE computes more than one transformation function, where each transformation function is computed at different processing area in the input image. LHE is an adaptive enhancement technique due to each computed transformation represents the intensity distribution of the particular area of the image only. Therefore, the resultant image is generated base on different unique transformation function at each different region in the input image. If the image contains non-uniform intensity distribution, LHE is preferred as compared to GHE in enhancing contrast of the image, whereby artifact likes over-enhance or under-enhance is suppressed. With this inherent strength, LHE is regarded as robust approach to be used to improve the contrast of the image, specifically for non-uniform intensity image.

Although the strength of LHE is able to perform better enhancement on non-uniform intensity image, this approach has its disadvantages due to its local processing methodology. One of the major disadvantages of LHE is this approach tends to amplify existing noises that appear in the image, specifically if the noise is located at the region of the image that contains uniform intensity. Due to this disadvantage, Contrast Limited Adaptive Histogram Equalization (CLAHE) is the most popular LHE approach. The popularity is due to CLAHE is able to minimize the noise amplification effect. This is done by introducing a clipping method at the maximum height of each local processing step (Pizer, Amburn et al., 1987, Kim and Paik, 2008, Chen Hee, Kong et al., 2009, Chen

Hee and Isa, 2010). Despite CLAHE able to reduce the effect of noise amplification, the resultant image still contains noises at the region where the intensity distribution is homogeneous (Josephus and Remya, 2011). This limitation had motivated researchers to continue to improve the method used in CLAHE and there is also exist of approach that combines CLAHE with other variation of HE for enhancement improvement purpose. (Mohan and Ravishankar, 2013).

As CLAHE is still suffering from over contrast and unwanted noise amplification, the author (Lidong, Wei et al., 2015) addresses a method known as CLAHE-discrete wavelet transform (CLAHEDWT). This method utilizes the advantage of DWT into CLAHE. CLAHEDWT comprises of three major computation steps. The first step involves the use of discrete wavelet transform to divides the input image into low and high frequency components. Secondly, CLAHE is used to improve the low frequency coefficients and the high frequency coefficients remain unchanged. By remaining the high frequency components unchanged, the important details in the image is preserved and to mitigate the presence of unwanted noises. The improved low coefficients are inversed by discrete wavelet transform (DWT) and at the same time the inverse results are averaged with a weighting factor. The final resultant image is obtained with less noise and highlighting the pertinent information in the image. The weighted averaging procedures enhances the dark region more than the brightness regions. Therefore, CLAHEDWT enhances the low frequency component while maintaining the value of the high frequency components.

There is another variation of LHE that is applied to MMBEBHE. This variation is known as the Local MMBEBHE or LMMBEBHE (Tian, Wan et al., 2007). For this algorithm, the MMBEBHE procedure is applied to multiple regions in the image. However, there is a need to decide the number of blocks and step size of the movement blocks. After all the blocks are processed by MMBEBHE, the output of each of the block

will be divided by the frequency matrix of the image histogram. Although this approach is exhibiting local information, the efficiency of this method is still low as compared to the origin method, MMBEBHE.

Arici introduced a new framework for histogram equalization. This framework is referred as Histogram Modification Framework (HMF) (Arici, 2009). HMF enhances the contrast of the image by optimizing a contrast model which is mathematically modelled with a cost function. The author emphasizes that this framework is flexible to suit various application where the cost function is able to be manipulated with different constraint terms.

It is noted that majority of the LHE approaches are mainly addresses the issue of noise amplification in the resultant image. The noise amplification issue is due to the nature of intensity remapping technique. In LHE, the outlier intensity either is too high or too low in the homogeneous region will change the overall pattern of the cumulative distribution function. This type of cumulative distribution function resulting the new intensity has a major difference from the original intensity. The major difference of intensity is referred as noise in the resultant image. Due to this limitation, LHE is not commonly used in practical application in enhancing the quality of the image as it will change the appearance of the pertinent information in the original image.

2.4 State-of-the-art Contrast Enhancement Algorithms

The mentioned contrast enhancement methods discussed under GHE approach are presented in Table 2.5, chronologically. In Table 2.5, GHE approach started with the general procedure of flattening the histogram of the image into full intensity range. Therefore, the standard GHE is a single equalized histogram method with the aim of preserving the overall contrast of the image. The limitation of this standard approach is consideration of brightness shift and detail loss are neglected. Throughout the years, GHE approach is improved from single histogram to two histograms where the original histogram is divided into two parts via a separating point. This separating point can be the mean, median, peak, local minima or value in the valid intensity range that has minimum error of the brightness shift. This GHE approach is categorized as bi-histogram via a separating point, where the aim of this type of approach is mainly preserving the brightness shift and contrast. However, the detail loss is not been considered in this improved GHE approach. From bi-histogram, recursive histogram is introduced where multiple histograms are recursively generated using valid separating points. Other variation of recursive histogram includes weightage recursive histogram and clustering method. Although multiple histogram is generated with the intention of preserving the brightness shift and contrast, the prominent details are still not been considered. The advancement of GHE approach is extended to applying clipped limit to recursive histogram. With the introduction of clipping procedure in the recursive histogram, the detail loss and brightness shift are considered, but the contrast is not enhanced optimally. In this thesis, the proposed contrast enhancement framework does not extend from the mentioned approaches in Table 2.5, but the method of using separating points to divide histogram into multiple histogram and applying clipped limit to each divided histogram are applied. In this proposed contrast enhancement framework, the separating point and clipped limit are determined, optimally.

Table 2.5: A comparison of GHE approaches in contrast enhancement

Approach	Method	Aim	Limitation
Standard GHE	Single equalized histogram	Preserve overall contrast	Not consider brightness shift and detail loss
BBHE	Bi-histogram via mean value	Preserve brightness shift	Not consider detail loss
MPHEBP	Bi-histogram via peak value	Preserve brightness shift and contrast (not optimal)	Not consider detail loss
DSIHE	Bi-histogram via median value	Preserve brightness shift	Not consider detail loss
MMBEBHE	Bi-histogram via minimum brightness error	Preserve brightness shift and contrast (not optimal)	Not consider detail loss
RMSHE	Recursive histogram via mean value	Preserve brightness shift and contrast (not optimal)	Not consider detail loss
DQHEPL	Recursive histogram with clipping	Preserve brightness shift	Not consider detail loss
RSIHE	Recursive histogram via median	Preserve brightness shift and contrast (not optimal)	Not consider detail loss
RWSHE	Weighted recursive histogram	Preserve brightness shift and contrast (not optimal)	Not consider detail loss
DHE	Bi-histogram via local minima	Preserve brightness shift and contrast (not optimal)	Not consider detail loss
WCHE	Recursive histogram via clustering	Preserve brightness shift and contrast (not optimal)	Not consider detail loss
BHEPL	Recursive histogram with clipping	Preserve detail and brightness shift(not optimal)	Contrast is not optimal

Other than HE approaches, new direction of contrast enhancement approach is introduced, lately. Although this new approach delivers promising enhancement result, this method is more suitable to natural colour image rather than medical image, specifically hand bone image. The following section is devoted to this new direction of contrast enhancement.

In the paper of “An Effective Gaussian Fitting Approach for Image Contrast Enhancement” (Sun, Xu et al., 2019), the authors utilized and adopted 24 Kodak test images and retinex database in order to assess the performance of the proposed algorithm with other contrast enhancement techniques (Sazonova and Schuckers, 2010, Tsai, 2013). The database is retrieved from the following links: (www.cs.tut.fi/~lasip/cfai) and (<https://dragon.larc.nasa.gov/retinex/>), respectively. The objective of this method is to enhance image details at specific region-of-interest area in the image. This enhancement step exploited the high frequency distribution of an image by estimating the intensity weighting matrix. With the estimated intensity weighting matrix, the contrast gain distribution is obtained by manipulating the Gaussian fitting curve on the matrix. This method can be summarized with three major steps that includes the extraction process of region-of-interest area, generation of intensity weighting matrix and finally, the Gaussian fitting transformation. The objective assessment is performed with three metrics although there is a lack of an acceptable objective criteria towards obtaining good enhancement results for all images. The three used metrics are peak signal-to-noise (PSNR) (Horé and Ziou, 2010), structural similarity (SSIM) (Zhou, Bovik et al., 2004) and entropy. PSNR is a measurement metric to compare the ratio of signal to background noises. Main focus of SSIM is to measure the preservation of original content in the image. On the other hand, entropy measures the richness of the preserved information after enhancement.

In the paper “Contrast Enhancement Based on Intrinsic Image Decomposition” (Yue, Yang et al., 2017), the authors proposed a framework to improve contrast enhancement of the input image by utilizing retinex theory (Land and McCann, 1971). Retinex-based contrast enhancement approach decomposes the original image into two layers, which is the illumination and reflectance layers, respectively. According to the human visual system, it is assumed that the visual image is resembling the product of illumination and reflectance layers. In this context, the illumination is the reflected brightness from the

surface objects in the scene and the reflectance is the measure of the proportion brightness that is reflected from the surface objects in the scene (Wang and Luo, 2018). In human visual system, the colour sensation is correlated to the reflectance (Land and McCann 1971). In this paper, the decomposition approach can be divided into three types, which is Poisson equation based, filtering based and variational based. The proposed contrast enhancement approach comprises of four major steps. The first step involves colour transformation where the RGB colour model of the input image is transformed to HSV colour model. The second step utilizes modified intrinsic decomposition technique to separate the HSV colour model into two different layers, which is the illumination and reflectance layers. Then, the illumination layer undergoes an intensity adjustment prior to the integration with reflectance layer. The authors introduced a weighted L1 norm constraint on the neighborhood pixels based on colour similarity. The purpose of this constraint is to reduce the effect of illumination on the reflectance layer. Finally, the last step is devoted to an enhancement process on the integrated image by applying CLAHE. The enhanced image in HSV colour format is retransformed back to RGB colour format as the final resultant image. The optimization technique used in this paper is Split Bregman algorithm (Goldstein and Osher, 2009). In this optimization process, three image quality measurement metrics are applied in order to evaluate the quality of the contrast enhancement. The three metrics are no-reference image quality metric for contrast distortion (NIQMC) (Gu, Lin et al., 2017), reference based colourfulness-based Patch-based Contrast Quality index (PCQI) known as (CPCQO) (Wang, Ma et al., 2015), and reference based edge visibility (EV) enhancement ratio (Gu, Tao et al., 2018).

In the paper “Naturalness Preserved Image Enhancement Using a Priori Multi-Layer Lightness Statistics” (Wang and Luo, 2018), the authors proposed a framework with an objective to solve the issue of over enhancement and effect of unnaturalness on the resultant image after contrast enhancement process. The introduced solution is utilizing

the priori multi-layer lightness statistics technique in which this statistic information is extracted from high quality image. There is three main contributions from this paper in the context of contrast enhancement approach. The first and second contribution of this paper is devoted to the generation of multi-layer image enhancement model and multi-layer lightness statistics of high quality outdoor images. The multi-layer lightness statistics will be combined to the multi-layer image enhancement model at the latter stage. The third contribution is demonstrated by the effectiveness of the proposed approach in ensuring consistent and high quality rating of the all the resultant images. In order to achieve the stated efficiency, two assumptions are made. The first assumption is highly important where the details of the image are across the spatial frequency bands. Based on the first assumption, the authors applied low-pass filter continuously towards the image in order to separate the image into multiple frequency bands. The low-pass filter iteration is terminated when the outcome of the lowest frequency band shows non-uniform characteristic. This type of filtering is different from the known multi-scale retinex (MSR) model, where the existing MSR estimates reflectance at several scales. The second assumption made is that the used high-quality images inherits standard statistical characteristic information in the lightness properties. The lightness properties are related to multiple scale layer and the statistical characteristics are different from the non-uniform illuminated images.

The authors in “Graph-Based Joint Dequantization and Contrast Enhancement of Poorly Lit JPEG Images” (Liu, Cheung et al., 2019) paper performs improvement on the images that suffers from poor lighting condition. This type of inferior image has unwanted artifacts in terms of luminance contrast and it exhibits quantization effects due to lossy compression. The authors claimed that dequantization and contrast enhancement technique using separate back-back steps will reduces the low visual quality of the image that is caused by undesirable compression effect. Based on the recent developed graph

signal processing, the authors adopted single graph-signal restoration framework by dequantizing and contrast-enhancing images. Based on Retinex theory, observed pixel patch is separated into reflectance and illuminance properties in which generalized smoothness prior and signed graph smoothness prior are obtained. The author computes the robust edge weights by using transform-coded image patch. Low pass filtering in the dual graph domain is used for the computation of robust edge weights. In this paper, Accelerated Proximal Gradient (APG) algorithms in transform domain is adopted in order to have illumination and reflectance components. One of the major limitations of this approach is the manual user intervention is needed to determine the positive and negative edge weights. If manual tune is needed, then the final resultant image may not be optimal.

2.5 State-of-the-art Segmentation Algorithms

Object segmentation is regarded as one of the important processes in the medical image processing and it is widely studied. Its contribution is able to augment the power of clinical analysis. One of the well-known and popular segmentation approaches is referred as active contour. One of the latest active contour algorithms by Wang (Qian, Wang et al., 2013) is a family member of learning active contour model that improve the performance of the current pixel wise segmentation method. The objective of this algorithm is to solve the presence limitations of active contour. The introduced method by Wang considers the area within and outside of the region of interest of the segmentation as well as the learning size of segmented boundaries. This consideration is done by having a design of a new loss function that include area and size information, where this information is further incorporated into a deep learning model. In this paper, the segmented results are analyzed and assessed using a comprehensive dataset that contains more than 2000 cardiac MRI scan images. From the analysis and assessment,

the loss function has a better and robust performance if this function incorporates hyper parameters in the design. The performance of this function is better as compared to other loss function such as cross-entropy on two common segmentation networks.

As artificial intelligence field is progressing, deep learning has been adopted to perform image segmentation in medical field. The segmentation applications include optic disc segmentation, lung segmentation and cell outline segmentation. This type of object segmentation using deep learning method is known as semantic segmentation. Semantic segmentation is a process of assigning each pixel in the image to a predefined set of classes. Deep learning becomes popular as Alexnet (Krizhevsky, 2012) is introduced as a convolutional neural network (CNN). This model of CNN shows a high accuracy result on image classification on ImageNet dataset. Due to its tremendous success in image classification field, various extensible techniques are designed and developed from CNN. These sophisticated deep learning models include R-CNN (Girshick, 2014), Faster R-CNN (Ren, 2016), VGG (Simonyan, 2014) and ResNet (He, 2016). All this model depicts up trending accuracy results. In term of image segmentation perspective in the medical field, it is always remain high demands to develop an automated segmentation technique that able to segment region of interest, precisely. Good segmentation result permits improvement in disease inspection and diagnosis, such as osteoarthritis assessment through segmented knee's cartilage (Prasoon, 2013, Faisal, 2015) and determination of prostate's volume from 3D prostate segmentation (Milletari, 2016). For image segmentation in medical imaging field, the recent two well-established and active explored neural network model are fully convolutional neural network (FCN) (Abadi, 2016) and Mask R-CNN (He, 2017).

The U-net (Falk, Mai et al., 2019) is one of the promising type of convolutional network architectures that is aiming to achieve rapid and accurate image segmentation

results. U-net successfully outperforms past segmentation results on the IEEE International Symposium on Biomedical Imaging (ISBI) challenge. This challenge relates to the task of segmenting the neuronal structures in electron microscopic stacks. In addition, U-net also won the grand challenge at ISBI 2015, which dedicated to the task of Computer-Automated Detection of Caries in Bitewing Radiography. Besides, U-net also shows its segmentation ability in the two most challenging transmitted light microscopy categories which are the phase contrast and DIC microscopy. U-net is developed from the convolutional neural network and shows tremendous result in the realm of pixel-based image segmentation. The early use case of U-net is first demonstrated on biomedical images. Each approach may has its own limitation and U-net is not excluded from this too. The limitation of U-net is further improved and inspired with others state-of-the-art including the CE-net: The U-net and its similar variations (Shaziya, Shyamala et al., 2018, Yang and Song, 2018, Zhang and Xu, 2018, Zhang, Chen et al., 2018, Waktola, Grudzien et al., 2019). The variations of U-net is a series of operations of pooling and convolution that mitigate the feature resolution in order to achieve higher level of feature representations.

The CE-net (Gu, Cheng et al., 2019), enhances the limitations of existing U-net based segmentation, which includes loss of spatial information from consecutive pooling and convolutional operations. The authors proposed a context encoder network or also known as CE-net, to exploit high-level information as well as preserve spatial information for 2D medical image segmentation. The CE-net comprises of three processing parts. The first part is the feature encoder module, the second part is the context extractor and the third part is the feature decoder module. The above mentioned context extractor module is established by the proposed design of dense atrous convolution (DAC) (Zhang and Xu, 2018) block and also the residual multi-kernel pooling (RMP) block. The authors experimented and analyzed the CE-net on variety of segmentation tasks that involves two

–dimensional medical images. The experiment result shows the CE-net performs better than the compared state-of-the-art methods. In this paper, the carried out experiments include optic disc segmentation, lung segmentation, cell outline segmentation and retinal optical coherence tomography layer segmentation.

Segmentation advancement is also contributed by the recent development of neural network based image classification such as Auto-DeepLab. This advancement is known as semantic segmentation. Prior to the detail discussion on Auto-DeepLab, NASNet search space (Zoph, Vasudevan et al., 2017) by the paper of “Learning Transferable Architectures for Scalable Image Recognition” is firstly discussed. NASNet search space is a method that has the ability to automatically learn the model architectures based on the targeted dataset. Due to the learning mechanism is based on the amount of dataset, the introduced method is highly computational expensive if the size of the dataset is large. This disadvantage cost motivates the authors to design a new searching method. In this method, the first search is devoted to smaller dataset in order to form an initial architectural building block. This block is transferable to form a new block to match a larger dataset. This type of search space is known as NASNet search space that permits transferability ability. The author carried out experiments on CIFAR-10 dataset in order to seek the optimum convolutional layer in which known as cell. In the following process, the cell layer is applied to the ImageNet dataset by stacking multiple cells together with its own individual parameter in order to form a convolutional architecture. The author addresses this convolutional architecture as NASNet architecture. Besides, the author also developed a new regularization technique known as ScheduleDropPath, which able to generalize into NASNet models. On CIFAR-10 dataset, NASNet is able to achieved 2.4% error rate which is already considered a low error rate as compared to others state-of-the-art methods. Besides, authors constructed a NASNet from the best cell, and recorded 1.2% better rate and also recorded 28% reduction in computational resource as compared to the

previous best human designed architectural model. In addition, the author also assessed the performance of NASNets at different computational cost and the accuracies of NASNets. NASNets also has been applied on object detection, the learned features by NASNet together with Faster-RCNN framework, the result surpasses state-of-the-art by 4%. Next we will discuss the application of NASNet in image semantic segmentation.

This work of Auto-DeepLab (Liu, Chen et al., 2019) is done together with John Hopkins Hospital, Google and Stanford University. Basically, this work utilized Neural Architecture Search (NAS) onto the work of semantic image segmentation. Semantic image segmentation associates each pixel in an image with a class label. NAS is chosen due to its recent report of NAS that documented high success on large-scale image classification. This work started by reviewing the existing works and discovered that existing works invariably put their interest on capturing repeatable cell and lack of automated outer network structure that manipulates the changes in spatial resolution. From the review, NAS exhibits difficulty in dense image prediction although it is able to simply searching space. It is known that dense image prediction contains multitude of variations in network level architecture. Therefore, the authors propose searching of network level structure other than the cell level structure. By doing this, a hierarchical architecture search space is formed. This presented network level search space comprises multitude state-of-the-art designs as well as formulates robust architecture search based on gradient. The success of this work has been proven by using several datasets that includes ADE20K datasets, Cityscapes, PASCAL VOL 2012. The authors named this search architecture as Auto-DeepLab. It is claimed that this method outperforms state-of-the-art methods even without pre-training using any ImageNet. The authors also claimed that this is the first attempt in extending the NAS other than application of image classification into dense image prediction problems. They stressed the significance of spatial resolution changes instead of fixating on cell level. In addition, they consider the

architectural variations by combining the network level with search space. A differentiable formulation has also been developed by the authors that permit effective architecture search for the two level hierarchical search space and the claimed speed is thousand times faster than DPC (Chen, Collins et al., 2018). The latest version of Auto-DeepLab is DeepLabv3+ (Chen et al., 2018). DeepLabv3+ extends DeepLabv3 by introducing an effective decoder module in order to enhance the segmentation results, targeting the outline of the objects. DeepLabv3+ combines Atrous Spatial Pyramid Pooling (ASPP) that is able to encode multi-scale contextual information with Encode-Decoder architecture, in which the location or spatial information is recovered. In this version, a faster and stronger network is developed by utilizing Modified Aligned Xception and Atrous Separable Convolution. The effectiveness of DeepLabv3+ is tested in PASCAL VOC 2012 and Cityscapes datasets. The performance is achieved at 89% and 82.1% without any post-processing on the two datasets, respectively.

2.6 Summary of Literature Review of Contrast Enhancement and Segmentation

From the review of the literature works in HE, the first observation is the GHE is a preferable approach in the HE field, as the nature of the GHE is simple and efficient in terms of computation resources. This makes the GHE is suitable to be used in real time applications. On the other hand, LHE is not favorable to be used in real time application as this approach demands a high computation time. Despite the high computation time of LHE, this approach is still an active research topic due to its advantage in preserving local information in the image. Thus, GHE and LHE have its own advantages and limitation. In this thesis, bi-histogram equalization and clipped histogram equalization approaches are chosen as an enhancement approach due to its high computation efficiency and at the

same time maintains the local information of the original image. The local information is maintained without performing the windowing technique in LHE approach.

The second observation is the majority of research works focus in enhancing the image by solving a single objective. For example, the single objective is devoted to brightness preservation and this is done by utilizing the mean brightness metric. Another example is endorsing entropy metric in preserving the details in the original image. In medical image application, an enhanced image with solving a single objective is not sufficient, and it demands an enhancement result that considers all factors. All factors are highly important in terms of enhancement perspective. A holistic enhancement approach is highly needed to produce a resultant image, where the contrast is enhanced, brightness is preserved and details is maintained. Therefore, in this thesis, a contrast enhancement framework based on histogram equalization that consider all these three factors is needed.

From the review of the literature works in segmentation, the early promising approach before the arrival of deep learning approach is to use active contour to obtain the outline of the objects in the image. One of the successful criteria in active contour approach is to initialize the primary object shape at the right location of the targeted object in the image. Therefore, the segmentation result of the targeted object is not accurate if the initialization of the primary object shape is not correct. The effectiveness of the active contour is also dependent to the complexity of the target objects in terms of shape, number of objects and the degree of overlapping of all the targeted object for segmentation. In active contour approach, the accuracy of the segmentation result is lower if the shape of the targeted object is complex, high number of objects in the image to be segmented and high overlapping degree between the targeted object. Due to this limitation of active contour, this approach is not suitable to be used in segmenting bone structures in the hand bone image. For example, carpals in hand bone image has high number of bone structures and

the shape of all the bone structure in the carpal is almost similar as well as they are overlapping to each other.

With the advancement of deep learning approach, active research based on this approach is showing significant improvement in segmenting objects in the medical images. Due to the dense architecture of the deep learning approach, the prominent features that is related to the targeted object for segmentation is automatically computed. Prior to the advancement in deep learning approach, the suitable and best features that represent the targeted object needs to be defined. From the past research works in finding the best feature, there is no single best representation features and each feature has its own trade off. In deep learning approach, the convolutional neural network able to extract mixture of good features, automatically. With this advantage, deep learning approach able to segments the bone structures in the hand bone image, successfully.

In general, the overall semantic segmentation approach is based on the encoder-decoder architecture. The encoder is expected to mitigate gradually the feature spatial dimension and learn more high-level semantic features. Meanwhile, the decoder acts to recover the objects spatial dimension. Although the higher level of feature representation is crucial for object classification and detection related tasks, this representation is an obstacle to dense prediction task in which this task demands detail spatial information. Despite the high level of feature representation is an important component in contributing to higher segmentation accuracy, the high volume of feature map will impede the training speed and increase the complexity of optimization. Due to this challenge, various semantic segmentation architecture is explored and introduced. The promising semantic segmentation approach includes some early method such as R-CNN (Regions with CNN feature) and FCN (Fully Convolutional Network), to more complicated architecture such as U-net, CE-net and AutoDeepLab approaches. As the semantic segmentation approach

is getting more advance, high accuracy of the segmentation results is successfully obtained although the image contains complex, high number and high overlapping of objects in the image. However, this advanced methods such as U-net, CE-net and AutoDeepLab demands extensive computation and large dataset for training. In addition, these methods suffers from parameter value ambiguity. In medical image analysis field, it is common that the number of available dataset is small and limited. Although the semantic segmentation with complex architecture able to segment the targeted object in the medical image, it is not suitable to be used due to its demand of large dataset for training. Therefore, an effective and automated segmentation technique with computationally feasible and less demanding large dataset for training is needed.

CHAPTER 3: METHODOLOGY

This chapter is devoted to the methodology of the proposed contrast enhancement and segmentation framework for the purpose of skeletal age assessment.

3.1 Overview of the Contrast Enhancement Methodology

With an intention to create a holistic histogram equalization framework, the principle objectives of the histogram equalization is studied. It is very clear direction that the contextual purposes are focused into having an enhancement framework that considers brightness preservation, detail preservation and enhanced contrast, simultaneously. Once the contextual purposes are clear, various mechanism in histogram equalization is evaluated in order to achieve the stated purposes. The evaluated mechanism includes bi-histogram equalization approach and clipped histogram equalization approach. In order to optimize each of these techniques, the value of the stopping criteria for each of the mechanism technique is determined. For example, separating point value is one of the stopping criteria need to be determine in the bi-histogram equalization approach. On the other hands, clipped histogram equalization approach needs to determine the value of the clipped limit.

The derivation of holistic contrast enhancement framework is motivated with four principle steps that is illustrated in Figure 3.1. In this figure, the formulation of multiple objective functions is identified as the first step. The second step involves modelling of analytical functions that comprised of multiple objective functions that are corresponding to preservation of brightness shift, contrast and detail loss. The third step is devoted to optimization of analytical functions. Lastly, the last step of this proposed framework is the selection of optimized solutions.

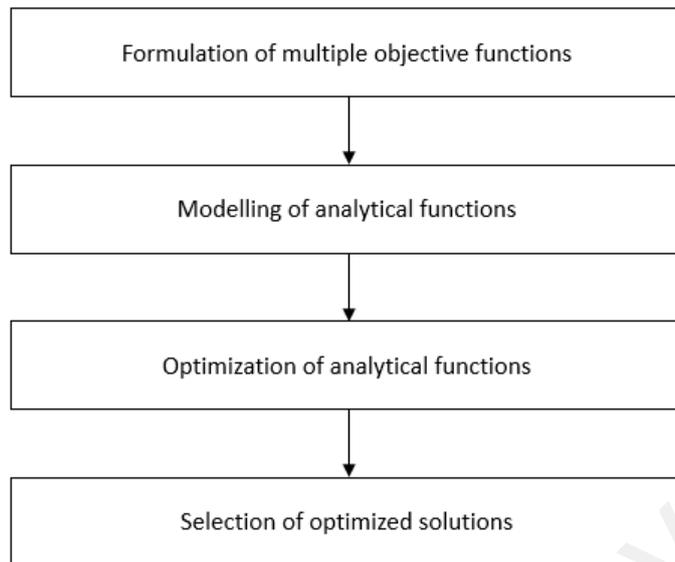


Figure 3.1: Work flow of holistic contrast enhancement framework

3.1.1 Formulation of Multiple Objective Functions

As opposed to the typical approach on using one objective for problem formulation as seen in the literature review chapter, approach of utilizing multiple objectives is put in focus in this thesis. Contrast enhancement is a complex problem and the nature of this type of problem is depending to various consideration factors in order to obtain substantial solution. The consideration factor includes correlation relationship between multiple objectives, type of parameters in each objective and their parameter values of the chosen objective. With the intention to enhance the visual appearance of the skeletal image for evaluation of maturity score, the principal objectives for histogram equalization implementation are designed for this context. With this context, three principal objectives are defined as the desired outcome of output image that is processed by histogram equalization. The objectives are i) enhanced contrast, ii) low mean brightness deviation and iii) low detail loss. The three principal objectives are defined systematically to resemble the human visual perception. In this thesis, the three principal objectives guide

the computation of histogram equalization approach towards improving the visibility of the carpal ossification sites in the radiograph hand bone image. Figure 3.2 depicts the identification of principal objectives towards enhancing the quality appearance of the hand bone image.

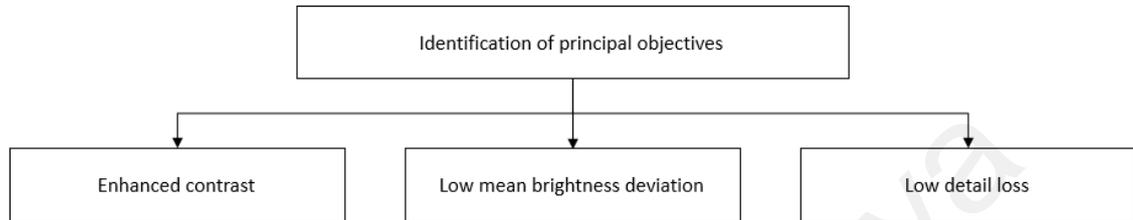


Figure 3.2: Identification of principal objectives

Once the principal objectives are identified, it is crucial to determine the quantitative measurement technique for each identified principal objective. The measurement technique formulates a mathematical model, which it is able to evaluate and quantify the effectiveness of the histogram equalization implementation in achieving the aim of each principal objective. From the existing research works, there is a range of metric techniques that can be used and adopted as a measurement technique for each principal objective. Each of its metric technique has its individual strengths and limitations. In this thesis, mechanism to select the best metric to be adopted for each principal objectives is not the main contribution of this thesis. The main focus of this thesis is the framework that optimize the multiple objectives instead of a single objective that has the ability to enhance the appearance quality of the hand bone image. However, the metric is not only properly chosen to provide measureable quantifier for each principal objective but also ensuring that the final resultant image is correlated to human visual perception. The quantitative measurement of the three principal objectives are described in the following sub-sections. Figure 3.3 illustrates the summary of the chosen quantitative measurement

on the principal objectives that are focus in terms of contrast, brightness and detail considerations.

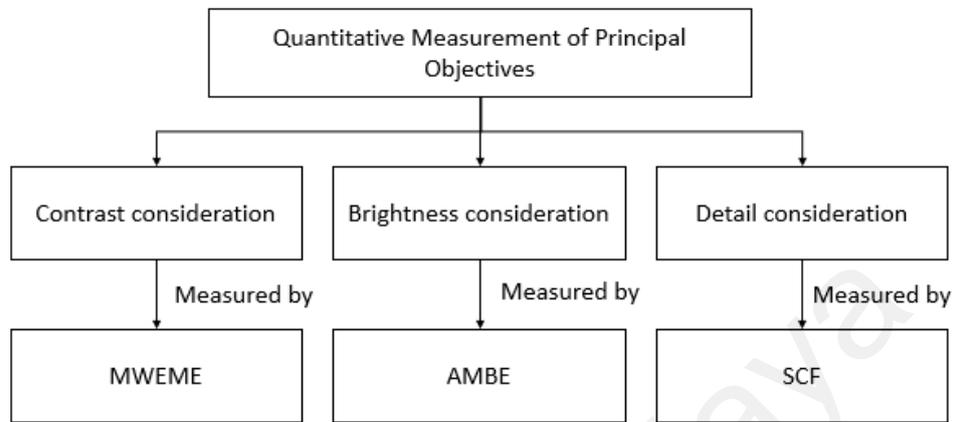


Figure 3.3: Quantitative measurement for contrast, brightness and detail considerations

3.1.1.1 The Measurement of the Enhanced Contrast

In this sub-section, few classical quantitative measurement technique on contrast are discussed. The classical techniques include Michelson Contrast Enhancement (MCM) (Michelson, 1927), Weber contrast enhancement (Peli, 1990) and Measurement of Image Enhancement (EME) (Sos S. Aгаian, 2000; Sundaram, Ramar et al., 2011). From the three classical contrast measurement, researchers also combines the three techniques together with an intention to leverage the strengths from each contrast measurement technique. The combination technique comprises of Weber contrast based EME, Michelson contrast based EME and Weber-Michelson contrast based EME (Hossain and Alsharif, 2007). In this thesis, the contrast measurement technique is reviewed and the relevant technique is chosen based on its ability to enhance the visual appearance of the hand bone image and resemblance of the human visual perception, subsequently.

MCM is designed to measure contrast of the input image that contains periodic intensity patterns. MCM does not exhibits the characteristic of human visual perception

due to its sensitivity to extreme low or high intensity values, which these extreme intensity values may appear only on a small number of pixels in the image. In this context, human is highly robust in terms of slight changes of high or low intensity in the image. The changes do not affect the overall contrast quality of the image in human perception. For Weber contrast enhancement technique, it works well in the image that has homogeneous intensity at the background of the image and the interest object does not covers the entire space of the image. However, hand bone image does not exhibits the characteristic of the image that able to fulfills the criteria that is stated in the Weber contrast enhancement (Hossain and Alsharif, 2007). Therefore, Weber contrast enhancement may not be the most suitable contrast measurement technique for the enhancement of hand bone image.

Equation (3.1) describes the MCM definition for contrast measurement.

$$MCM = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \quad (3.1)$$

where I_{max} and I_{min} denotes the maximum value of the pixel intensity and minimum value of the pixel intensity throughout the image, respectively. From the MCM definition, it is clear that the contrast value measured by MCM is very sensitive to the change of maximum or minimum intensity value, even the outlier value only happened in one single pixel of the image. Due to this over sensitivity to intensity change, MCM is not the most optimum contrast measurement technique if it is used independently and not resembling the human visual perception.

Although global contrast measurement technique such as MCM is very simple to be implemented but this technique is not suitable to be standalone measurement used as contrast difference always appeared in various location in the image. In this thesis, contrast measurement is locally measured using combination of previous adopted contrast measurement techniques in the research works. The combination technique involves

mixture technique of MCM, Weber contrast measurement and EME (MWEME). The local contrast measurement technique, MWEME is defined in Equation (3.2).

$$MWEME = \frac{1}{k_1 k_2} \sum_{m=1}^{k_1} \sum_{n=1}^{k_2} \alpha \left[\frac{I_{max;m,n}^S - I_{min;m,n}^S}{I_{max;m,n}^S + I_{min;m,n}^S + c} \right]^\alpha \ln \left[\frac{I_{max;m,n}^S - I_{min;m,n}^S}{I_{max;m,n}^S + I_{min;m,n}^S + c} \right] \quad (3.2)$$

where $k_1 k_2$ denote the totals number of sub-images in terms of row and column; m by n denotes each sub-image horizontal and vertical size. $I_{max;m,n}^S$ and $I_{min;m,n}^S$ denote the maximum pixel intensity and minimum pixel intensity of each sub-image $I_{S(m,n)}$, respectively. The fraction formula of $\frac{I_{max;m,n}^S - I_{min;m,n}^S}{I_{max;m,n}^S + I_{min;m,n}^S + c}$ represents the MCM formula. The constant c value is included within the formula in order to ensure the stability in the numerical computation. One of the most determinant factors in this MWEME is the parameter α , which this parameter determines the desired range of the most optimum of the output of the contrast enhancement. Thus, α value is defined based on the application objective and problem context. For contrast enhancement on hand bone image, α value is defined in a way to focus to the pertinent features in the image. The α value is also considered with other principal objectives in order to ensure optimal contrast is blended well with others enhancement in terms of brightness and detail. This harmonization is important as these three principal objectives are highly correlated and each enhancement has its own trade-off.

Figure 3.4 illustrates the correspondence value between MCM and MWEME by using different α value. The purpose of this mapping is to illustrate the corresponding intuitive relationship value of MCM and MWEME for different used weight value of α . In Figure 3.4, five α values are used to examine the relationship between value of MCM and MWEME. The α values are 0.5, 1.0, 1.5, 2.0 and 6.5. Other α values are not shown in Figure 3.4 as they show almost similar relationship curve with the five chosen α values.

Thus, only five α values is sufficient to study the relationship between MCM and MWEME as contrast measurement technique. For example at $\alpha = 0.5$, MWEME value is highest at the MCM value approximate to 0.1. The highest MWEME value is reflected at different MCM value if different α is used. For instance, the highest MWEME value is located at MCM value of 0.5 if $\alpha = 1.5$ is used, as compared to the previous example of the highest MWEME value is at $\alpha = 0.5$ for MCM value of 0.1. With this observation of the MWEME value by different α value, the setting of α value is dynamic in the development of holistic equalization. The α value is chosen based on the characteristic of the input image and also the correlation with others principal objectives.

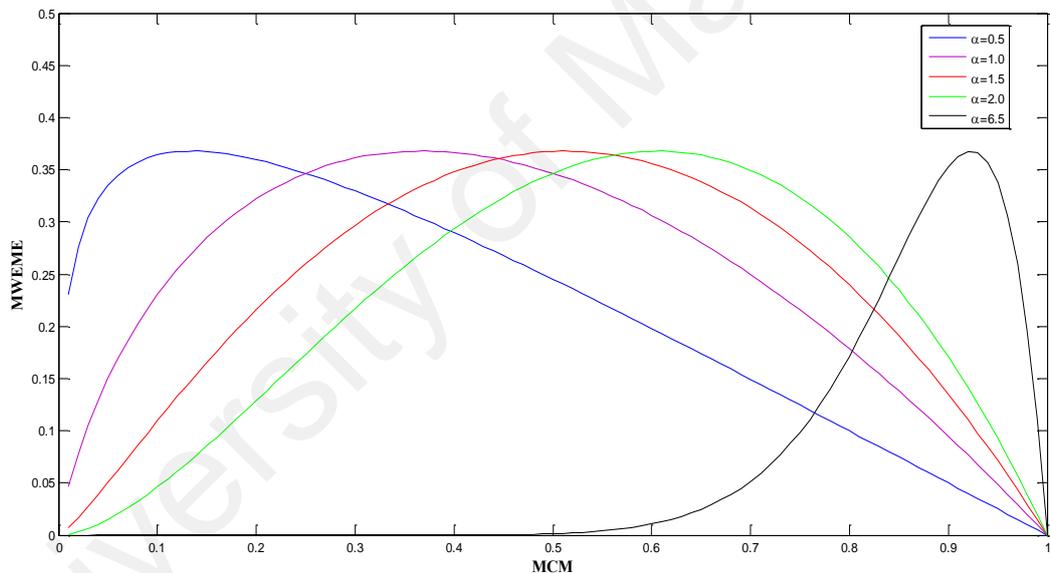


Figure 3.4 The correspondence value between MCM and MWEME with α variation values

Figure 3.5 illustrates the original image of radius ossification site and three resultant images with different visual artifacts from enhancement. The implication of visual artifacts is observed from various parameters, such as k_1, k_2, α of MWEME. The observation also motivates the usefulness of local contrast enhancement as compared to the global enhancement and at the same time, relationship of MWEME to human visual

perception is observed. From Figure 3.5, it is observed that histogram equalization generates visual artifacts on the enhanced image. This visual artifacts hinder practitioner to inspect the radiographic image and also degraded the performance of the automated skeletal age assessment system in measuring the development stage of the skeletal bone.

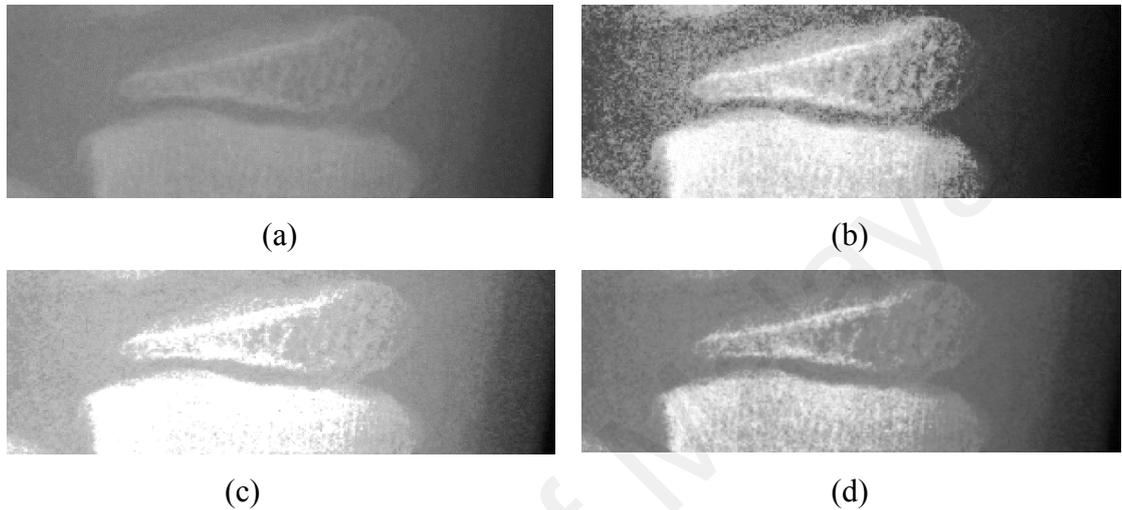


Figure 3.5 (a) Original Image of radius ossification sites (b) Amplified background noise found at the histogram equalized image (c) Unwanted intensity saturation and high mean brightness shift of the histogram equalized image (d) Well enhanced image without obvious unwanted artifacts.

Table 3.1 describes the contrast measurement value using MWEME and MCM technique for each sub-figure (a) to sub-figure (d) from Figure 3.5, where different parameters value of k_1 , k_2 and α are used. As described in previous paragraph, MCM is the partial input to the MWEME formula. As MCM does not have α value, there is no variation for MCM as illustrated in Table 3.1. However, the α value in MWEME reflects the relative importance of this value in emphasizing pertinent information in the image for contrast enhancement purpose.

In this table, it is noted that contrast measurement using $k_1 = 1, k_2 = 1$ equivalents to global measurement technique. For global measurement technique, the majority of the contrast measurement result value is either close to zero (0) or one (1). This result value

is expected as the MCM considers the minimum and maximum intensity values only for contrast measurement value. If the whole image is under consideration, it is common to obtain the contrast measurement result value is either close to zero (0) or one (1) as the minimum and maximum intensity value is zero and 255, respectively. Therefore, the contrast measurement value using MCM will be close to one and MWEME will returns value closely to zero. As the four sub-figures has different visual qualities, the value of the global contrast measurement is almost the same (zero or one) reflects that this global technique is not reflecting human visual perception. For example, the global contrast measurement value in Figure 3.5(b) and Figure 3.5(d) is identical although the image quality in these two sub-figure is different as perceived by human perception. Therefore, the local processing that perform contrast measurement at different region in the image is more favor to visual human perception.

From Figure 3.5, it is observed that the local MWEME value for the four sub-figures fall within the specific minimum MWEME (k_1, k_2, α) value and maximum MWEME (k_1, k_2, α) value, regardless the quality of the image as perceived by human. Therefore, a typical histogram technique that only put focus in improving the contrast quality of the image may not able to obtain the output image that is equivalent to visual human perception. It is clearly reflected by the local contrast measurement value at Figure 3.5(c), where the value is almost identical to Figure 3.5(d) although the image in Figure 3.5(c) is suffers from brightness and detail preservation. The observation from the results in Table 3.1 strengthen the motivation of this thesis, which is to develop a holistic histogram equalization approach for the enhancement of hand bone image.

Table 3.1 Contrast measurement value with MWEME and MCM with different k_1 , k_2 and α values

Figure 3.5	Configuration of Contrast Metrics	$k_1 = 1;$	$k_1 = 5;$	$k_1 = 10;$	$k_1 = 15;$	$k_1 = 20;$	$k_1 = 40;$
		$k_2 = 1$	$k_2 = 5$	$k_2 = 10$	$k_2 = 15$	$k_2 = 20$	$k_2 = 40$
(a)	MWEME ($\alpha = 0.5$)	0.0693	0.3066	0.3372	0.3464	0.3479	0.3407
	MWEME ($\alpha = 1.0$)	0.1286	0.2834	0.2647	0.2417	0.2257	0.1816
	MWEME ($\alpha = 1.5$)	0.1790	0.2117	0.1675	0.1351	0.1173	0.0777
	MWEME ($\alpha = 2.0$)	0.2215	0.1548	0.1047	0.0747	0.0609	0.0338
	MCM	0.8613	0.3128	0.2098	0.1593	0.1349	0.0867
(b)	MWEME ($\alpha = 0.5$)	0	0.1785	0.2397	0.2740	0.2891	0.3227
	MWEME ($\alpha = 1.0$)	0	0.2469	0.3013	0.3179	0.3186	0.3061
	MWEME ($\alpha = 1.5$)	0	0.2659	0.2966	0.2904	0.2772	0.2337
	MWEME ($\alpha = 2.0$)	0	0.2615	0.2677	0.2444	0.2232	0.1680
	MCM	1	0.6307	0.4959	0.4148	0.3766	0.2744
(c)	MWEME ($\alpha = 0.5$)	0.0078	0.2901	0.3262	0.3305	0.3350	0.3199
	MWEME ($\alpha = 1.0$)	0.0154	0.2877	0.2818	0.2590	0.2502	0.2014
	MWEME ($\alpha = 1.5$)	0.0230	0.2225	0.1925	0.1611	0.1487	0.1027
	MWEME ($\alpha = 2.0$)	0.0304	0.1611	0.1243	0.0952	0.0844	0.0508
	MCM	0.9844	0.3656	0.2537	0.1945	0.1723	0.1125
(d)	MWEME ($\alpha = 0.5$)	0	0.2591	0.3121	0.3332	0.3412	0.3496
	MWEME ($\alpha = 1.0$)	0	0.2995	0.3092	0.2964	0.2856	0.2432
	MWEME ($\alpha = 1.5$)	0	0.2674	0.2383	0.2064	0.1880	0.1352
	MWEME ($\alpha = 2.0$)	0	0.2188	0.1697	0.1339	0.1159	0.0717
	MCM	1	0.4512	0.3165	0.2491	0.2169	0.1451

3.1.1.2 The Measurement of the Mean Brightness Deviation

Brightness preservation is indicated by the low value of mean brightness deviation. Mean brightness deviation is one of the measurement of the artifacts that caused by the histogram equalization. High value of the mean brightness deviation in the enhanced

image deteriorates the quality of the input image and at the same time creates unnatural visual effects. The effect of mean brightness deviation is reflected in the resultant image at Figure 3.5(c) compared to original image in Figure 3.5(a), where the increase brightness of the resultant image reduces the contrast of the original image. The contrast reduction of the resultant image at Figure 3.5(c) as compared to Figure 3.5(d) is illustrated in Table 3.1.

In order to address this problem, various approaches on histogram equalization are introduced to minimize the enhancement artifacts that related to mean brightness deviation. The introduced histogram equalization approaches include brightness preserving bi-histogram equalization (BBHE) (Yeong-Taeg, 1997, Yeong-Taeg, 1997), dualistic sub-image histogram equalization (DSIHE) (Yu, Qian et al., 1999), minimum mean brightness error bi-histogram equalization (MMBEBHE) (Soong-Der and Ramli, 2003), brightness preserving histogram equalization with maximum entropy (BPHEME) (Chao and Zhongfu, 2005), recursive mean-separate histogram equalization (RMSHE) (Sim, Tso et al., 2007) and brightness preserving dynamic histogram equalization (BPDHE) (Ibrahim and Kong, 2007).

The above mentioned proposed histogram equalization approaches use brightness preservation measurement method in order to evaluate the effectiveness of the enhancement approach in preserving the brightness of the image. The brightness preservation measurement method includes absolute mean brightness error (AMBE) or the Mean of AMBE (MAMBE). The equation of AMBE and MAMBE is illustrated in Equation (3.3) and Equation (3.4), respectively.

$$AMBE(X, Y) = |E(X) - E(Y)| \quad (3.3)$$

where $E(X)$ denotes the mean brightness of input image and $E(Y)$ denotes the mean brightness of resultant image. For multiple resultant images, N involve form the computation of the histogram equalization, the mean of the AMBE (MAMBE) is computed to measure the overall effectiveness of the proposed algorithm in terms of brightness preservation:

$$MAMBE(X, Y) = \frac{1}{N} \sum_{i=1}^N |E(X) - E(Y)| \quad (3.4)$$

At the early days of research works, low value of AMBE and MAMBE show high brightness preservation in the enhanced image. There is also findings about the importance of brightness preservation while performing contrast enhancement. Thus, this group of researcher designed various histogram equalization algorithms that will enhance the input image with minimum value of AMBE. However, this is only partly true due to few reasons. Firstly, the resultant image with low value of AMBE may insignificantly enhanced. The second reason is devoted to the resultant image that just focus in brightness preservation might not produce perceptually meaningful enhanced image to human vision system (Singh and Vikram, 2002, Zuo, Chen et al., 2013).

The authors in the paper (Chen Hee and Isa, 2010) who designed MMBEBHE method used AMBE as the metric for brightness preservation. They claimed and justified that the AMBE as the brightness preservation metric does not resemble human visual perception in terms of quality measurement of histogram equalization. Despite to this drawbacks of AMBE (Konak, Coit et al., 2006), the measurement metric for the brightness preservation is still adopting AMBE technique in this thesis. This adoption is reasonable due to the following justifications:

- 1) The limitation of AMBE in achieving enhancement result that resembling human visual perception is compensated by introducing optimization of multiple objectives that includes enhancement on contrast, detail lost and occurrences of other artifacts. In this thesis, the aggregation of various standard measurement metric as a constraint to each principal objective is one of the contributions in order to solve the limitation of the standard metric by performing optimization on the aggregated multiple objective function.
- 2) The simplicity of AMBE compared to other brightness preservation metrics that requires a number of user-specified parameters is one of the important factors to be considered as a metric measurement in the optimization of multiple objective function. AMBE does not contain user-specified parameters and can be executed automatically without human intervention. These two properties make the execution of histogram equalization is practical to be used in the image enhancement process as it is done without any presence of human expert. In addition, the simplicity of AMBE assures feasibility in terms of computational complexity in solving the optimization of multiple objectives.

3.1.1.3 The Measurement of Detail Loss

For the artifacts that related to detail loss, it can be described by the artifacts caused by intensity saturation and amplification of background noise in the enhanced image. In this context, the details represent the highly important information of the interest objects in the processed image. The loss of this highly important information contributes to the failure of visual inspection or automatic classification of skeletal age assessment system due to the pertinent information does not exist in the processed image. For skeletal age assessment, the pertinent information exists in the ossification sites of the hand bone

image. Thus, artifacts that lead to loss of information at the ossification sites will directly hinder radiographer or pediatrician to perform visual inspection in order to make bone age decision if GP method is used. For automated system such as CASAS, the loss of this information hinders the right features to be extracted from this region of interest and subsequently causing the inaccurate of the prediction of actual bone age. Therefore, it is obvious that consideration to preserve the details in the enhanced image is highly crucial in skeletal age assessment and construction of multiple objectives function is necessary to ensure the resultant image from histogram equalization is resemble visual human perception. Due to the urgency to preserve the details in the ossification sites, a measurement metric for detail loss is needed for this purpose. Firstly, various possible measurement metric related to detail loss that is adopted by previous researchers is studied. Then, the most suitable detail loss measurement is chosen as the metric component in the multiple principal objectives function.

Shannon entropy (Chao and Zhongfu, 2005) is used as a measurement metric to determine the efficiency of histogram equalization in highlighting details, although this metric also been used for contrast preservation. In (Chen Hee and Isa, 2010) research work, they claimed that the higher value of entropy indicates the better contrast enhancement or detail preservation. Besides entropy, authors in this paper (Sengee, Bazarragchaa et al., 2009) evaluate the detail preservation after the enhancement process by using the Peak noise signal ratio (PNSR). Research work claimed that the higher value of PSNR represents a better visual effect. It is also noted that PSNR also been used to evaluate the efficiency of contrast enhancement (Rajavel, 2010). Therefore, detail loss measurement metric such as entropy and PSNR may not be directly correlated to the detail preservation as the metrics are generic for overall contrast measurement metric. Thus, entropy and PSNR are not chosen as metric for measuring the degree of detail loss in this thesis. In this context, structure comparison function (Zhou and Bovik, 2002) is adopted

for detail loss measurement in histogram equalization. Structure comparison function is termed as SCF. It is one of the essential components from the structural similarity index (Zhou, Bovik et al., 2004). SCF is described as follows in Equation (3.5) and Equation (3.6).

$$SCF(X, Y) = \frac{\sigma_{XY} + C}{\sigma_X \sigma_Y + C} \quad (3.5)$$

$$\sigma_{XY} = \frac{1}{N-1} \sum_{i=1}^N (X_i - \mu_X)(Y_i - \mu_Y) \quad (3.6)$$

X_i and Y_i denote the i -th pixel intensity in input image and output image, respectively. μ_X and μ_Y denote the mean intensity of input image and output image, respectively. C denotes a small constant value that ensures computational stability.

From the literature reviews, histogram equalization that has a single objective in removing specific artifact in the image fails to ensure high visual quality of the enhanced image. The single objective can be the minimization of mean brightness shift or maximization of contrast in the image. One of the reasons is that the adopted metrics such as entropy and AMBE fail to reflect the actual perceived visual effect that corresponding to human visual perception. Therefore, majority of the previous histogram equalization focuses on one single objectives while neglecting others important objective. BPHEME does consider both objectives of contrast and brightness for the optimal histogram equalization result using histogram specification technique. However, simultaneous consideration of two objectives is not sufficient for enhancement purpose in complex images such as medical images, specifically to hand bone image.

3.1.2 Modelling of Analytical Functions

The general principle to evaluate the performance of each objective function is to compare the fitness results of each computed objective function. However, independent comparison of the fitness value for each objective function is not an effective method as all these objective functions are correlated to each other. Therefore, the overall fitness value is determined from a single solution, where this single fitness solution is computed by aggregating the objective functions. The aggregation of multiple objective functions is known as modelling of analytical functions. Figure 3.6 illustrates the three major components that model the analytical function. The three components are the objective functions, decision variables and constraints. In Figure 3.6, the objective functions define the effectiveness of the possible feasible solutions in the proposed approach. Meanwhile, the decision variables values generate all the possible feasible solutions based on the proposed approach. In the optimization process in getting the feasible solution, there is no ideal solution that is able to simultaneously optimize all the objective functions. In fact, a certain trade-off exists in each optimized objective function. This trade-off is defined by the initial constraint set based on the desired goal of the optimization. The constraint set reduces the searching space in obtaining a feasible solution. Therefore, the searching space is clearly defined only certain range of solutions are feasible to the problem, where not all range contains possible solutions. This feasibility is defined by the problem constraints according to the desired goal, which greatly reduces the searching space in order to ensure a feasible solution is obtained. The goal of optimization is to obtain a solution that is feasible by adhering to the constraints. It can be concluded that the constraints for each of the objective functions and their trade-off are ideally set in accordance to the problem context.

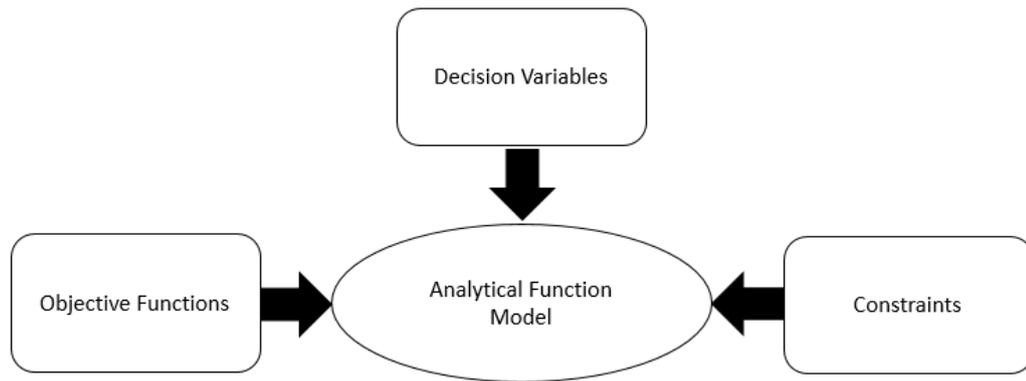


Figure 3.6: Formulation of analytical function model by objective functions, decision variables and constraints.

In the proposed contrast enhancement framework, histogram equalization (HE) is chosen as the approach to enhance the visibility of the hand bone image. Thus, the analytical function model is defined as histogram equalization model. Figure 3.7 describes the histogram equalization model, which this model is formed from three objective functions, two decision variables and two constraints. In Figure 3.7, the three objective functions are the principal objectives that are formulated in Section 3.1.1. These objective functions are enhanced contrast, low mean brightness deviation and low detail loss. Due to the proposed contrast enhancement framework is utilizing bi-histogram equalization approach and clipped histogram equalization approach, the two decision variables are separating point and clipped limit values. In order to limit the searching range of the decision variable that reflects the possible feasible solutions, two constraints are introduced in the histogram equalization model. The constraints are search range separating point and search range clipped limit. The following section describes the details of the decision variables and constraints used in the proposed contrast enhancement framework.

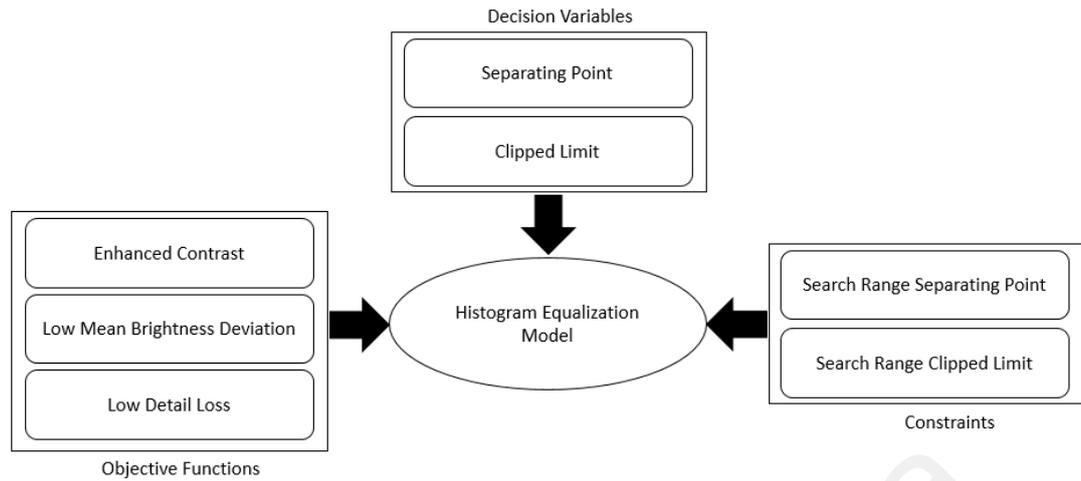


Figure 3.7: Formulation of histogram equalization model.

In sub-histogram equalization approach, the separating point is a threshold value that is used to separate the histogram of the input image into two sub-histograms. For each sub-histogram, histogram equalization is performed independently over each sub-histogram. The suitability of the calculated separating point is highly important as the intention of having separated sub-histogram is to preserve the mean brightness of the input image. The second decision variable, which is the “Clipped Limit” clips both the sub-histograms prior to histogram equalization process. The clipping step is highly crucial in order to prevent over-enhancement and detail loss that may effects the output image after histogram equalization process (Chen Hee, Kong et al., 2009, Pizer, Amburn et al., 1987). The “Clipped Limit” consists of “Clipped Limit-L” and “Clipped Limit-R”. “Clipped Limit-L” depicts the clipped limit that is applied into the first sub-histogram after histogram separation. On the other hand, “Clipped Limit-R” is computed for the clipped limit for the second sub-histogram after histogram separation. In this thesis, Non-dominated Sorting Genetic Algorithm (known as NSGA-II) (Deb, Pratap et al., 2002) is adopted as solution to the multi-objective optimization problem. The output of the NSGA-II is the non-dominated decision variables vector, which is denoted as \mathbf{X}^S . Once

the non-dominated decision variables vector is obtained, a set of enhanced resultant images by using \mathbf{X}^s is generated.

The average intensity of an image, I , is defined as A^I :

$$A^I = \frac{1}{(i_{max})(j_{max})} \sum_{i=1}^{i_{max}} \sum_{j=1}^{j_{max}} I(i, j) \quad (3.7)$$

where $I(i, j)$ describes the intensity values at coordinate (i, j) of a two-dimensional grayscale image. The multiplication of i_{max} and j_{max} resemble the total number of pixels in the image. The computed value of A^I is used to define the sub-images and their constraints. The sub-histograms are defined as following:

$$H^I = H^L \cup H^U \quad (3.8)$$

In Equation (3.8), H^I describes the histogram distribution of intensity of image I . H^L resembles the lower sub-histogram with bin index that is distributed from 0 to A^I . H^U represents the upper sub-histogram with bin index that is distributed from $A^I + 1$ to maximum gray level. In intensity histogram context, the histogram with bin index, B describes the gray level intensity value, where the value of this specific bin, describes the total number of pixels in the image that having the same gray level intensity value, B . The value of the histogram bin, which denoted as n_B is actually represents the height of the histogram bin. Let A^L denotes the average pixel value of intensity in H^L . It is defined as the summation value of all the multiplication of each bin index with its bin value, then divided by the total number of bin values within H^L denoted as N^L .

$$A^L = \frac{1}{N^L} \sum_{B=1}^{B=A^I} (n_B) (B) \quad (3.9)$$

The average pixel value of intensity in H^U is computed similar as H^L , and it is denoted as A^U . In this definition, N^U represents the total bin values across H^U and $Lmax$ is the maximum gray level intensity value. A^U is illustrated in Equation (3.10) as follows:

$$A^U = \frac{1}{N^U} \sum_{B=A^L+1}^{B=Lmax} (n_B) (B) \quad (3.10)$$

The solutions \mathbf{x}^s subject to three constraints, denoted as $g_1(\mathbf{X}^s)$, $g_2(\mathbf{X}^s)$ and $g_3(\mathbf{X}^s)$:

$$g_1(\mathbf{X}^s) \text{ is termed as:} \quad (3.11)$$

$$A^L < x_1^s < A^U$$

Let N^{bL} is defined as total number of bins in H^L ,

$g_2(\mathbf{X}^s)$ is termed as:

$$A^{bL} - \sigma^L < x_2^s < A^{bL} + \sigma^L$$

where

$$(3.12)$$

$$\sigma^L = \sqrt{\frac{1}{N^{bL}} \sum_{B=1}^{B=A^L} (A^{bL} - n_B)^2}$$

$$A^{bL} = \frac{1}{N^{bL}} \sum_{B=1}^{B=A^L} n^B$$

Let N^{bU} is defined as total number of bins in H^U ,

$$g_3(\mathbf{X}^s) \text{ is defined as:} \quad (3.13)$$

$$A^{bU} - \sigma^U < x_3^s < A^{bU} + \sigma^U$$

where

$$\sigma^U = \sqrt{\frac{1}{N^{bU}} \sum_{B=A^L+1}^{B=Lmax} (A^{bU} - n_B)^2}$$

$$A^{bU} = \frac{1}{N^{bU}} \sum_{B=A^L+1}^{B=Lmax} n^B$$

x_1^s , x_2^s and x_3^s represent the three non-dominated decision variable, which is representing “Separating point”, “Clipped Limit-L” and “Clipped Limit-R”, respectively. The notation “ σ^L ” and “ σ^U ” denote the standard deviation of bin value for H^L and H^U respectively. The notation “ A^{bL} ” and “ A^{bU} ” represent the average bin value for H^L and H^U , respectively. It is highly crucial to introduce constraint to the decision variables in order to limit the feasible search space for optimization purpose. With this limitation, the duration for optimization is accelerated and optimization result is directed to the initial set objectives. For standard deviation of $g_2(\mathbf{X}^s)$ and $g_3(\mathbf{X}^s)$, both values provide an adaptive search space with the reference at the average bin value for both lower and upper sub-histogram, respectively. For sub-histogram with less variation, standard deviation is adjusted to allow narrower search space as compared to the sub-histogram with bigger variation. Therefore, the adjustment of the standard deviation of $g_2(\mathbf{X}^s)$ and $g_3(\mathbf{X}^s)$ is highly dependent to the characteristic of the sub-histogram in different input image.

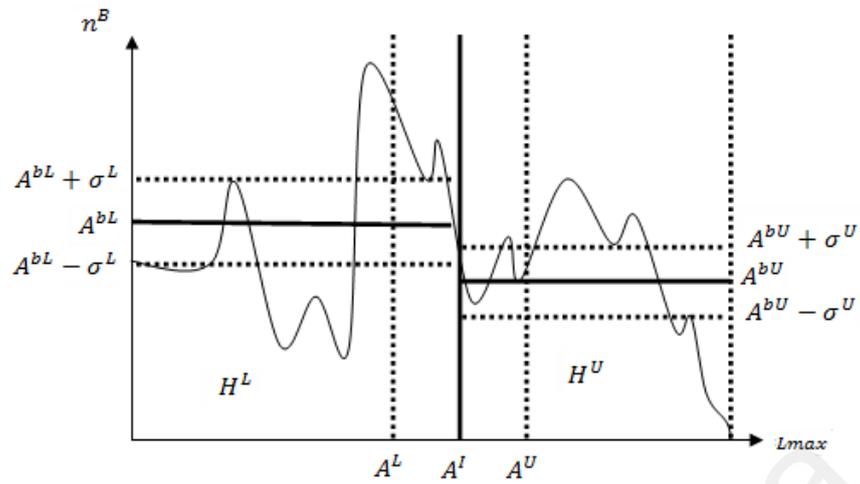


Figure 3.8: Pre-determined decision variables search ranges of “Clipped Limit – L” (ranges from $A^{bL} - \sigma^L$ to $A^{bL} + \sigma^L$), “Clipped Limit – R” (ranges from $A^{bU} - \sigma^U$ to $A^{bU} + \sigma^U$) and “Separating Point”(ranges from A^L to A^U).

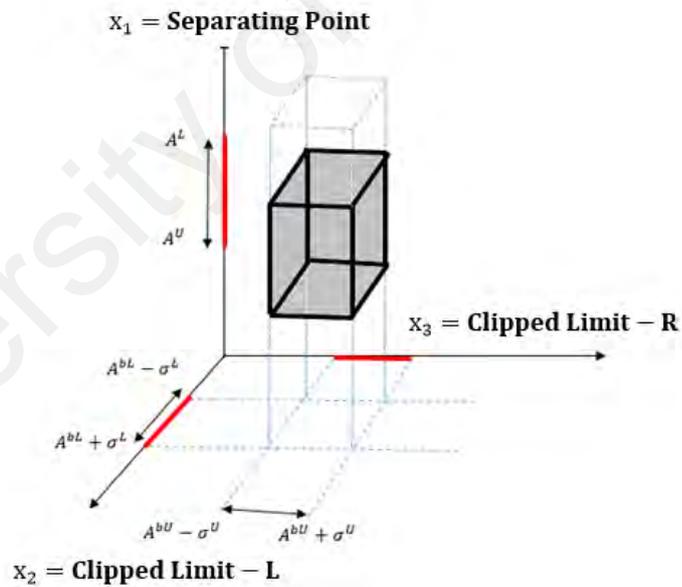


Figure 3.9: Solution Space, X.

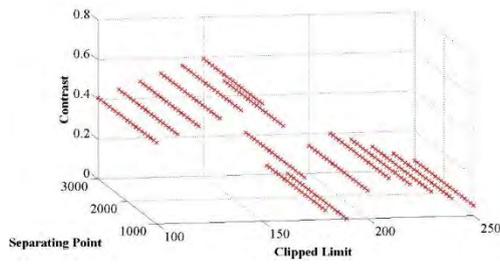
In Figure 3.8, the search space for the decision variables are prior determined for optimization purpose. The “Clipped Limit-L” is ranged from $A^{bL} - \sigma^L$ to $A^{bL} + \sigma^L$ and “Clipped Limit-R” is ranged from $A^{bU} - \sigma^U$ to $A^{bU} + \sigma^U$. For “Separating Point”, it is

ranges from A^L to A^U . Figure 3.9 illustrates the solution space, X , within the search space of three decision variables, which are the “Separating Point”, “Clipped Limit-L” and “Clipped Limit-R”. In this thesis, the dimension of decision variable is further reduce with the intention to enhance the efficiency of the optimization process and also reduce the computation time to obtain the optimum decision variables. The two decision variables: “Clipped Limit-L” and “Clipped Limit-R” to only one decision variable, known as “Clipped Limit”. These two decision variables are valid to be combined as the gray level intensity distribution in the region of interest of the carpal bone radiographic image is almost symmetrical in histogram representation. As the objective of the clipping process is to reduce the disadvantage of over enhancement that may results washed out effect, this over enhancement effect is mainly contributed by intensity bin with high excess value. With this understanding, the objective of the clipping process is still achievable by having one clipping limit value for both the lower and upper histograms. By reducing one decision variable, the computation time to obtain optimum decision variable values is faster as the computation time for additional decision variables is at the exponential rate. Therefore, only two decision variables are considered in the optimization, which is the x_1 as “Separating Point” and x_2 as “Clipped Limit”. The range of Clipped Limit is termed in Equation (3.14).

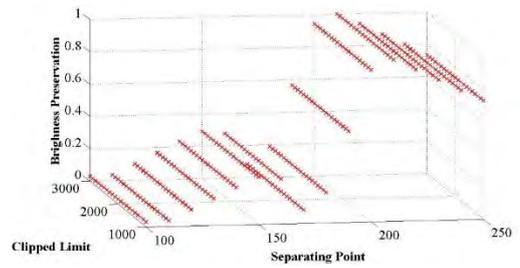
$$\frac{(A^{bL} + A^{bU})}{2} - \frac{(\sigma^L + \sigma^U)}{2} < x_2^s < \frac{(A^{bL} + A^{bU})}{2} + \frac{(\sigma^L + \sigma^U)}{2} \quad (3.14)$$

Prior to optimization process in obtaining the optimized solutions, it is an essential step to compute the defined analytical function, which it describes the relationship between each determined objective function with the decision variables. It has four fundamental steps to compute this analytical function. The first step involves the processing operation of the input image with specific values of separating point, x_1 and

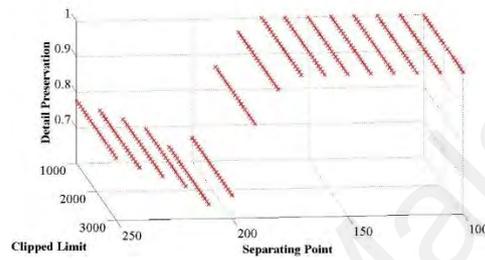
clipped limit, x_2 . The second step is devoted to the computation of the corresponding objective functions with $\mathbf{X} = \{x_1, x_2\}$. The corresponding objective functions are represented as $z_1(\mathbf{X})$, $z_2(\mathbf{X})$ and $z_3(\mathbf{X})$ for contrast, brightness preservation and detail preservation, respectively. The third step involves the sampling process of $\mathbf{X} = \{x_1, x_2\}$ in order to obtain a set x_1 and x_2 values with the final objective in mind to obtain the optimized solution. Optimized solution is termed as $\mathbf{X}^s = \{x_1^s, x_2^s\}$ where the first solution is labelled as $\mathbf{X}^{s=1} = \{x_1^{s=1}, x_2^{s=1}\}$ and the following solution is labelled with the similar representation. The size of the $\mathbf{X} = \{x_1, x_2\}$ is depending to the resolution of the sampling. Higher resolution sampling corresponds to the higher accuracy estimation of the relationship between the objective functions with the decision variables. However, the balance of the resolution sampling and the relationship accuracy is justified as the higher resolution sampling demands a longer computation time. After decision of the sampling resolution and subsequent generation of the corresponding objective value, curve fitting process is generated to model the relationship between the decision variables and objective functions in the fourth step. The final outcome of the curve fitting process is an analytical function. This analytical function is used in performing the multi-objective optimization. Figure 3.10 illustrates the three figures of the generated values of objective functions for a region of interest in carpal bone, Hamate. For example of Figure 3.10(a), each point in this figure represents the contrast value of the resultant image of bi-histogram equalization, where the image is equalized using a certain value of Clipped Limit and Separating Point. Figure 3.10(b) illustrates the brightness preservation value of the resultant image from equalization and Figure 3.10(c) shows the detail preservation value of the similar resultant image. The total number of points in each figure is corresponding to the defined sampling resolution for each decision variable as well as the intensity attributes of the input image that determine the searching range. The searching range includes $\sigma^L, \sigma^U, A^L, A^U$.



(a)

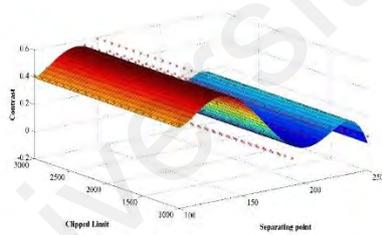


(b)

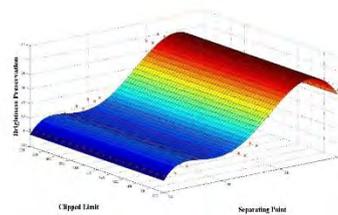


(c)

Figure 3.10: Computation of each objective value of the resultant image after histogram equalization process with the sampled decision variables values (a) Contrast value (b) Brightness Preservation value (c) Detail Preservation value.



(a)



(b)

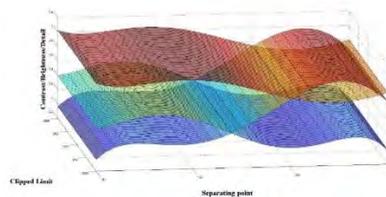
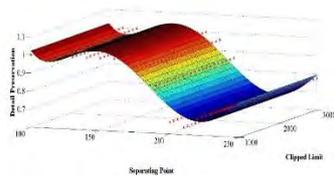


Figure 3.11: Fitted curve plane for each objective (a) Contrast (b) Brightness Preservation (c) Detail Preservation (d) Combinations of all fitted curve planes in the same Cartesian space.

Figure 3.11 illustrates the three figures of the fitted curve plan of the three objective functions as depicted in Figure 3.10. Figure 3.11 (d) combines all the fitted curve plans in the same Cartesian space.

For each computed set of values of objective function, a curve fitting procedure is applied to these points in order to obtain an analytical surface. This analytical surface fits the points, optimally. Optimum fitting of the analytical surface is the most practical solution to describe the relationship between the objective function and the decision variables. This is due to the natural characteristic of the data point values that may contain noises, which known as outlier points. A well fitted analytic surface may include these noises into the fitting function. With the noise presence in it, the relationship between the objective function and the decision variables is not well represented. Thus, an approximate function that is able to describe the data points as an overall representation is a better solution with a broader perspective sense. This approach is termed as curve fitting process. In this curve fitting process, a vector of points that is equally spaced in the interval is generated. This vector of points is represented by a plane function, $p(\cdot)$. The plane function is generated to fit the points of objective function that corresponding with the two decision variables. The plane function is a three dimensional polynomial curve function with a degree N that utilizes least square method to perform the fitting onto the vector of points. The implementation of this procedure generates coefficients for a polynomial curve (or “plane” for 3 dimensional) of degree N that fit optimally with least square method onto the generated data points. The coefficients of the polynomial function that is represented as m parameters, r_1, \dots, r_m are in the order of descending powers and the length of the polynomial function is described as m. The functionality of the least squares method is as an optimizer by adjusting the values of m parameters, r_1, \dots, r_m , for a certain functional form $p(\mathbf{X}, r_1, \dots, r_m)$ such that to attempt fitting onto the given set of points $\mathbf{X} = \{x_1^i, x_2^i, o_z^i\}_{i=1}^{n+1}, z=1,2,3$. o_z^i denotes objective function, where $o_{z=1}^i$ represents

contrast objective, $o_{z=2}^i$ represents brightness preservation objective and $o_{z=3}^i$ represents detail preservation, respectively. In this formulation, $n+1$ represents total number of data in the vector points. The curve fitting is performed by minimizing the error term of $e^2 = \sum_{i=1}^{n+1} (p(x_1^i, x_2^i, o_z^i, r_1, \dots, r_m) - x_1^i, x_2^i, o_z^i)^2$. The following illustrates the example of description of fitted function for contrast objective function. Let the m coefficient parameters for contrast objective function to be denoted as $\{r_1^{z=1}, \dots, r_m^{z=1}\}$, the fitted function is illustrated as below. The fitted function for another two objectives, which is brightness preservation and detail preservation is similar as contrast but with different value of parameters.

$$\begin{aligned}
p(x_1, x_2)^{z=1} = & r_1^{z=1} + r_2^{z=1}x_1 + r_3^{z=1}x_2 + r_4^{z=1}x_1^2 + r_5^{z=1}x_1x_2 + r_6^{z=1}x_2^2 \\
& + r_7^{z=1}x_1^3 + r_8^{z=1}x_1^2x_2 + r_9^{z=1}x_1x_2^2 + r_{10}^{z=1}x_2^3 + r_{11}^{z=1}x_1^4 \\
& + r_{12}^{z=1}x_1^3x_2 + r_{13}^{z=1}x_1^2x_2^2 + r_{14}^{z=1}x_1x_2^3 + r_{15}^{z=1}x_2^4 + r_{16}^{z=1}x_1^5 \\
& + r_{17}^{z=1}x_1^4x_2 + r_{18}^{z=1}x_1^3x_2^2 + r_{19}^{z=1}x_1^2x_2^3 + r_{20}^{z=1}x_1x_2^4 + r_{21}^{z=1}x_2^5
\end{aligned}$$

3.1.3 Optimization of Analytical Function

Optimization method is applied to the defined analytical function model in order to obtain optimum solutions for the defined problem context. In the proposed contrast enhancement framework, the histogram equalization model is optimized. In mathematical expression, this model is constructed prior to the optimization (Srinivas and Deb, 1994, Zitzler and Thiele, 1999) as follows:

$$z = \operatorname{argmin}[z_1(\mathbf{X}), z_2(\mathbf{X}), z_3(\mathbf{X}), \dots, z_K(\mathbf{X})] \quad (3.15)$$

K denotes the number of objectives, $z_i(\cdot)$. $\mathbf{X} = (x_1, x_2, x_3, \dots, x_N)^T$ denotes the vector of N number of decision variables. In this optimization, a set of optimum preference for

each objective is searched by emphasizing the features contain in the carpal ossification sites. The outcome of this optimization is to determine the favorable trade-offs that compromise all the desired criteria of the objective functions instead of finding the global optimized solution. The outcome of the multiple objectives optimization is a set of preference solutions. However, a set of preference solutions is not applicable to the process of histogram equalization as this enhancement process only demands a single set of solution. Therefore, a set of relative preference solution from each criteria in the multiple objective functions is computed for this purpose. This relative preference solution is depends to the chosen problem context. For this thesis, the problem context is to enhance the visual effect of the pertinent features lies in the ossification sites of the carpals. In this thesis, GHE (Bull, Edwards et al., 1999) is chosen as the contrast enhancement approach to be used.

In this thesis, an evolutionary algorithm method, which is known as Non-Dominated Sorting Genetic Algorithm (NSGA II) technique (Deb, Pratap et al., 2002) is chosen as an optimization technique to be applied to the histogram equalization model. The output of the NSGA-II is the non-dominated solution, which is also known as Pareto-optimal solution. In the Pareto-optimal solution, this is the most ideal and optimum solution as any further improvement towards the overall fitness value of the objective functions will degrade the fitness value of one of the objective functions. In the proposed contrast enhancement framework, the Pareto-optimal solution is the vector of decision variables that optimally enhance the hand bone image by considering the preservation of contrast, brightness and detail. The following section addresses the non-dominated solution in mathematical expression.

In general, optimization of the objective function is classified as either minimization process or maximization process. In minimization process, the optimization is carried out

in order to achieve the lowest value of the objective function. On the other hand, maximization process is executed in order to achieve the highest value of the objective function. Given a N -dimensional decision variable vector $\mathbf{X} = \{x_1, \dots, x_N\}$ in the solution space \mathbf{X} , find a vector \mathbf{X}^s that optimizes a given set of K objective functions in order to obtain non-dominate solutions: $\mathbf{Z}(\mathbf{X}^s) = \{z_1(\mathbf{X}^s), \dots, z_K(\mathbf{X}^s)\}$ for $s = 1, \dots, T$, where T represents the total number of non-dominated solutions. The solution space \mathbf{X} is designed to be restricted by M number of constraints, such as $g_j(\mathbf{X}^s)$, where $j = 1, \dots, M$, which bounds on the decision variables: $\mathbf{X} = \{\mathbf{X} \in R^N | g_j(\mathbf{X}^s) \leq 0, j = 1, \dots, M\}$. In the minimization of the objective functions, a feasible solution \mathbf{X}^1 is defined to dominate another feasible solution \mathbf{X}^2 (which can be expressed as $\mathbf{X}^1 > \mathbf{X}^2$), if and only if $\mathbf{Z}(\mathbf{X}^1)$ is partially less than $\mathbf{Z}(\mathbf{X}^2)$, i.e., if and only if $z_i(\mathbf{X}^1) \leq z_i(\mathbf{X}^2) \forall i = 1, \dots, K$ and $\exists i \in \{1, \dots, K\}$ with $z_i(\mathbf{X}^1) < z_i(\mathbf{X}^2)$. The solution space \mathbf{X} is generally restricted by M constraints, such as $g_j(\mathbf{X}^s) \leq 0$ for $j = 1, \dots, M$ bounds on the decision variables: $\mathbf{X} = \{\mathbf{X} \in R^N | g_j(\mathbf{X}^s), j = 1, \dots, M\}$. This relationship is expressed as following:

$$(\mathbf{X}^1 > \mathbf{X}^2) \leftrightarrow (\forall i)(z_i(\mathbf{X}^1) \leq z_i(\mathbf{X}^2)) \wedge (\exists i)(z_i(\mathbf{X}^1) < z_i(\mathbf{X}^2)) \quad (3.16)$$

In Equation (3.16), decision variable, \mathbf{X}^1 is said to dominate another decision variable, \mathbf{X}^2 , if and only if there is no objective value of \mathbf{X}^1 is worse than the objective value of \mathbf{X}^2 and there is at least one objective value of \mathbf{X}^1 is better than the objective value of \mathbf{X}^2 . In the following section, the background of the NSGA II and the process flow of the algorithm to obtain non-dominated solutions are addressed.

The former technique of this evolutionary algorithm is referred as NSGA I, which is introduced in 1994 (Srinivas and Deb, 1994). This evolutionary algorithm utilizes genetic algorithm in searching and determining optimal solutions. NSGA II is chosen from NSGA I as the latter technique able to solve the associated drawbacks that are found in the former

technique (Srinivas, 1994, Konak, 2006). There are three drawbacks that are encountered in the NSGA I. The first drawback is related to the high computation resources needed in performing the NSGA I. In NSGA I, intensive resources are needed to compute the sorting process in order to obtain the non-dominated solutions. This type of sorting comes with high computation complexity of $O(MN^3)$, where M is the number of objective functions and N is the number of decision variables. High computation complexity exists in the sorting process because each of the solutions needs to be compared with other solutions at the preceding position in order to examine the non-domination criteria. Therefore, any increase of the M and N values will contribute to the tremendous increase in computation complexity. On the other hand, NSGA II has lower computation complexity due to its quick non-dominated sorting technique. It only requires $O(MN^2)$ complexity as compared to $O(MN^3)$ complexity as observed in NSGA I. However, the quick non-dominated sorting technique in NSGA II demands a higher space complexity, in which NSGA II needs $O(N^2)$ memory requirement as compared to $O(N)$ from NSGA I. Despite this fact, NSGA II is chosen as the preferred approach as compared to NSGA I as it is able to shorten the running time of the process algorithm. At the same time, NSGA exhibits additional advantages as a robust optimization approach in finding optimum solutions.

The second drawback of NSGA I is related to the weakness of this approach in emphasizing the quality candidates in the searching process, where it neglects elitism. This neglect slows down the searching speed in obtaining non-dominated solutions via genetic algorithms. Due to this drawback, NSGA II has the ability to identify and retain the best possible non-dominated solution from parent to child generation. The best retention is possible in NSGA II due to its elitist-preserving algorithm in its technique.

The third drawback of NSGA I is related to user intervention in specifying parameter that assures the diversity in each population cycle in the genetic algorithm. Diversity attributes in genetic algorithm is highly important in order to ensure the Pareto optimal solution is optimum and widely spread. This parameter is known as sharing parameter and labelled as σ_{share} . In NSGA I, this sharing parameter needs to be specified, explicitly. If there is a parameter need to specified, means the success criteria for the optimal solution is also dependent to the suitability of this specified parameter value. Therefore, NSGA I is more complicated to be used as optimization approach. For NSGA II, the diversity in the population cycle is preserved without needs any user specified parameter values. With the advantages upon the three drawbacks from NSGA I, NSGA II is chosen as the optimization approach used in this thesis.

This section describes the process flow of NSGA II algorithm. Figure 3.12 illustrates this algorithm until the non-dominated solutions are achieved. In NSGA II, it is initiated with a population, $P_{t=0}$ for each objective function. This first population is generated with a random values. With these random values, the solutions are calculated, accordingly. For each calculated solution, fast sorting technique and crowding distance evaluation are used to assigns the fitness value and non-dominated level. Subsequently, offspring population, which is denoted as $Q_{t=0}$ is derived through genetic algorithm operators. The operators includes crossover, mutation and elitism rating steps. The resultant of $P_{t=0}$ and $Q_{t=0}$ are combined as $R_{t=0}$. The subsequent process is followed by ranking step into front F via non-dominated sorting technique. With reference to the sorted front F, non-dominated solution is selected from the best rank of front F, which is denoted as F_1 . The size of F_1 is compared with the size of the initial population. If the size of F_1 is smaller, size of F_1 is expanded until the size is equivalent to the size of the initial population. Otherwise, if the size of F_1 is not smaller than initial population, the above process above is repeated into

second generation, $t = t + 1$. For the next generation, the population, $P_{t=1}$ is determined from the selected non-dominated set in first generation. This process loop for each subsequent generation is continued until the pre-set maximum generation is reached. The resultant solution from the maximum generation is known as the Pareto-optimal solutions. With this Pareto-optimal solutions, their corresponding generated image from the set of solution in the objective space is referred as Pareto front. The Pareto-optimal solution refers to the non-dominated solutions that fulfills the trade-offs among all the objective functions.

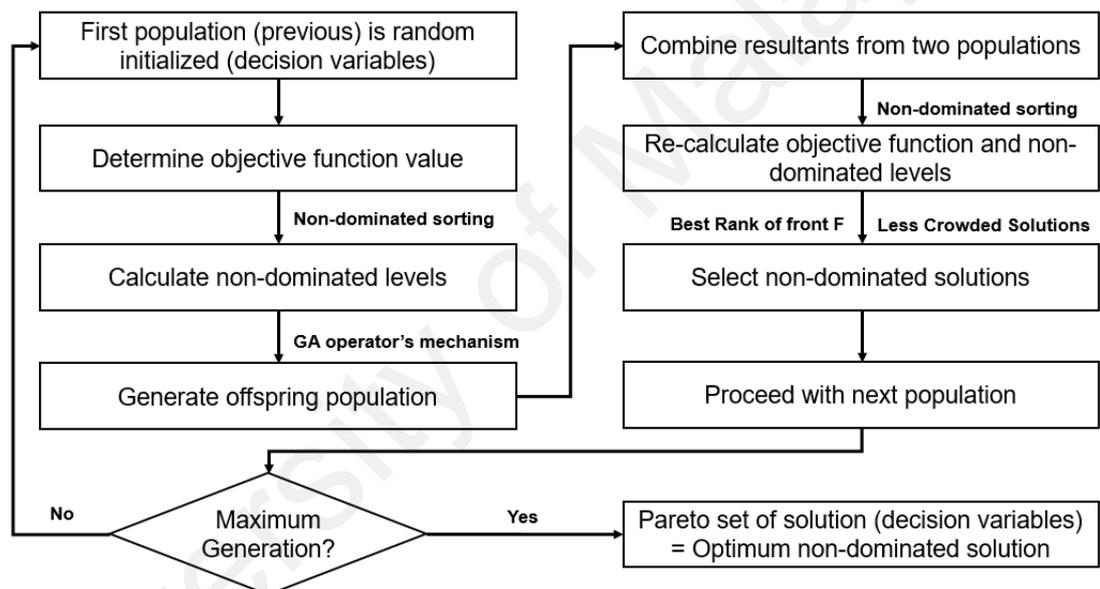


Figure 3.12: Process flow of NSGA II in obtaining non-dominated solutions

3.1.4 Selection of Optimized Solutions

Figure 3.13 illustrates the overall process in the proposed contrast enhancement framework. Figure 3.13 describes the determination of search range of the decision variables, which is the separating point and clipped limit. The search range is correspond to the three objective functions, which are Contrast, Brightness Preservation and Detail Preservation. The desired value of the decision variables for these three objective

functions are modelled using a curve fitting process, where three analytical functions for the three objective functions are computed. These three analytical functions are combined into a Cartesian space. From these analytical functions, Pareto-optimal solution or known as non-dominated solution is computed via NSGA II. At the end of the NSGA II process, selection process is carried out from the set of non-dominated solution in order to determine the desired non-dominant solution that is suitable for the enhancement of hand-bone image. This desired non-dominant solution is selected to compute enhanced version of the hand bone image via histogram equalization. The desired non-dominant solution is selected from the set of non-dominated solution that has the highest fitness value of each principal objective, in terms of preservation of contrast, brightness and detail.

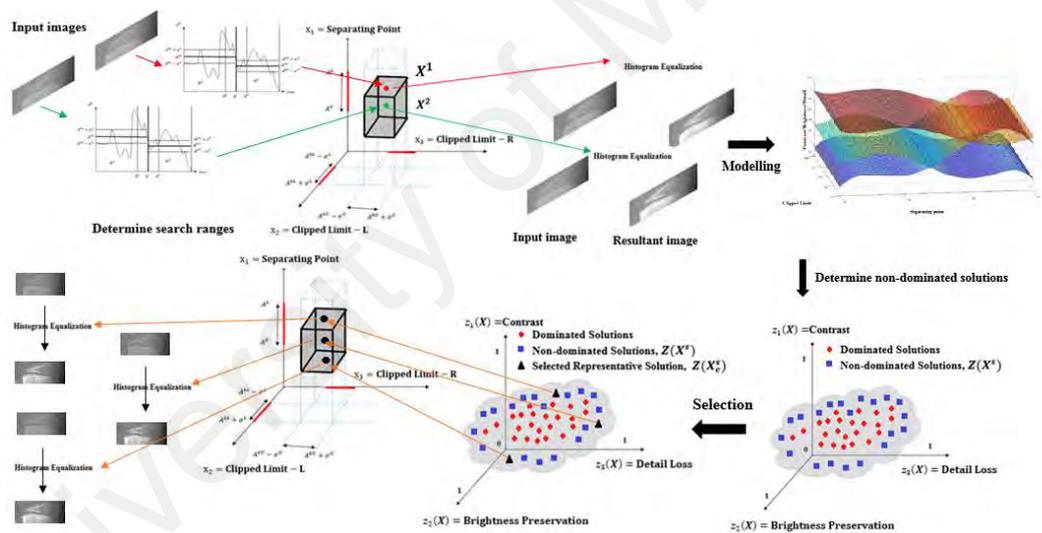


Figure 3.13: Illustration of the overall procedures including the search ranges determination, objective function modelling, multi-objective optimization solver searching for non-dominated solutions, and selection of representative solutions out of the non-dominated solutions.

3.2 Denoising of the Hand Bone

In the context of denoising process, it is important to remove or eliminate non-bone blobs after the binarization stage. Binarization will highlight the object of interest as foreground and other non-object of interest as background. However, the highlighted foreground may still contains unwanted noise that does not belong to the object of interest. Therefore, it is highly important to include a denoising step that able to eliminate anomalies in the detected region. The proposed anomalies elimination method consists of three major steps. The first step involves fitting procedure of the emphasis density distribution onto the detected region and follow by the second step that comprises of reference point computation. The third step will identify and eliminate the anomalies. The details of each step is further elaborated in the following section.

Generally, the proposed anomalies elimination in detected region from bone detection stage is a combination of two significant procedures, which is the anomalies accommodation and anomalies identification. The functionality of the anomalies accommodation process is to amplify the effect of anomalies from the object of interest, which is known as true signal. In anomalies identification process, it determines whether the detected region is anomalies or not.

Anomalies accommodation procedure is a series of steps of assigning a value to each pixel in the hand bone image. This value resembles the degree of relevancy of this pixel been emphasis as a bone blob or not. In another words, this value will categorize each pixel to bone pixel (true signal) or non-bone pixel (anomalies). In this procedure, a reference point is introduced and the computation of this point is based on the involvement of true signal and anomalies. Therefore, the intention of the anomalies accommodation is to increase the involvement of the true signal in the computation of reference point as compared to the involvement by the anomalies. In order for the

reference point has better approximation of true signal, weight distribution function known as emphasis distribution function (EDD) is introduced. EDD is designed with the assumption that the anomalies are located around the true signals. With this assumption, three steps are introduced for this purpose. The first step is devoted to the pre-selection of a continuous parametric distribution model based on hand bone properties. This is followed by the computation of the parameters that is related to the distribution model. The third step involves the discretization of the continuous distribution with value assignment on each pixel of the hand bone image.

Based on the assumption of the observation that the anomalies are commonly located at the vicinity of the true signals, a two dimensional Gaussian distribution is chosen. This type of distribution is estimated by two parameters, which are covariance and mean. In this, covariance consists of variance and correlation coefficients. A 2D Gaussian distribution is fitted into the hand bone image by overlaying this joint density function according to $f(x|\mu, \Sigma)$ for pixels in the image, where R is the pixel on the detected region and 0 is in other region other than R such that:

$$G(C, L) = \begin{cases} f(X|\mu, \Sigma) = \frac{1}{\sqrt{|\Sigma|} (2\pi)^d} e^{-0.5(X-\mu)\Sigma^{-1}(X-\mu)^T} & \text{if } (C, L) \in R \\ 0 & \text{if otherwise} \end{cases} \quad (3.17)$$

where $G(C, L)$ resembles the EDD in coordinate (C, L) of the detected region. In Equation (3.17), C and L represents column and row in R , respectively, X refers to 1-by-2 vectors, $[C \ L]$ describing C and L coordinates of the rectangle detected region, R and Σ represents covariance of vector X , and Σ is a symmetric positive definite matrix. In the example of two (2) properties of blob that are available in bone image, which referred as $P1$ and $P2$, then $X = [P1 \ P2]$ and the μ in Equation (3.17) is determined as the center

point of the region of interest. In μ , the x-coordinate of the center point is calculated as (bone width/2) and y-coordinate of the center point is measured as (bone height/2). The center point as described above is chosen as μ and this selection fulfills the assumption where the pixels near to the center point should be given a higher weight from the pixels that far from the center point, for example the side pixels in the bone image. Higher weight means higher tendency to be true signal and the anomalies are scattered around the center point. The decision in making the center point as μ is still valid for hand bone image without anomalies due to the reason that the side pixels that are away from the center point is susceptible to uneven illumination. In terms of correlation, this technique can be enhanced by adding the correlation variable, which this variable is calculated from the analysis on the directions of the blob. In this thesis, correlation variable is not added as the detected region in the hand bone image is assumed to be well-skewed already. Thus, correlations are defined as zero. In terms of variance at vertical side and horizontal side of the hand bone image, both variance value is dependent to the bone width and bone height, respectively.

This section describes the computation of the reference point vector, which considering all the blobs in the hand bone image. In this context, a reference point is a point in D-dimensional feature space that exhibits central tendency properties. This reference point serves as an approximation reference to true signal or near to true signal. This reference point may exist in any of the blobs in the hand bone image if D-dimensional properties or blob features are computed. One of the possible intuitive ways to calculate this reference point is by computing either the mean, median or mode value of all the available feature points. However, each of this computed value has its own limitation. For example, mean value will include anomalies in the final output. For median, this value does not include other data points except the median point itself in which the important data point that belongs to the detected region may be missing in the

calculation too. It is also known that mode value always fail in the scenario that has datasets with low standard deviation distribution. Although mean value maybe the best approximation to the reference point, a better reference point should be computed with the motivation that this reference point is further away from the anomalies as compared to the conventional mean value. This new reference point has the tendency to alleviate the difference between central tendency and anomalies. The following section is devoted to detail explanation of mechanism to compute EDD that is overlayed on the detected region in order to derive a new reference point that fulfills the above stated motivation.

In order to simplify the explanation of the computation of the reference point, two blobs are to be considered as compared to all blobs that exists in the hand bone image. These two blobs are referred as P1 and P2. In order to compute the reference point, one of the initial steps is performing binarization onto the hand bone image. The output image from the binarization process is having two colour pixels, which is white pixels with value equals to one and black pixels with value equals to zero. In binarization process, the term “blob” is introduced and this term means a group of connected pixels having the same properties in the detected region. Second term to introduce is $I(C, L)_B$ where subscript, B describes the blob index in which $B \in [1, 2, 3 \dots, total\ blobs]$, C and L depict column or x-coordinate and row or y-coordinate in detected region, R , respectively. For example, $I(23, 44)_2$ depicts the pixel value in image coordinate (23, 44) and this pixel belongs to second blob; if that particular pixel is a white pixel, then $I(23, 44)_2 = 1$ or otherwise, $I(23, 44)_2 = 0$.

$$\text{Reference point} = E[E[G_{B, P1, P2}]]$$

in which

$$\begin{aligned}
& E[E[G_{B,P_1,P_2}]] \\
&= \frac{1}{(T_1)(T_2)} \sum_{B=1}^{T_2} (P_{1_B} \\
&+ P_{2_B}) \sum_{C=x_{1_B}}^{C=x_{2_B}} \sum_{L=y_{1_B}}^{L=y_{2_B}} I(C, L)_B G_{B,P_1,P_2}(C, L)
\end{aligned} \tag{3.18}$$

in which,

T_1 = total density on all white pixels =

$$\sum_{B=1}^{T_2} \sum_{C=x_{1_B}}^{C=x_{2_B}} \sum_{L=y_{1_B}}^{L=y_{2_B}} I(C, L)_B G_{B,P_1,P_2}(C, L)$$

$$G_{B,P_1,P_2}(C, L) = (P_{1_B} + P_{2_B}) [I(C, L)_B][G(C, L)]$$

$$I(C, L)_B \begin{cases} 1 & \text{if } I(C, L) \text{ is a white pixel} \\ 0 & \text{otherwise} \end{cases}$$

S = total number of blobs

It can be further generalized as:

$$E[E[G_{B,P(i)}]] = \frac{1}{(T_1)(S)} \sum_{B=1}^{T_2} (P(i)_B) \sum_{C=x_{1_B}}^{C=x_{2_B}} \sum_{L=y_{1_B}}^{L=y_{2_B}} I(C, L)_B G_{B,P(i)}(C, L) \tag{3.19}$$

where x_{1_B} , x_{2_B} , y_{1_B} , y_{2_B} represent the coordinates of the bounded rectangle of the particular blob B ; $P(i)_B$ represents the i^{th} feature of B^{th} blob in which $B \in [1, 2, 3, \dots, S]$.

The last procedure of anomalies identification intended to identify blobs that are categorized as anomalies. Anomalies identification is a challenging procedure. This is due to characteristic of the anomalies where their existence, number of anomalies and type of anomalies is unknown. The unknown status of the anomaly existence is challenging to certain extent. If the existence is known, a proper identification algorithm is designed to allocate them only. If the existence is unknown, the identification algorithm can be complicated. Similarly, this procedure is less complicated if the number of anomalies is known, where the fix number of unlikely-to-be-bone can be identify. By knowing the type of anomalies, whether the anomalies are extreme anomalies (exists gap

difference between feature value extracted from the anomalies and bone blob) or bone-resembled anomalies gives a very important clue for the design of effective identification algorithm. With the presence of all these unknown about anomalies, identification of anomalies turns into a non-trivial problem. In this thesis, a framework that deal with these unknown in detection anomalies is proposed.

With the reference to the early computed reference point, hierarchical anomalies identification framework is designed in which the extreme anomalies will be first eliminated prior to the removal of bone-resembled anomalies. This hierarchical framework is also effective for bone image without anomalies as actual bone blob has a gap difference from the extreme anomalies.

The anomalies identification method is proposed based on Chebyshev's inequality. Once the reference point is determined, the discrepancy between each two-dimensional feature point from the blob (P_B^1, P_B^2) (of each B^{th} blob) and reference point (r_{p1}, r_{p2}) in the vector space is calculated. Let D_B denotes the real-valued random variable that depicts the calculated discrepancy in Euclidean distance, which subscripts $B \in [1,2,3, \dots \text{Total blob } (S)]$. The computed Euclidean distance between each feature vector and reference point is depicted in Equation (3.20):

$$D_B = \sqrt{(|P_B^1 - r_{p1}|)(|P_B^2 - r_{p2}|)}$$
(3.20)

where $B \in [1,2,3, \dots \text{Total blob}]$

Equation (3.21) illustrates the discrepancy in Euclidean distance between N-dimensional feature point, $(P_B^1, P_B^2, P_B^3, \dots, P_B^N)$ (of each B^{th} blob) and reference point $(r_{p1}, r_{p2}, r_{p3}, \dots, r_{pN})$ in the vector space.

$$D_B = \sqrt{\prod_{n=1}^{n=N} (|P_B^n - r_{pn}|)} \quad (3.21)$$

where $B \in [1,2,3, \dots Total\ blob]$

After the discrepancy in Euclidean distance is computed, a measure of spread $M(D)$ and expected value of D , $E(D)$ are calculated based on the following equations.

$$E(D) = \frac{1}{S} \sum_{B=1}^{B=S} (D_B) \quad (3.22)$$

$$M(D) = \sqrt{\frac{1}{S} \sum_{B=1}^{B=S} (D_B - E(D))^2} \quad (3.23)$$

Further analysis is done on the computed measure of spread $M(D)$ and expected value of D , $E(D)$ in order to eliminate different type of anomalies. Extreme anomalies that have big gap difference from true signal in the detected region will be eliminated by the following inequality:

$$\begin{aligned} P[D_B \geq k] &= P[D_B \geq [E(D) + |k - E(D)|]] \\ &\leq \frac{M(D)^2}{M(D)^2 + |k - E(D)|^2} \end{aligned} \quad (3.24)$$

One of the constraints in Chebyshev's inequality is not assume any underlying probability distribution of the data. As the number of discrepancies is usually appears in small number, law of large number might not be applicable as a suitable estimation to assume normal distribution on discrepancies. Decision rule is defined with C percent confidence interval as follows:

$$\text{Reject if } P[D_B \geq k] \leq (1 - \frac{c}{100}) \quad (3.25)$$

In this thesis, all the features from the hand bone image will be considered simultaneously in order to detect an accurate anomalies. If each feature is considered individually, the accuracy of identifying anomalies maybe drop as the anomaly may not be distinguish able by certain specific feature. For example, if a single height feature height is used to differentiate bone blob and anomalies, the detection will be fail as the height of the anomalies is similar to actual bone blob. Therefore, discrepancy of all the relevant features should be considered in multivariate setting in order to have higher successful rate in detecting anomalies. For example, even the height of some anomalies blob resembles the true bone blob, the width of the anomalies blob might be different from true bone blob. With this, anomalies is successfully detected as outlier to the hand bone image. From empirical observation, good features for anomalies identification consists of confidence level of blob recognition by classifier, texture of blobs and shape properties of blobs. These features demonstrate high discriminative ability in practical implementation. Therefore, relevant multiple features and simultaneous consideration is an important strategy to ensure the successful of the algorithm in detecting anomalies. In the feature space, each point in this space resembles the feature vector that is obtained from each blob in the detected region.

One of the most promising techniques in identifying anomalies with multiple good features is multivariate feature vector approach. In this thesis, Mahalanobis distance is chosen as the multivariate feature vector approached to be used for this purpose. Mahalanobis distance is a significant distance metric in statistical field been introduced since 1936 by P.C. Mahalanobis. In term of considering the correlation and difference between multiple features, Mahalanobis distance shows better consideration as compared to Euclidean distance. In this section, Mahalanobis distance, D_M with T dimensional

feature vector = $[P^1, P^2, P^3, \dots, P^T]$ in which is blob properties feature, the mean vector for the features vector $\bar{Z} = [\bar{Z}_1, \bar{Z}_2, \bar{Z}_3, \dots, \bar{Z}_T]$, and covariance depicted as Σ , is defined as

$$D_M(i) = \sqrt{(Y_i - \bar{Z})^T \Sigma^{-1} (Y_i - \bar{Z})} \quad (3.26)$$

$$\text{where } \Sigma = \frac{1}{(N-1)} (Z - \bar{Z})^T (Z - \bar{Z})$$

$D_M(i)$ represents the Mahalanobis distance of i^{th} blob features vector in the distribution of features vector data Z . Let the features vector is a T -dimensional vector and the number of blobs in the detected region represented by $T2$. Meanwhile, Σ^{-1} is a $T \times T$ matrix depicting the inverse covariance of data, Y_i is a $1 \times T$ matrix corresponding to features value of i^{th} blob in the detected region. \bar{Z} is a $1 \times T$ matrix as defined in Equation (3.27).

$$\begin{aligned} \bar{Z} &= [E[E[G_{B,P(1)}]], E[E[G_{B,P(2)}]], E[E[G_{B,P(3)}]], \dots, E[E[G_{B,P(T)}]]] \end{aligned} \quad (3.27)$$

As similar to the reconstruction of reference point concept, the influence of anomalies should be reduced in the covariance, Σ if the anomalies are considered during the forming of the Mahalanobis distance. Σ is defined by the following equation:

$$\Sigma = \frac{1}{(N-1)} (Z - \bar{Z})^T (Z - \bar{Z})$$

where $Z \in \mathbb{R}^{T-by-T2}$

$$\Sigma_{P^{(i)}P^{(i)}}^W = \frac{1}{(N-1)} (Z_{P^{(i)}B}^W - \bar{Z}_{P^{(i)}}^W)^T (Z_{P^{(i)}B}^W - \bar{Z}_{P^{(i)}}^W) \quad (3.28)$$

where $\sum_{P(i)P(i)}^W \in \mathbb{R}^{1-by-1}$, $\sum^W \in \mathbb{R}^{T-by-T}$

$$W_B = \frac{1}{(X_B^2 - X_B^1)(Y_B^2 - Y_B^1)} \sum_{C=X_B^1}^{X_B^2} \sum_{L=Y_B^1}^{Y_B^2} G(C, L)$$

$$N_{W_B} = \frac{W_B}{\sum_{B=1}^S W_B}$$

so that

$$\sum_{B=1}^S N_{W_B} = 1$$

where S = total blob, $X_B^2 X_B^1 Y_B^2 Y_B^1$

$$\overline{Z_P^W} = [E[E[G_{B,P(i)}]]] \text{ as defined in Equation (3.19)}$$

in which

$$Z_{P(i)B}^W = N_{W_B} Z_{P(i)B}$$

for $i = 1, 2, 3, \dots, T$, $B = 1, 2, 3, \dots, T2$

where $Z_{P(i)B}^W$ depicts element in matrix Z^W in index i row and index B column,

the subscripts W is an indicator representing that it is a weighted Z

\therefore Equation (3.26) is modified to

$$D_M^W(i) = \sqrt{(Y_i - \overline{Z^W})^T \Sigma^{W-1} (Y_i - \overline{Z^W})}$$

(3.29)

$D_M^W(i)$ shows the Mahalanobis distance of i^{th} blob features vector in the distribution of weighted features vector data Z^W for $i = 1, 2, 3, \dots, T2$.

Equation (3.28) depicts the elements of the covariance matrix in which they form the covariance matrix, Σ^W . Next, an inverse of this covariance matrix is calculated in order to define a new distribution of Mahalanobis distance for each blob's feature vector, Y_i , with regards to this distribution center. This new defined distribution of Mahalanobis distance is proven to less effect from the anomalies feature vector. Less effect is due to the weight of each feature vector is represented by corresponding EDD as defined in Equation (3.19).

Finally, the anomalies are determined by comparing $D_M^W(i)$ of each i^{th} blob feature vector in Equation (3.29) with a calculated threshold. This threshold is calculated with respects to the standard deviation of the $D_M^W(i)$ for all blob feature vector:

$$S_M^W = \frac{1}{N-1} \sum_{i=1}^{T2} [D_M^W(i) - \frac{1}{N} \sum_{i=1}^{T2} D_M^W(i)]^2 \quad (3.30)$$

The anomalies identification decision rule is defined in Equation (3.31).

$$i - th \ blob = \begin{cases} Anomaly & \text{if } D_M^W(i) > K(S_M^W) \\ Signal & \text{otherwise} \end{cases} \quad (3.31)$$

K in Equation (3.31) serves as a constant regulator in order to manipulate the tolerance level in anomalies elimination procedure. Higher K value means has higher tolerance to anomalies and higher possibility of identifying anomaly as true signal. On the other hand, lower K value resembles intolerance to anomalies and higher possibility of identifying true signal as anomaly.

3.3 Segmenting the Hand Bone

In this thesis, fully convolutional neural network (FCN) is chosen as deep learning model for hand bone segmentation. FCN-8 is chosen as specific network type of FCN to be used for segmentation purpose. The focus segment is at the carpal bones. Current bone age assessment is commonly carried out through qualitative visual inspection. Due to the limitation of this manual visual inspection, automated bone age assessment system is in active research. One of the procedures that impacts the accuracy of the bone age assessment system is the accuracy of the segmented bone structure in the hand bone image. A good segmentation output provides physician with a better visual representation for bone age determination. Due to the importance of accurate bone segmentation for bone age assessment system, the applicability of deep learning neural network model is demonstrated with hands and wrist radiograph image. This radiograph image consists of both gender, which is female and male and their age is ranging from zero to 18. The segmentation technique is applied to entire hand bone image in the database and its accuracy is illustrated in the result session.

The training and testing dataset used to compute the segmentation results are obtained from open database of hand bone radiograph image. The radiograph images are collected from Children's Hospital Los Angeles with effort of University of Southern California. This collection is funded by National Institutes of Health (NIH). This training and testing utilizes 348 left hand and wrist radiograph images. Selected hand bone images is extracted from this database according to their suitability to the objective of segmentation. From 348 images, 290 images is selected as training set, while the remaining 58 images are reserved for testing and evaluation purpose. The adopted images are belongs to normal children together with patients' demographic data and radiologists' reading. This data is distributed to 19 age groups that include newborn and age 1 to age 18. It also covers two genders, male and female as well as four ethnicity group, which is Caucasian, African-

American, Hispanic and Asian. The distribution of number of hand bone image for different ethnicity group and gender is depicted in Appendix A. This database contains 1390 hand bone images, various from clear to blur quality and the available format is DICOM and JPEG. In this research work, only JPEG image is used. The resolution of the JPEG format is various from approximate 928 x 1218 to 1904 x 2653. This database is generated for research and academic use and it is available for download (Joseph, L., 2005).

Prior to the dataset training using FCN model, data preparation need to be carried out. This preparation consists of two main steps in order to setup the relevant dataset for training purpose, which comprises of extraction of the region of interests of the bone structures in the carpal bone and removal of unwanted object from the radiograph image. A masking process is carried out on the original image in order to get nine classes, where the nine classes include eight (8) bone structure of carpal bone and one (1) background. In the masking process, the region of interest of each bone structure in the carpal bone is assigned to different colour that each unique colour is corresponds to each respective class of the bone structure and background. Figure 3.14 illustrates the distribution of age and gender in the training dataset. In this figure, the number of training sample for male gender and female gender in each age category is quite balance. It is also can be noticed that the training sample for newborn is limited due to the hand bone image acquisition for newborn is challenging and inconvenient.

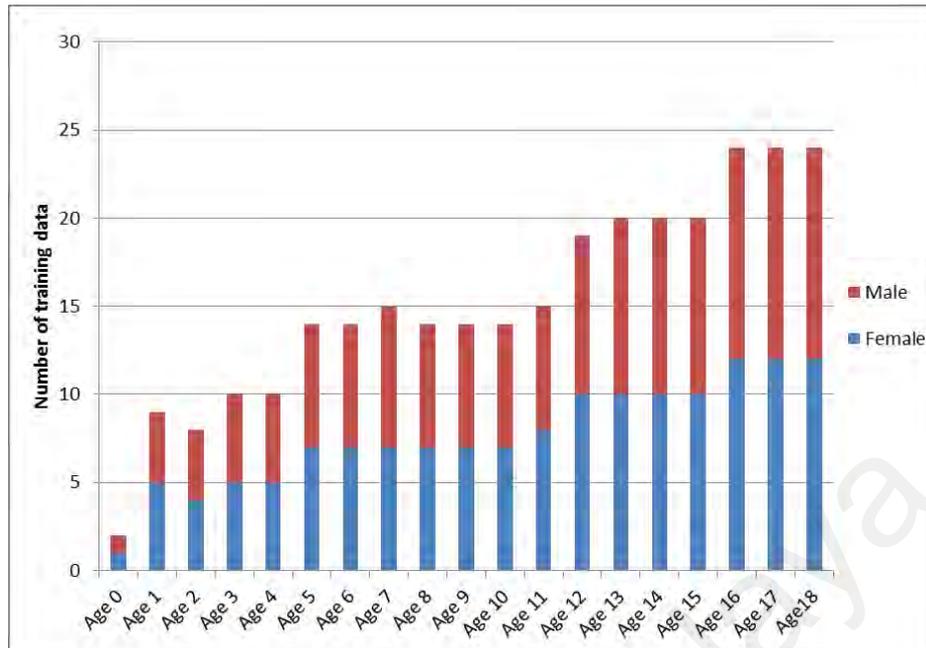


Figure 3.14: Age and gender distribution in the selected training dataset

Left image in Figure 3.15 depicts the original radiographic image used in the training and right image Figure 3.15 illustrates the labelled image. The labelled image shows the region of interest of the nine classes in different colour that corresponding to respective bone structures in the carpal bone and the background of the hand bone image.

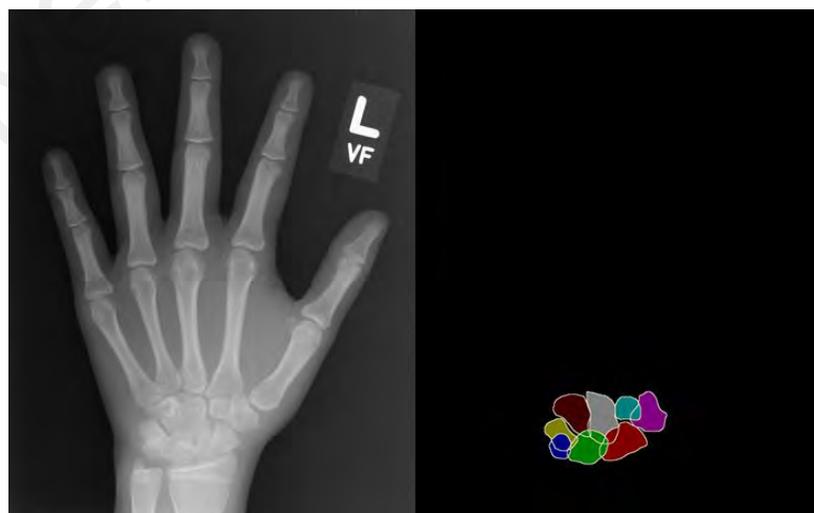


Figure 3.15: Original hand bone radiographic image (left image) and the labelled image (right image).

The following section describes the training architecture and training environment of FCN model in performing segmentation of carpal bone in the hand bone image. FCN is derived from one of the classification architectures, which known as VGG-16. VGG-16 that comprises of 16 convolutional layers and 3 fully connected layers is enhanced into convolutional layers with one kernel size, which resulting into FCN. In FCN, the convolutional layers is a fully connected layers and the process in computing the weighted sum is different from VGG-16. FCN computes the weighted sum is at per receptive field while the weighted sum computation of VGG-16 is happened at per image (Long, Shelhamer et al., 2015). The motivation of FCN in computing the weighted sum at per receptive field is to conduct the abstraction process that not only consider the extracted features at the labelled region of interest but also with its location information. In this thesis, FCN-8 is chosen from the other two versions of FCN, which is FCN-16 and FCN-32. This is due to FCN-8 exhibits ability to process various sizes of input images, where prediction happened at small scales of 8x8 kernels. Figure 3.16 illustrates the network architecture of FCN-8. In this figure, FCN-8 consists of 19 convolutional layers and performs dense predictions. Due to the characteristic of dense prediction process, loss of information is unavoidable throughout the training process. Thus, skip connections are utilized by adding up all the final prediction layers in FCN-8 with the earlier layers. By doing this, the loss information is able to be retrieved back. From Figure 3.16, the output from conv7 is upsampled with a stride 2 transposed convolution. The upsampled result is summed up with pool4. The following skip connection process is to upsampled them with a stride 2 transposed convolution, and followed by summing them with pool3. The final segmentation map is obtained by applying a transposed convolution layer with stride 8 on the resulting feature map.

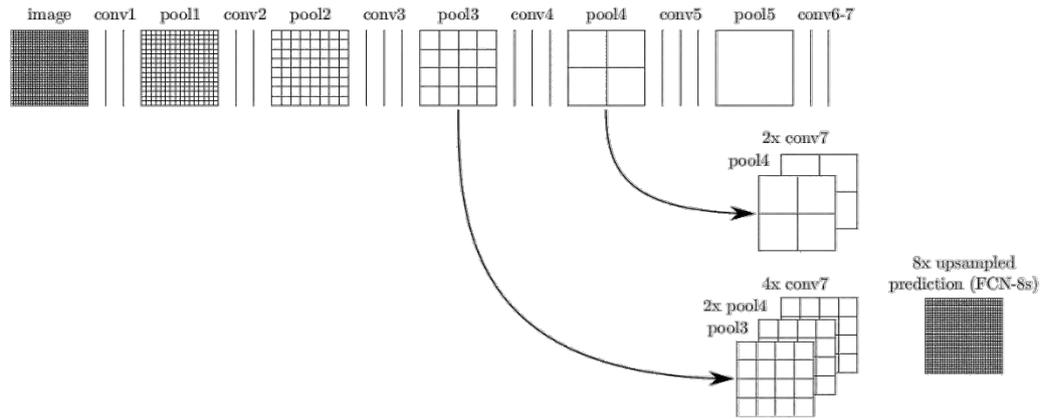


Figure 3.16: Network architecture of FCN-8

In this thesis, a transfer learning technique using FCN-8 heavy pascal caffemodel (Jia, Shelhamer et al., 2014) is employed in obtaining segmentation model. Transfer learning technique is used due to the limited labelled hand bone image. The chosen number of epoch for carrying out the training process is 50 and base learning rate is set at 0.0001. The training process of FCN-8 is executed via Nvidia DIGITS 6.0 in which is a Deep Learning GPU Training System. Nvidia DIGITS 6.0 is running on GPU machine with the specification of Intel(R) Xeon (R) CPU E5-26200, 2.00 GHz processing speed and GPU Nvidia card of GTX 1080Ti 11GB.

For testing and evaluation of the resultant model from the FCN-8 model training using 290 labelled image with transfer learning technique, the segmentation accuracy on 58 labeled image is obtained by comparing the segmentation result with the ground truth image. The segmented contours of the resultant image of FCN-8 segmentation model and ground truth image is compared using two statistical validation metric. The two metrics are Dice similarity coefficient (DSC) and Hausdorff distance (HD). The output of these two validation metrics are important in order to gauge the effectiveness of the trained FCN-8 model in performing segmentation on carpal bone of the hand bone image. The

computed validation metric value indicates the similarity value of the segmented result from the FCN-8 segmentation model with the segmented outlines from the ground truth image.

The DSC measures the spatial overlap index value between region of interest in labelled image and the region of interest in ground truth image. The spatial overlap index indicates the overlapping area of the two regions in terms of location and region size. The DSC is formulated as:

$$DSC = \frac{2(R_A \cap R_B)}{R_A + R_B} \quad (3.32)$$

where R_A and R_B resemble the region of interest of image A and image B, respectively. DSC metric value falls in the range between zero to one. Zero value depicts the condition of total no area overlapping between two regions. Meanwhile, DSC value of one resembles 100 percent overlapping match between the two regions.

On the other hand, HD measures the mutual proximity between two regions, where the proximity value is indicated by the maximum distance between any point in the first region with any point in the second region. The HD metric is formulated as:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (3.33)$$

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (3.34)$$

where A and B depicts two set of contour points in image A and image B, respectively. The overall workflow of the proposed segmentation approach in outlining the region of bone structure in carpal bone in hand bone image is illustrated in Figure 3.17. In this figure, the evaluation of FCN-8 model by testing the segmentation accuracy with the ground truth image is implemented in MATLAB (vR2014a, Mathworks, Natick, USA)

with processing machine that runs on an Intel(R) Xeon (R) CPU E5-26200 @ 2.00 GHz processing speed.

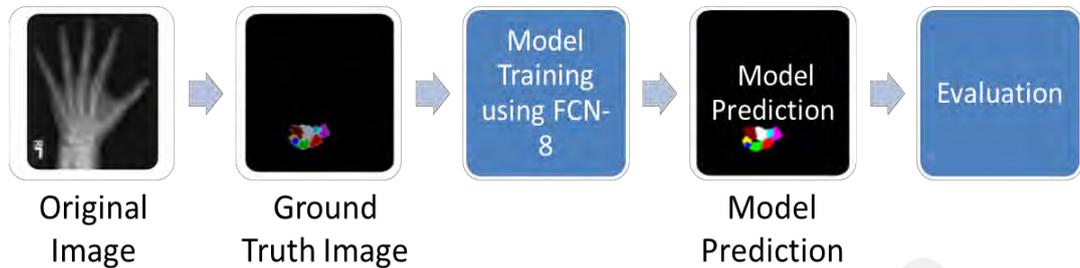


Figure 3.17: Workflow for image segmentation of carpal bones.

In the training the proposed FCN-8 model stage, the computed model for each epoch will be tested by validation dataset. For this training, one epoch equals to one full training cycle where all the provided training dataset had gone through the training for optimization. Number of epoch is set to standard value that is equals to 50. For each training cycle, the training lost rate, the accuracy rate, and validation loss rate are computed and their values are illustrated in Figure 3.18. Based on the GPU machine with the stated specification at the previous section, the training time of FCC-8 model is recorded at 175 minutes by completing 50 epochs of training and the total training images is 290.

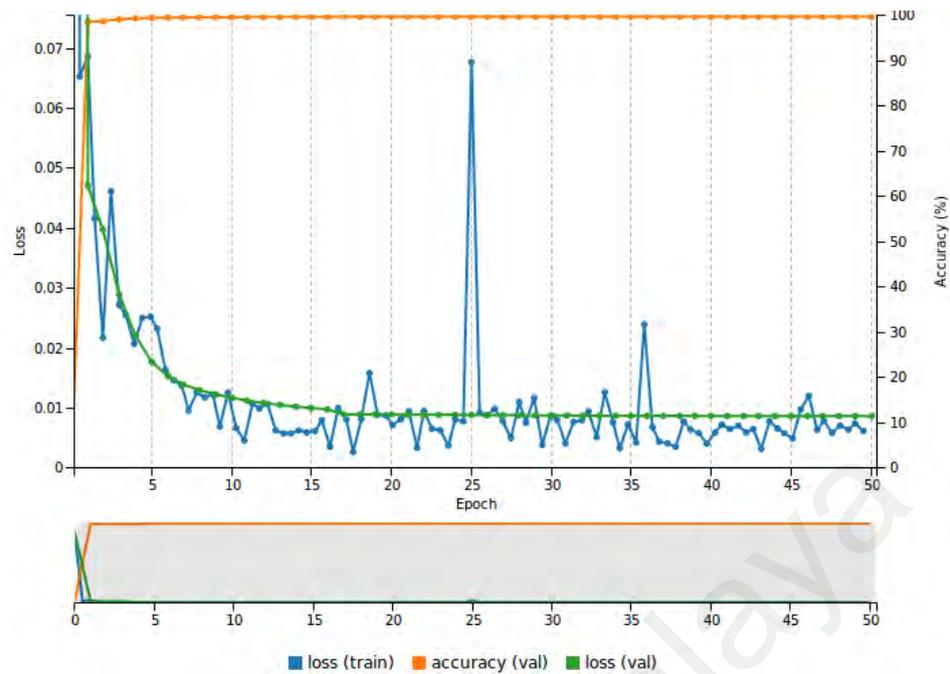


Figure 3.18: Training loss rate (blue), accuracy rate (orange) and validation loss rate (green) with respect to epochs.

Figure 3.19 illustrates the variation of the learning rate been used and set during the training stage of FCN-8 model at each epoch. In order to obtain the most optimal model from the training, it is always a good practice to control and adjust learning rate at different epoch. DIGITS allows the flexibility in adjusting the learning rate during the training. By adjusting the learning rate, the time taken for the training to achieve optimum solution is shorter as the weights of the deep learning model are optimized better without falls into the local minima of optimization. As illustrated in Figure 3.18 and Figure 3.19, it is observed that the training model is optimized and ready to be stopped from further training at epoch beginning from 17. In Figure 3.18, the performance of trained model is already at the optimum when the training reached at 17 epoch. From epoch of 17 onwards, the accuracy rate and validation loss rate is already stagnant and reach plateau behaviour.

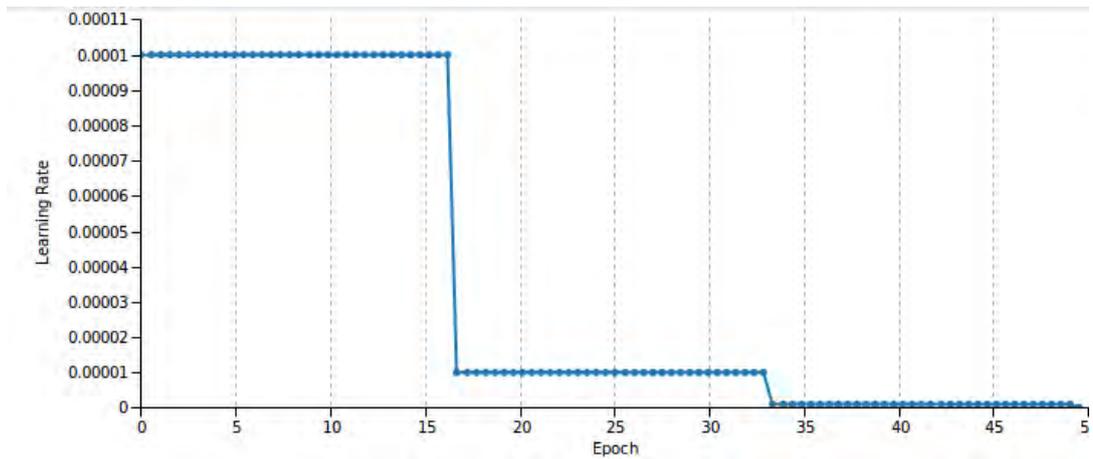


Figure 3.19: Learning Rate variation during the training of FCN-8 model.

Figure 3.20 illustrate the results of the carpal bone segmentation using the trained model of FCN-8. For each segmented result, DIGITS display each carpal bone with a different colour based on its corresponding segmented classes. In Figure 3.20, the outline of the carpal bone will be highlighted if the particular carpal bone exists in the hand bone radiographic image.

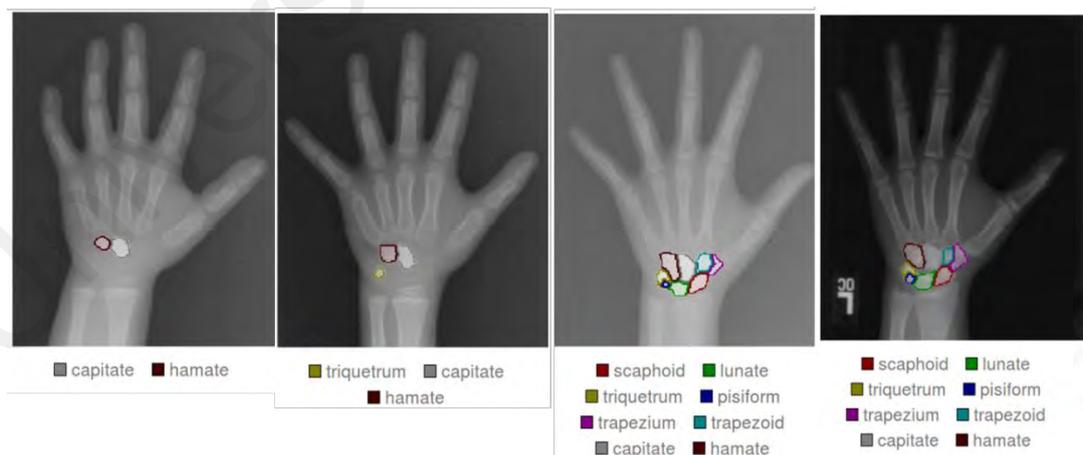


Figure 3.20: Segmented carpal bone using the trained FCN-8 model

From the DSC and HD metric values from 58 labelled hand bone image that is used for testing, the mean of DSC and HD with its corresponding standard deviation are calculated and tabulated in Table 3.2.

Table 3.2 Validation of segmentation accuracy as measured using DSC and HD

Number of images	DSC (mean \pm SD)	HD (mean \pm SD)
58	0.78 \pm 0.06	1.56 \pm 0.30 mm

From the evaluation results on the testing dataset using the trained FCN-8 model, the model performance in carpal segmentation with regards to the age is accessed. The DSC and HD accuracy results are compared with respect to age since the visualization of carpal bones quantity varies with the age variation. It is known that the difficulty level for model training increases with the increasing number of feature class and parameter.

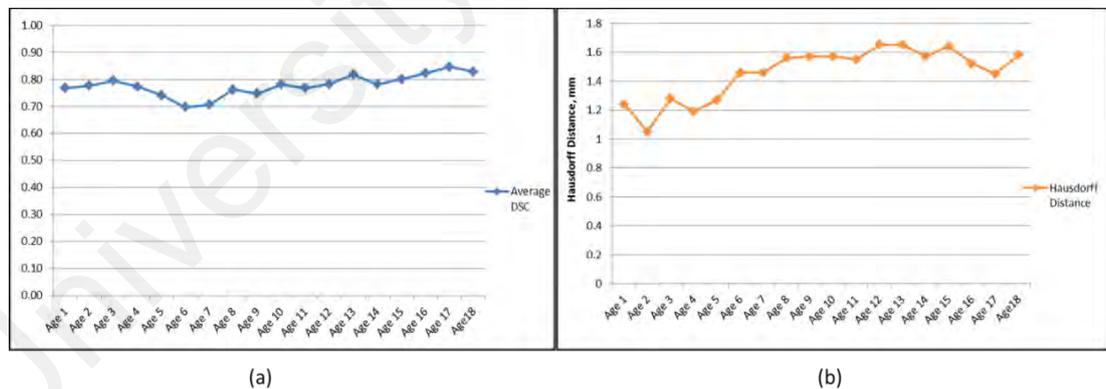


Figure 3.21 Comparison of (a) DSC and (b) HD with respect to age.

Thus, it is make sense that the accuracy of the training performance is dropped with the increasing of age. As the age increase, the number of carpal bones in the hand bone image will increase. Figure 3.21(a) and Figure 3.21(b) depicts the DSC and HD values with respect to age, respectively. In Figure 3.21(a), the highest value of DSC metric is

found at the age of 12 and the corresponding DSC value is recorded as 0.85. Meanwhile, the lowest value of DSC metric is found at the age of 6 and the corresponding DSC value is recorded as 0.67. In DSC metric, the higher the value then the better the network performance is. The result illustrated in Figure 3.21(a) has indicate that the network performance is improved as the age increases. This improve performance may due to the imbalance number of training data for age below of 11, which is only had 15 images as compared to the 20 images for the age ranges from 12 to 18. The information related to the number of training data is depicted in Figure 3.14. Therefore, the number of images used for training plays an important role in determining the performance of the network model.

In the future enhancement, it is important to ensure that the number of training and testing images is almost equal in order to obtain consistency in the segmentation performance. Besides, there is also an observation that the HD metric value is higher when the age is increasing as illustrated in Figure 3.21(b). Higher HD metric values means the maximal distance is between the two regions is bigger, which indicating the network performance is dropping. Network performance in both Figure 3.21(a) and Figure 3.21(b) is not consistent as the network performance increases with increasing age for DSC metric. For HD metric value, the network performance is opposite to DSC metric performance. The inconsistency of the HD metric performance is due to its measurement is not based on the coefficient with respect to the total contour. Thus, a bigger contour derives a higher HD value. Thus, additional evaluation metric is recommended in order to improve the evaluation method. Possible evaluation method is Intersection over Union (IoU).

CHAPTER 4: RESULTS AND ANALYSIS

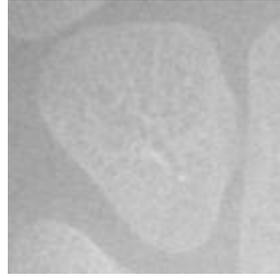
In this chapter, quantitative and qualitative evaluation of the enhancement result are established. These two evaluation results are presented and discussed, individually. In terms of qualitative result, visual quality of the resultant image from the proposed enhancement framework is illustrated. For this section, visual analysis is described in order to elaborate the significance of the enhancement framework in highlighting the pertinent information in the carpal bone image for better judgement on the bone age and at the same time, suppressing unwanted noises that impacting the manual and automatic bone age assessment techniques. In addition, the visual analysis of the resultant image also includes the explanation of specific resultant image is preferable in the bone age assessment as this enhancement produces unique characteristics according to assessment standard.

In terms of quantitative result, the significance of the enhancement output with optimization is justified by evaluating the accuracy of the classification result of the maturity stage via automated recognition classifier. The enhancement output contains the pertinent information that has been emphasized from the enhancement process with optimization. The automated recognition classifier is trained using convolutional neural network with standard hyper parameter values. The chosen database that is used in the training and testing is extracted from hand bone online database (Joseph, L., 2005). In this database, the hand bone images comprise from both genders in four populations. The four population consists of Caucasian, African American, Hispanic, and Asian. The age of children ranges from 0 to 18 years old. This database is collected from Children's Hospital Los Angeles (CHLA). In this quantitative evaluation, three types of input image set that is generated from the same set of images is used in the automated recognition classifier for maturity stage. The first type of input image is the original image, while the

second type of input image is the output of the proposed enhancement framework without optimization, which is the random image with clipped bi-histogram equalization. The second type of input images, which are generated from random value of separating point and clipped limit represents all the possible outcome for all the approaches that is discussed in Section 2.3.1 (literature review of global histogram equalization). In this quantitative evaluation, there is no comparison with specific approach of global histogram equalization due to two reasons. Firstly, it is not comparable as the proposed contrast enhancement uses Pareto optimization that will generate multiple enhanced solutions, while other methods usually generate only one enhanced solutions. Secondly, other methods focus only one of the preservation factors, while the proposed enhancement framework considers the three preservation factors in terms of contrast, brightness and detail. Therefore, direct comparison is not meaningful and invalid. The third type consists of enhanced image with optimization. The enhanced image with optimization is Pareto optimized selected input image with clipped bi-histogram equalization. Based on the final evaluation result, the accuracy of the recognition is higher if the input image are those that have been enhanced with optimization as compared to enhanced image without optimization.

4.1 Qualitative Analysis on Visual Appearance of the Resultant Images

This session is devoted to illustration of resultant image after deploying the proposed enhancement framework with optimization. With this illustration, the quality of the visual appearance of the resultant image is analysed and discussed. For example, the following figures depicts the image of the carpal bones that has been processed using the proposed enhancement approach with optimization. This figure shows the set of Pareto Front or the Pareto optimized bi-histogram equalized resultant image of the Hamate structure in the carpal bone. Each resultant image is an output of the enhancement with the two corresponding decision variable values and its three targeted objective function values. The two corresponding decision variables are annotated with x_1 and x_2 , which is the separating point and clipped limit, respectively. The three objective values are referred as C, B and D, which is the contrast objective, mean brightness preservation and detail preservation, respectively. The illustration is followed by detail discussion and analysis on the quality of the visual appearance of the image in order to measure the effect of the proposed approach in deriving the set of Pareto images.



Hamate Bone Image (Before Enhancement)

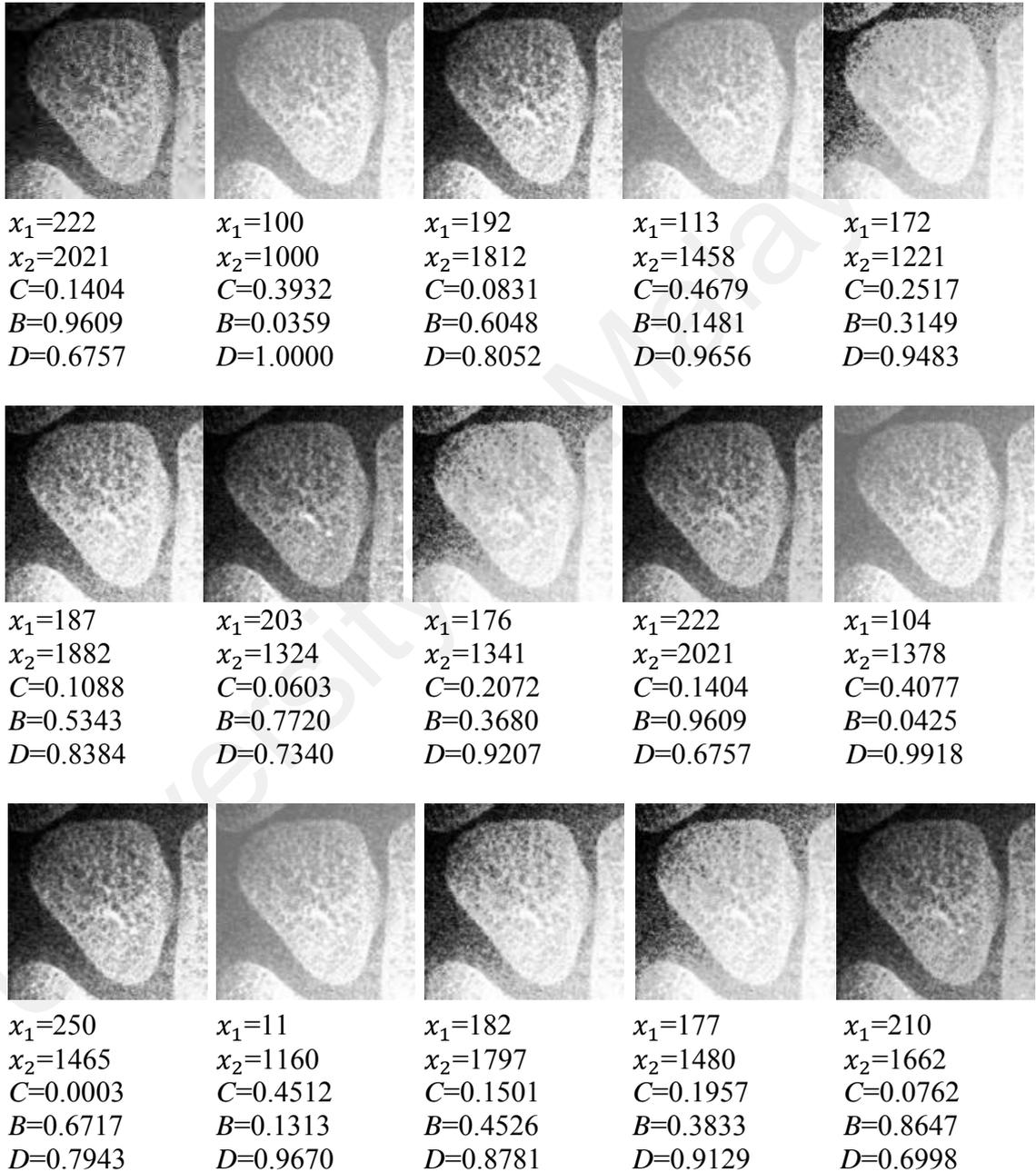


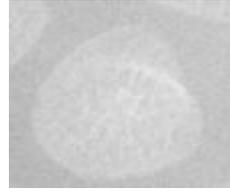
Figure 4.1: Output set of Pareto optimized bi-histogram equalized resultant images of the Hamate bone with its corresponding decision variables and each objective function values, in which x_1 represents the separating point, x_2 represents the clipped limit, C represents Contrast objective, B represents the mean brightness preservation and D represents the detail preservation, respectively.

Figure 4.1 illustrates Pareto optimized resultant images of Hamate bone. One of the measurement attribute in maturity assessment is the existence of calcium deposit or multiple deposits at the center position in the Hamate bone. This attribute is clearly highlighted in the shown Pareto optimized image as compared to the visibility of the similar attribute in the original image before enhancement of Hamate bone. In maturity assessment, the border attribute of the bone structure also represents the growth degree of a bone. The sharpness, smoothness and continuity of the border line are the maturity attributes shown in the enhanced image as compared to the original image before enhancement. In the enhancement image, the clarity of the border ensures the diameter measurement is successfully done, accurately. In bone age scoring system, the maturity stage is classified as stage D if the measured diameter is half or more than the width of the radial metaphysis. Apart of the diameter measurement, the border shape is also contributing to the decision of the maturity stage. In Hamate bone, stage D is clearly highlighted if the clear outline of the border is displaying the shape similar to character “D” in which the straight line appears diagonally to the long axis of the hand. Accurate classified maturity stage is highly important for the final maturity score. For example in Hamate bone, stage D carries score value of 79 and if it is classified wrongly to stage C due to unclear border, then the maturity score only has value of 75. Thus, the contribution of resultant image from the proposed enhancement approach is highly critical in measuring the accurate maturity stage.

The clarity of the border in Hamate bone image is crucial for the classification of higher maturity stage, such as stage E, F and stage above. The border of Hamate bone that is adjacent to capitate bone is used to determine the presence of stage E. This stage is reflected by the border formation with a slight bulge of half to two-third of the border and its flattened edges. Stage F is observed by appearance of concavity along the border that is near to triquetral bone. The subsequent higher stage is dictated by the existence of white

lines that are thickening along the distal border of Hamate bone and at the ulnar border of the hook, which is adjacent to metacarpals IV and metacarpals V. All the presence of thickening white lines at specific location in the Hamate bone is the pertinent features to be observed for maturity stage assessment. In Figure 4.1, the resultant image from the optimized Pareto is successfully observed in order to assist in the rating of maturity stage. This is due to the important cue of the shape characteristic and the contrast of the visual appearance of the Hamate bone are emphasized in the resultant image after undergone the proposed enhancement approach.

In Figure 4.1, selective images from the Pareto optimized resultant images of Hamate bone is selected to be the input images to the segmentation process, where this process is executed using deep learning approach. The images are selected based on their most optimum value for each objective values: C, B and D, where their value is near to 1.



Lunate Bone Image (Before Enhancement)

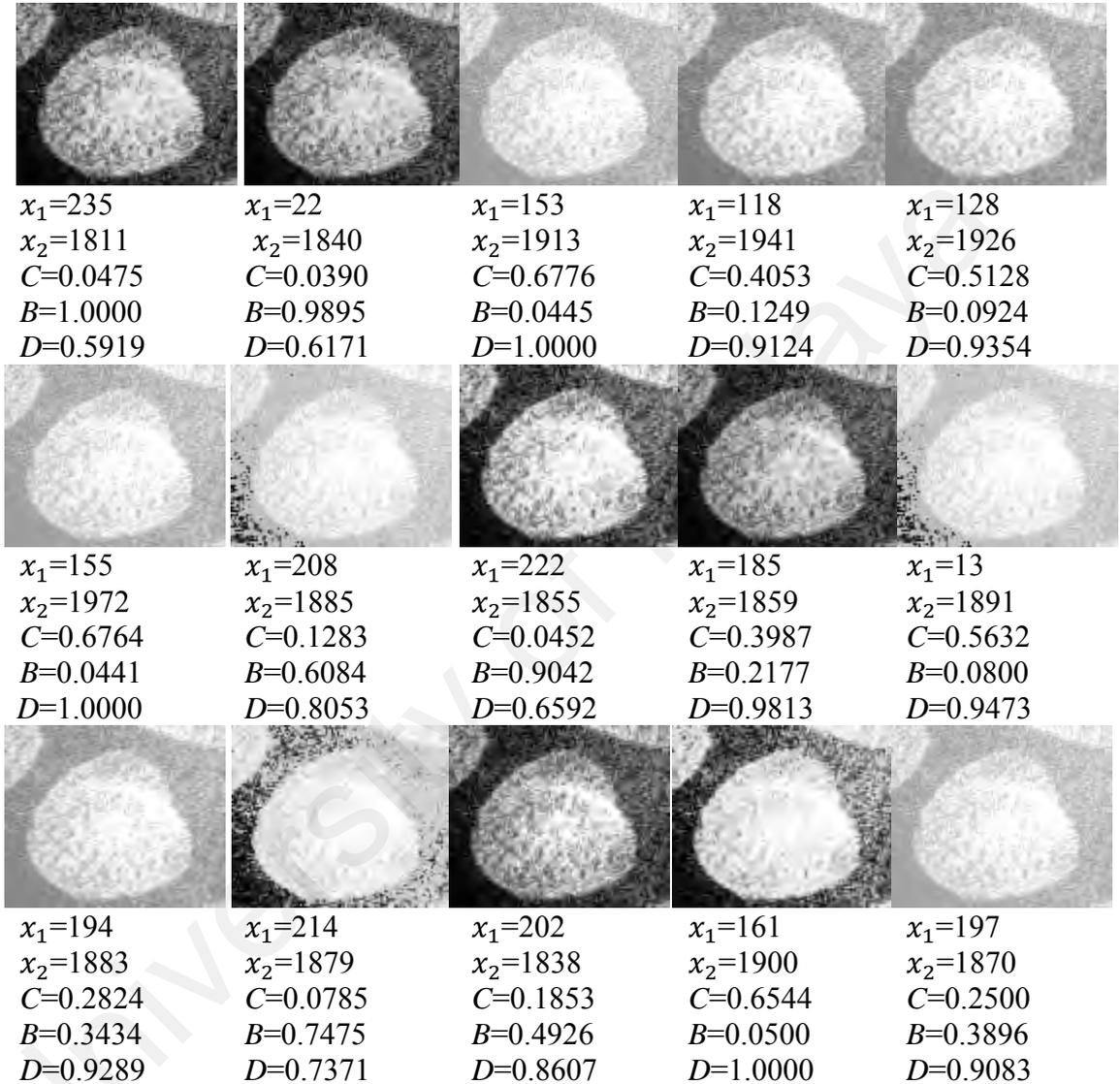


Figure 4.2: Output set of Pareto optimized bi-histogram equalized resultant images of the Lunate bone with its corresponding decision variables and each objective function values, in which x_1 represents the separating point, x_2 represents the clipped limit, C represents Contrast objective, B represents the mean brightness preservation and D represents the detail preservation, respectively.

Figure 4.2 depicts partial of the Pareto optimized resultant images for Lunate bone. In this figure, the Pareto optimized resultant images clearly exhibit significant improvement

on the contrast between the bone and the background. The border in the enhanced image is enhanced to certain extent, where the enhanced border is clearly delineated. As described in the previous section where resultant image of Hamate bone is shown, the border provides sufficient information details that can be analyzed to dictate the stage of maturation of the particular bone. The information related to maturity includes shape, border diameter, inherent border properties, flattening effect at specific side of the bone; the border saliency border and border shape is significantly useful in assisting in maturity rating for Lunate bone. For example, higher maturity stage of Lunate bone is identified by the definite saddle for articulation with outgrowth capitata of lunate dorsal part towards the scaphoid is formed at the distal surface of the Lunate bone. In this, the dorsal part extends out beyond the lateral edge of the palmar part of the saddle, which this extension is from the palmar edge to scaphoid edge. In Figure 4.2, the shown Lunate bone is classified as stage E due to major attributes that are observed in the shown Pareto optimized resultant images. The first attribute is the appearance of the flattening effect at the border that is adjacent to the upper side of radius bone. Meanwhile, the second attribute is the existence of the thicker white line at the distal border of the Lunate bone.

In Figure 4.2, the shown Pareto optimized resultant image exhibits enhanced image in which favor to specific enhancement objectives that had been discussed in the previous chapter in this thesis. From the enhanced images in Figure 4.2, exist image that displayed the white line at the distal part of the bone and also depicted low mean brightness difference with enhanced contrast as compared to over saturated original image. In other words, the set of Pareto optimized resultant image contains a number of enhanced image that successfully emphasizes the three enhancement objective that were previously discussed. The resultant image shows a well balance enhancement that considers the preservation of contrast, mean brightness and details. Some of the noises in the resultant image is easily been filtered away with some standard filtering techniques such as

anisotropic diffusion approach. This filtering technique can be implemented to the original image before the proposed enhancement approach is imposed. However, this filtering algorithm is not implemented in the experiment in order to avoid additional effect in the resultant image caused by other methods that is not related to the proposed approach.

In Figure 4.2, selective images from the Pareto optimized resultant images of Lunate bone is selected to be the input images to the segmentation process, where this process is executed using deep learning approach. The images are selected based on their most optimum value for each objective values: C, B and D, where their value is near to 1.

University of Malaya



Capitate Bone Image (Before Enhancement)

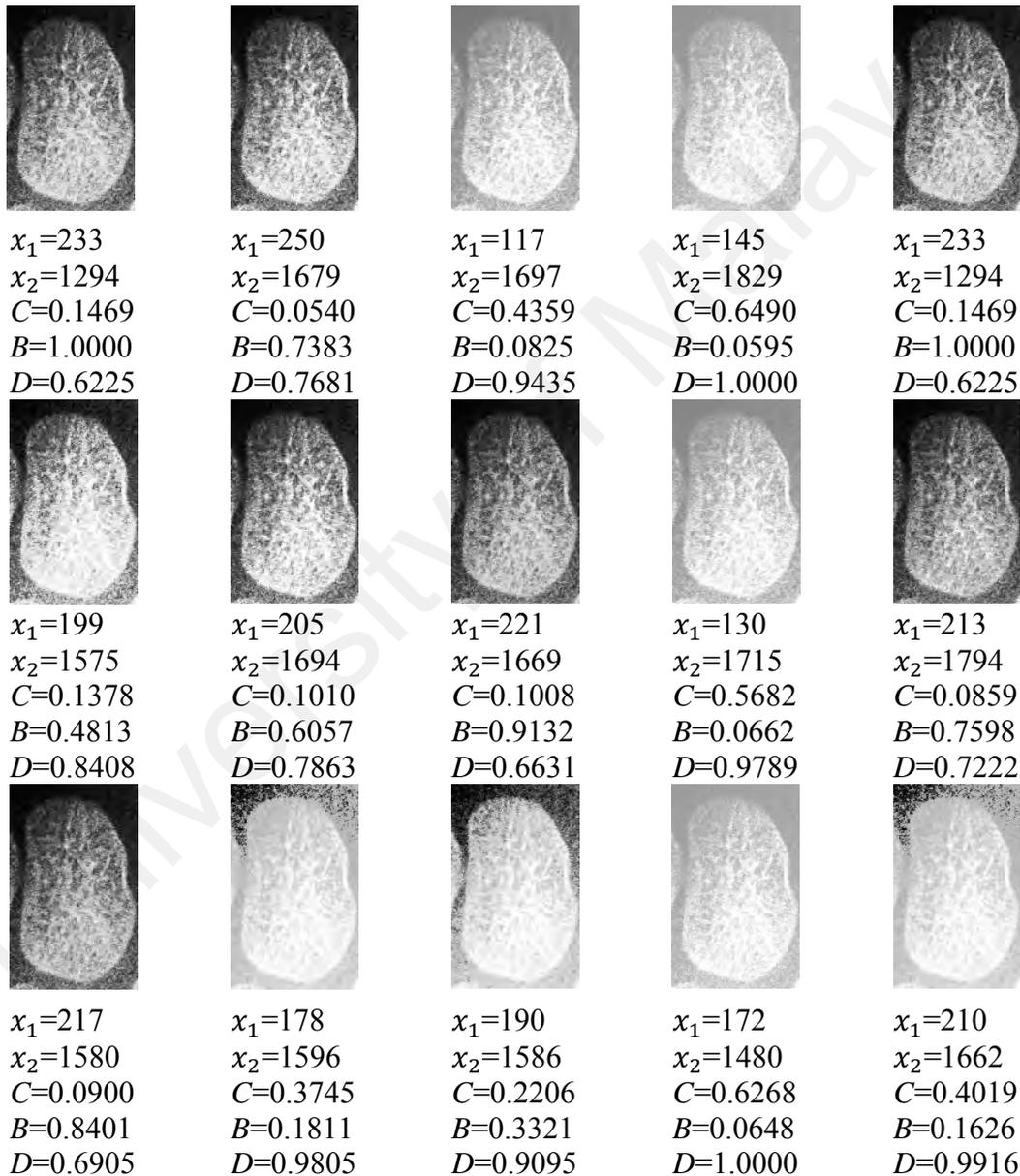


Figure 4.3: Output set of Pareto optimized bi-histogram equalized resultant images of the Capitata bone with its corresponding decision variables and each objective function values, in which x_1 represents the separating point, x_2 represents the clipped limit, C represents Contrast objective, B represents the mean brightness preservation and D represents the detail preservation, respectively.

Figure 4.3 depicts partial of the Pareto optimized resultant images for Capitate bone. In this figure, overall contrast of the original image is clearly enhanced as compared to the original image, where the region of interest of the Capitate bone is substantially visible from the background. In addition, the Pareto optimized resultant image exhibits enhanced internal contrast of the bone content in Capitate bone. It is also can be observed that the border shape, border outline and the contrast within the bone are improved, significantly. Besides, certain images in the set of Pareto optimized image has low mean brightness difference and this outcome is expected as the proposed approach has the ability to fulfills the preservation objectives in an optimal manner during the enhancement process.

Similar to other carpal bones that was previously discussed in the previous section, the enhanced properties such as the shape, border outline and the internal contrast within the bone is crucial for the maturity rating for TW3 system. For example, specific characteristics at the Capitate bone are referred as maturity stage D, which includes the maximum diameter is half or more from the radial metaphysis' width, the border outline that is beside the Hamate bone is flat and slightly convex and lastly, the border outline that is adjacent to the second metacarpal bone is distinct away from the Hamate bone in order to form a D shape. In Figure 4.3, the characteristics shown in the set of Pareto optimized resultant images of Capitate bone indicates maturity stage of E. In this set of images, the characteristics that resembles maturity stage of E are clearly visible. The displayed characteristics include the concavity of the bone border, thicken bone border and its longitudinal diameter that substantially greater than the transverse but less than the distance from its proximal border to the radial metaphysis. Visual observation on the resultant images in Figure 4.3 clearly indicates that all the features that resembles the maturity stage is easily identified and subsequently improving the efficient in maturity stage rating.

In Figure 4.3, selective images from the Pareto optimized resultant images of Capitane bone is selected to be the input images to the segmentation process, where this process is executed using deep learning approach. The images are selected based on their most optimum value for each objective values: C , B and D , where their value is near to 1.

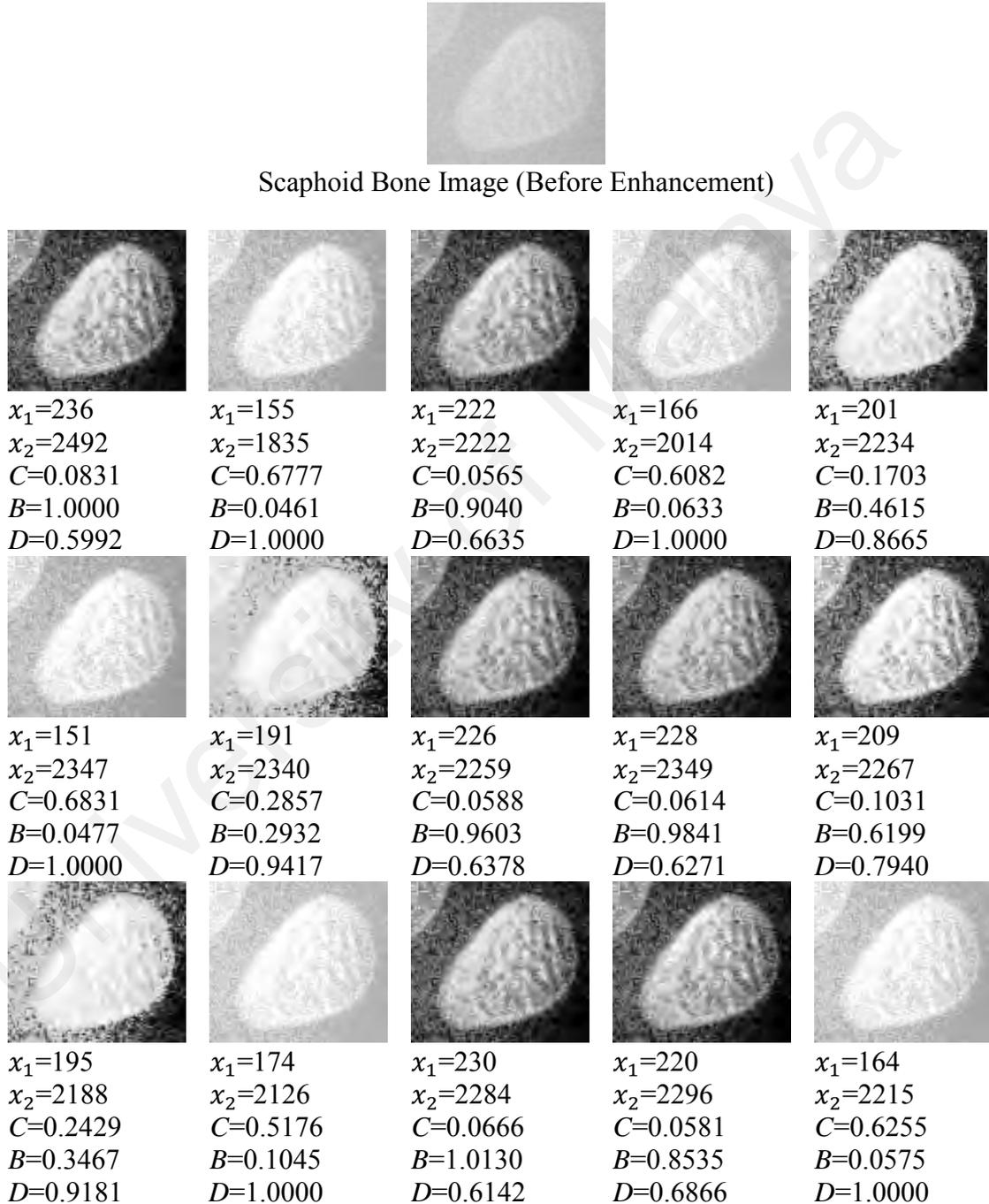


Figure 4.4: Output set of Pareto optimized bi-histogram equalized resultant images of the Scaphoid bone with its corresponding decision variables and each objective function values, in which x_1 represents the separating point, x_2 represents the clipped limit, C represents Contrast objective, B represents the mean brightness preservation and D represents the detail preservation, respectively.

In Figure 4.4, the resultant images in the set of Pareto optimized image exhibits various type of enhancement that fulfills the three objective functions, which is preservation of the contrast, mean brightness and details. Some of the images preserves the mean brightness of the original image with substantial enhancement on contrast. There are images that prone to the contrast objective in which the contrast of the Scaphoid bone is large as compared to the background. Enhanced images that prone to contrast are suitable to be used in evaluating higher maturity stage such as stage E, F, G and H. For each stage, specific attributes from the bone is identified as indicator to the specific maturity stage. For example, visibility of thicken white line that resembles the palmar articular surface is an indicator to stage E. At stage F, the maturity indicators are the concavity attribute of the surface facing the capitate and the flattened border next to the trapezium and trapezoid bones. The bone growth at the proximal and medial direction in which the dorsal surface is extending beyond the thickened white line towards the lunate, the proximal part of the capitate as well as the distinct border near to the lunate are the distinct indicator for maturity stage G. Lastly for the identification of maturity stage H, the enhanced image shows the distal portion of the bone has lateral enlargement and differentiation of the radial styloid articular surface appears in two location. The two locations are the distal part of its lateral border and distinct distal head.

In Figure 4.4, selective images from the Pareto optimized resultant images of Scaphoid bone is selected to be the input images to the segmentation process, where this process is executed using deep learning approach. The images are selected based on their most optimum value for each objective values: C, B and D, where their value is near to 1.

From Figure 4.1 to Figure 4.4, the set of Pareto optimized resultant images for the Carpal bone is visually enhanced to fulfill the three preservation objectives. The enhanced image exhibits clear border and contrast within the bone material, which is an essential

requirement in ensuring effectiveness of the maturity stage assessment. Pareto optimized images provide optimal enhanced output image that highlights various pertinent information in the bone material to the observer.

4.2 Improved Segmentation Result due to Enhanced Contrast

The definition of accurate segmentation of anatomical structures of hands bones means that the most of the regions of hand bones have been classified as truth class and the rest of the pixels are classified as background class. Accurate segmentation of this hand bone structure is instructive for subsequent steps in analyzing and determining bone age. It is known matter in image segmentation communities that the accuracy of segmentation is assessed by comparing the extent of overlapping between the ground truth areas and segmented areas. At first glance, the most intuitive measurement is pixel accuracy. Pixel accuracy here means the percentage of pixels that are classified correctly to truth class. This type of measurement is problematic when the ground truth area is only appears at a small percentage of the entire image. High percentage of accuracy in segmentation is still achievable if all the pixels in the image is classified into background. For example, if only 1% of the entire image is ground truth area, then the naive scheme by classifying all pixels into background pixels will derive segmentation accuracy as 99%. This is due to all background pixels are classified as background pixels and there are 99% of them. This issue is regarded as a class imbalance issue of pixel accuracy calculation. The domination of pixels of certain class will make this measurement fails.

Therefore, overlapping schemes that can avoid this problem are often being used such as the Intersection-Over-Union (IOU). IOU is also named as the “Jaccard Index”. IOU is currently the most conventionally adopted segmentation accuracy measurement since it is also a straightforward metric. IOU is defined as the ratio between the area of overlap

between the segmented pixels and the ground truth pixels, to the area of union between the segmented pixels and the ground truth pixels. The second segmentation accuracy metric that is also commonly adopted is Dice Coefficient or also known as F_1 score. Dice Coefficient is also based on overlapping scheme as well but different from IOU. Dice coefficient is just a ratio of two portions of areas of overlapping between the ground truth and the segmented pixels to the total number of pixels in both ground truth and segmented pixels. Dice Coefficient is different from IOU as the denominator of metric is not the union of ground truth and segmented pixels but the sum of these two areas instead. Dice Coefficient (DC) is adopted as segmentation metric to measure the computed hand bone segmentation result.

The Dice Coefficient (DC) is defined by the following formula

$$DC = \frac{2|A \cap B|}{|A| + |B|} \quad (4.1)$$

In the DC definition, the symbol “A” denotes the ground truth areas represented by the ground of pixels that are identified as the targeted area. In the hand bone context, these ground truth is defined by the areas that were delineated by three medical experts whereas the symbol “B” denotes the segmented pixels or the areas that are identified as hand-bone related. For each value in “A”, the value is either “0” or “1” in which “0” represent not ground truth, and “1” represents it is ground truth. Again the ground truth here represents delineated area of each hand bone image by experts. It is mathematically define $|A|$ as $|A| = \sum_i^{total\ i} a_i$ in which a_i denoted the i^{th} pixel in area “A”. Similarly, the background area $|B|$ can be defined as $|B| = \sum_j^{total\ j} b_j$ in which b_i denoted the j^{th} pixel in area “B”. Therefore, the nominator of DC representation, $|A \cap B|$ can be defined as $|A \cap B| = \sum_i \sum_j a_i b_j$. In other words, there is an entry in the summation if and only if both a_i and

b_j equals to 1. Please note that since both images in which consists of ground truth image and the segmented image are of same dimensions, therefore $i = j$ and thus definition of $|A|$ as $|A| = \sum_i a_i$ and $|B|$ as $|B| = \sum_i b_i$, and $|A \cap B| = \sum_i a_i b_i$.

Dice coefficient also known as F_1 score as stated. Our hand bone segmentation can be regarded as binary classification of each pixel. In statistical analysis for binary classification, the F_1 score is viewed as a measurement of accuracy in test. This measurement of test accuracy comprises of two components: 1) precision, p and 2) recall, r . First, the precision is defined as the ratio of true positive to the total positive results, which is the summation of true positive and false positive. In other words, precision can be viewed as “given the output is positive, how precise this positive result” or “out of all positive results, how many of them is truly a positive result” or “how many of the selected result are relevant”. On the other hands, the recall is calculated by the ratio of the number of true positive to the number of all true positive available in which is summation of true positive and false negative. In other words, recall can be viewed as “of all true positive results, how many of them will be identified correctly as true positive result” or “how many relevant results are selected”. The F_1 score is defined to be the harmonic mean of both these components (precision, denoted as p and recall denoted as r) as following:

$$F_1 = \frac{2}{p^{-1} + r^{-1}} \quad (4.2)$$

And thus can be then rearranged to become the following:

$$F_1 = \frac{2pr}{p+r} \quad (4.3)$$

Note the subscript “1” in F_1 can be generalized to any positive real number if different emphasis on p and r is assigned. If r is β times more important than p, then the general formula is as following:

$$F_\beta = \frac{(1 + \beta^2) pr}{\beta^2 p + r} \quad (4.4)$$

In the proposed hand bone segmentation measurement, F_1 metric is adopted such that both recall and precision is equally important. In the context of segmentation, precision here represents the ratio of pixels which are predicted correctly as bone to the total pixels which are ground truth as bones. On the other hand, the recall here represents the ratio of pixels which are predicted correctly as background to the total pixels of bone. In the hand bone segmentation context, true positive for $|A|$ and $|B|$ is different. Besides, it is important to note that the maximum value of DC is 1 and the minimum is 0 or ranges from 0 to 1 in which being 1 means that all the pixels predicted as bones are all actually bones as delineated by experts and being 0 means that all the pixels predicted as bones are all not bones.

The dataset that is adopted to generate the segmentation result is the database of hand bone, which acquired from an open database of hand images specifically for research purposes (Joseph, L., 2005). Experiments are conducted for images before and after the proposed contrast enhancement scheme. The experiments are designed in this way with the intention to emphasize the impact of the contrast of the image on the segmentation result. For each age group and gender group, most of the images in database are used to generate the segmentation result by using the segmentation quality metrics, dice coefficient with equal weight on p and r. In other words, one pixel misclassified as bone is the same penalty of one pixel misclassified as background and vice versa. All results

require ground truth and we obtained it from three experts by delineating the images for real areas of bones. For each age group and gender group, the median, 25th and 75th percentiles, maximum and minimum values of dice coefficient are computed in order to identify any performance discrepancy of the proposed segmentation in different age and gender group. Boxplot is used to demonstrate the result. Note that the soft tissue region is considered as background as well in this context and in fact this is actually the main challenge in the segmentation.

Table 4.1 shows the experimental result and the associated overall mean and standard deviation.

University of Malaya

Table 4.1: Experimental result for segmentation using Dice Coefficient across 19 age groups for male category. The value of each cell represents the mean dice coefficient of a particular age group using the segmentation scheme against the ground truth by a particular expert.

Age Group (male)	Before Contrast Enhancement			After Contrast Enhancement		
	Expert 1	Expert 2	Expert 3	Expert 1	Expert 2	Expert 3
0	0.9911	0.9937	0.9901	0.9925	0.9912	0.9929
1	0.9917	0.9936	0.9904	0.9963	0.9911	0.9985
2	0.9749	0.9632	0.9764	0.9571	0.9715	0.9933
3	0.9654	0.9559	0.9636	0.9612	0.9521	0.9874
4	0.8986	0.8871	0.8970	0.8982	0.8846	0.8791
5	0.8816	0.8481	0.8834	0.8640	0.8742	0.8409
6	0.8544	0.8320	0.8573	0.8541	0.8649	0.8493
7	0.8636	0.8147	0.8650	0.8736	0.8596	0.8551
8	0.8307	0.8177	0.8325	0.8305	0.8347	0.8343
9	0.7283	0.7662	0.7301	0.7705	0.7644	0.7307
10	0.7465	0.7565	0.7434	0.7831	0.7715	0.7428
11	0.7350	0.7328	0.7346	0.7821	0.7730	0.7432
12	0.7200	0.7309	0.7177	0.7774	0.7411	0.7509
13	0.7231	0.7227	0.7236	0.7790	0.7515	0.7314
14	0.7265	0.7115	0.7260	0.7853	0.7618	0.7438
15	0.7157	0.7025	0.7219	0.7702	0.7533	0.7659
16	0.7344	0.7407	0.7368	0.7730	0.7430	0.7405
17	0.7401	0.7216	0.7408	0.7689	0.7597	0.7202
18	0.7376	0.7531	0.7369	0.7648	0.7646	0.7652
Overall Mean	0.8189	0.8129	0.8193	0.8411	0.8320	0.8245
Overall Standard Deviation	0.1048	0.1005	0.1044	0.0829	0.0901	0.1017

Table 4.1 shows the experimental result for the proposed segmentation scheme using Dice Coefficient across 19 age groups for male category. The value of each cell represents

the mean dice coefficient of a particular age group using the segmentation scheme against the ground truth by a particular expert. From the table, there are several interesting results can be observed. First of all, the overall mean value is the first observation key point. This value indicates the average DC value of the segmented bone across all the age group in which ranging from 0 to 18, and for each of the three expert delineation of bone anatomical structure on the hand bone radiographic images. Firstly, it is observed that the overall mean for across the three expert is consistent (0.8189, 0.8129, 0.8193 in images without contrast enhancement and 0.8411, 0.8320, 0.8245 in images with contrast enhancement) which standard deviation results is only 0.036 for images without contrast enhancement and 0.083 for images with contrast enhancement. In other words, the inter-variability of ground truth delineated from the three experts are very consistent. This in turn also implies that segmentation from human vision system of the hand bone are consistent. By comparing these two inter-variability values, the observer's inter-variability for images with contrast enhancement is although small but higher than the observer's inter-variability for images without contrast enhancement. This result is expected since the image after contrast enhancement is not comparable to natural image from x-ray. However, the interesting result is that the overall mean across three experts for images after contrast enhancement is significantly higher than those without contrast enhancement, where mean of overall mean for three experts before contrast enhancement is 0.8170 whereas the mean of the overall mean for three experts after contrast enhancement is 0.8325. Figure 4.5 shows the overall mean before and after enhancement for each expert. In Figure 4.5, it is clearly highlighted that the overall mean for each expert is higher in the scenario after enhancement as compared to before enhancement. Hence, the contrast enhancement is noticed to be having the effect of enhancing the segmentation result as well. Lastly from this table, another important implication is that by using the segmentation scheme, the standard deviation is only around 0.1 although the hand bone

image across 19 age group is of enormous difference. This results are a promising indicator to the robustness of the proposed segmentation scheme. Last but not least from this table is that the standard deviation of DC values for images after contrast enhancement is consistently lower than the standard deviation for images before contrast enhancement indicating that although the observer's inter-variability is higher using human visual system, the most important is the analysis from machine can benefit more from the contrast enhancement scheme in segmentation performance.

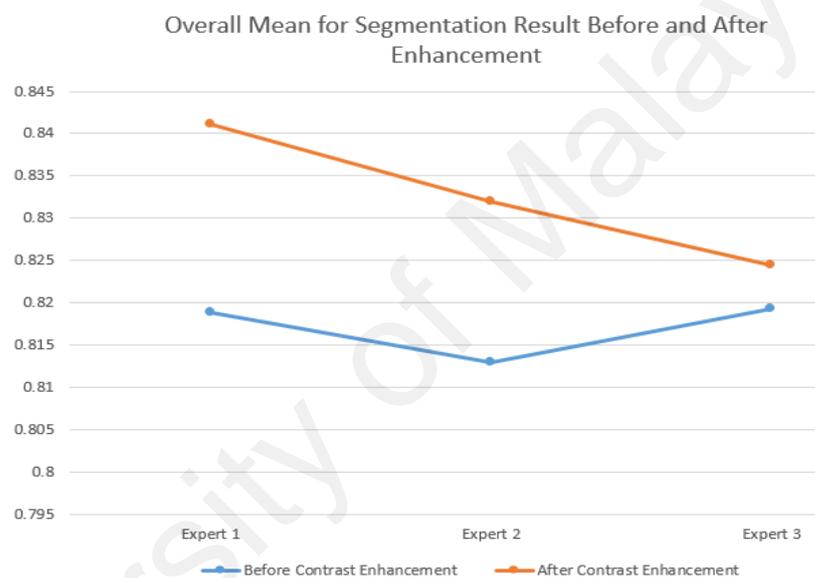


Figure 4.5: Overall mean for three experts for before and after enhancement

The significance difference before and after contrast enhancement for each expert is further verified using a statistical test, known as paired T-test. In this test, the null hypothesis refers to the mean before enhancement is equivalent to after enhancement with 5% significant level. In this paired T-test evaluation, it is assumed that the data follows bivariate normal distribution. The T-test is of the form sample mean difference/sample standard deviation of the sample mean difference. Table 4.2 illustrates the paired T-test sample for Expert 1.

Table 4.2: T-test: Paired two sample for means for Expert 1 for male category.

	<i>Before Enhancement</i>	<i>After Enhancement</i>
Mean	0.818905263	0.841147368
Variance	0.010977081	0.006864858
Observations	19	19
Pearson Correlation	0.987306068	
Hypothesized Mean Difference	0	
df	18	
t Stat	-3.662453026	
p(T<=t) one-tail	0.000890777	
t Critical one-tail	1.734063607	
p(T<=t) two-tail	0.001781554	
t Critical two-tail	2.10092204	

Table 4.2 shows the p value for Expert 1 is 0.001781554. This value depicts strong evidence (<0.05) to reject the null hypothesis. In other words, the difference between before and after enhancement is significant and not likely to occur by chance. Similarly, the same test is performed on Expert 2. Table 4.3 depicts the paired T-test sample for Expert 2.

Table 4.3: T-test: Paired two sample for means for Expert 2 for male category.

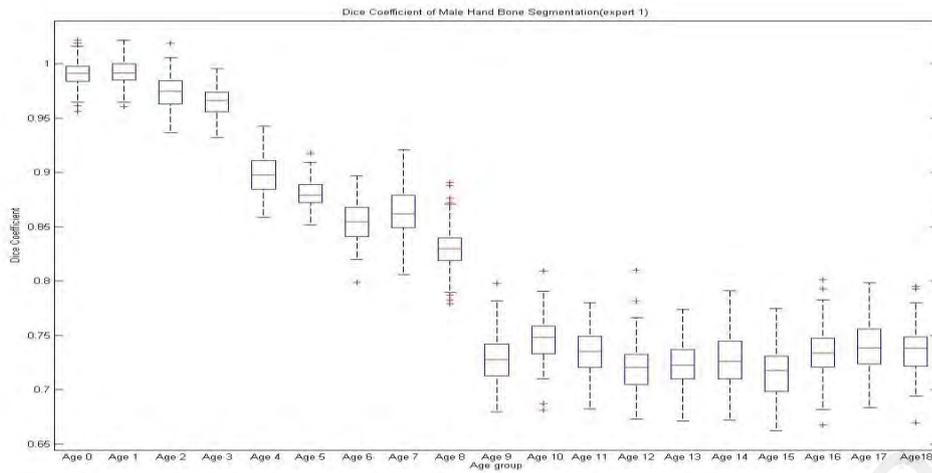
	<i>Before Enhancement</i>	<i>After Enhancement</i>
Mean	0.812868421	0.831989474
Variance	0.010097142	0.008122645
Observations	19	19
Pearson Correlation	0.985368655	
Hypothesized Mean Difference	0	
df	18	
t Stat	-4.319498021	
p(T<=t) one-tail	0.000206431	
t Critical one-tail	1.734063607	
p(T<=t) two-tail	0.000412863	
t Critical two-tail	2.10092204	

In Table 4.3, the null hypothesis for Expert 2 result on before and after enhancement images is that both means are equal, $H_0: \mu_{before} = \mu_{after}$. The p value obtained is 0.000412, showing strong evidence to reject the null hypothesis. For Expert 3, paired T-test is conducted too.

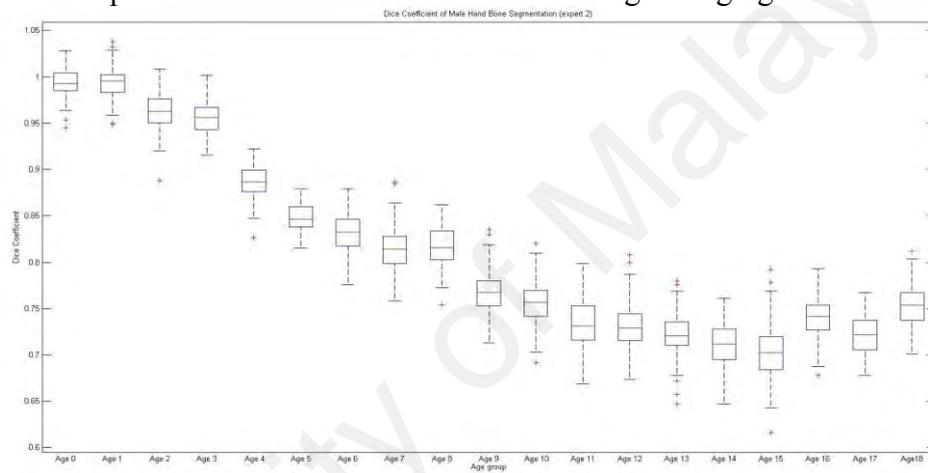
Table 4.4: T-test: Paired two sample for means for Expert 3 for male category.

	<i>Before Enhancement</i>	<i>After Enhancement</i>
Mean	0.819342105	0.824494737
Variance	0.010895891	0.010340388
Observations	19	19
Pearson Correlation	0.98086553	
Hypothesized Mean Difference	0	
df	18	
t Stat	-1.104542694	
p(T<=t) one-tail	0.14195228	
t Critical one-tail	1.734063607	
p(T<=t) two-tail	0.283904561	
t Critical two-tail	2.10092204	

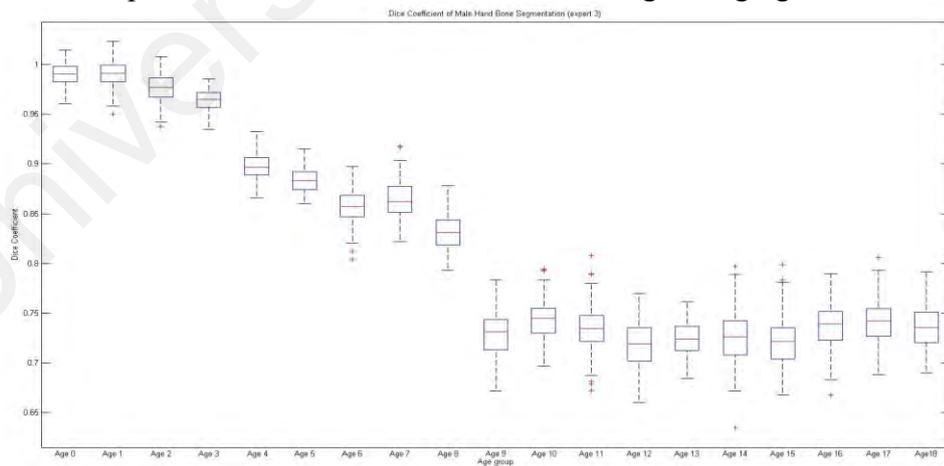
In Table 4.4, the p value obtained is 0.283904561 showing strong evidence to accept null hypothesis. From the statistical analysis, two out of three expert's validation results show significant difference before and after enhancement. Therefore, this significant difference shows the enhancement able to improve the segmentation results of hand bone image and contribute substantially in bone age assessment.



(a) Boxplot of DC for hand bone segmentation with ground truth delineated by expert 1 across all male hand bones with ages ranging from 0 to 18.



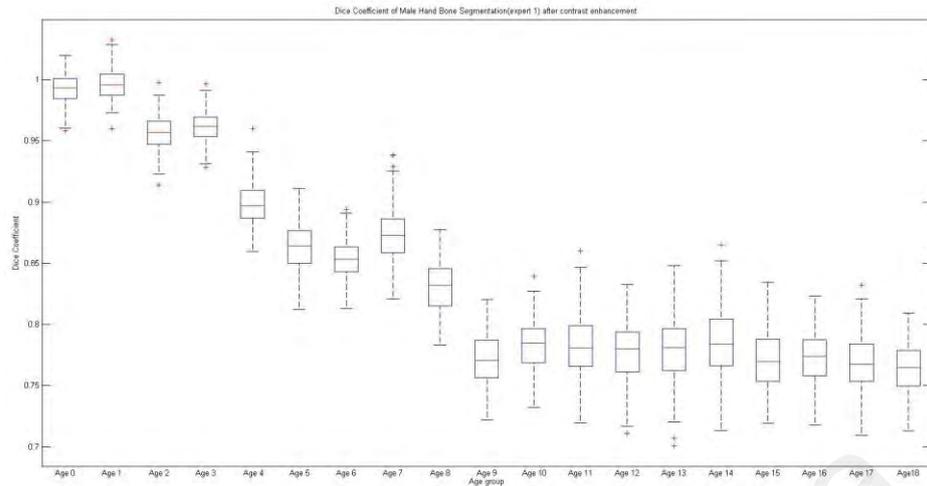
(b) Boxplot of DC for hand bone segmentation with ground truth delineated by expert 2 across all male hand bones with ages ranging from 0 to 18.



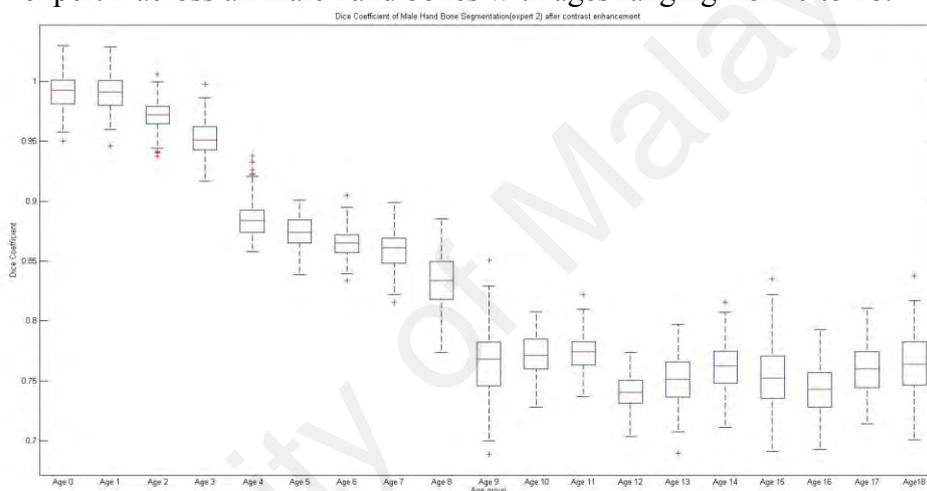
(c) Boxplot of DC for hand bone segmentation with ground truth delineated by expert 3 across all male hand bones with ages ranging from 0 to 18.

Figure 4.6: Boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1, 2 and 3 across all male hand bones with ages ranging from 0 to 18, before applying the contrast enhancement scheme.

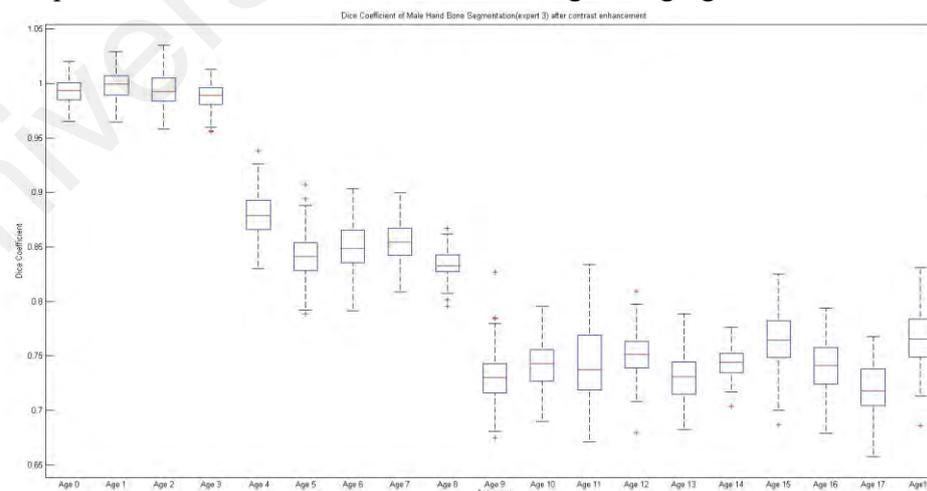
Figure 4.6 shows the boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert1, 2 and 3 across all male hand bones with ages ranging from 0 to 18, before applying the contrast enhancement scheme. Figure 4.6 (a) shows the boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1 across all male hand bones with ages ranging from 0 to 18. Figure 4.6 (b) shows boxplot of dice coefficient for hand bone segmentation with ground truth delineated by expert 2 across all male hand bones with ages ranging from 0 to 18. Figure 4.6 (c) shows boxplot of dice coefficient for hand bone segmentation with ground truth delineated by expert 3 across all male hand bones with ages ranging from 0 to 18. From these boxplots, it can be noticeable that the across 19 age groups, the median of DC drops sharply from age 4 to age 5 and from age 8 to age 9. Another finding from the three plots is that the DC drops consistently that from age 0 to age 8 (mean drops from 0.9911 to 0.8307 for expert 1, mean drops from 0.9937 to 0.8177 for expert 2, mean drops from 0.9901 to 0.8325 for expert 3). After age 8, it can be noticeable that the DC value is rather consistent. For example, the standard deviation for expert 1 from age 0 to age 8 is 0.0638 and the standard deviation for expert 1 from age 8 to age 18 is 0.0315; 0.0760 and 0.0318 for expert 2, 0.0626 and 0.0316 for expert 3. This result is aligned to the complexity of the anatomical structures in the hand bone throughout the growth period. From the hand bone radiographic image, it can be observed that age 8 onwards, as compared to age 7, differs much in the anatomical structure such that age 8 onwards, carpals and epiphyseal sites begins to overlap with each other. For the age 4 to age 5, the underlying reasons might be the appearance of epiphyseal plates in which the difficulty in segmentation of epiphyseal plates is higher than other anatomical structure of hand bone.



(a) Boxplot of DC for hand bone segmentation with ground truth delineated by expert 1 across all male hand bones with ages ranging from 0 to 18.



(b) Boxplot of DC for hand bone segmentation with ground truth delineated by expert 2 across all male hand bones with ages ranging from 0 to 18.



(c) Boxplot of DC for hand bone segmentation with ground truth delineated by expert 3 across all male hand bones with ages ranging from 0 to 18.

Figure 4.7 boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1, 2 and 3 across all male hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme.

Figure 4.7 shows the boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1, 2 and 3 across all male hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme. Figure 4.7(a) shows the boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1 across all male hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme. Figure 4.7(b) shows boxplot of dice coefficient for hand bone segmentation with ground truth delineated by expert 2 across all male hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme. Figure 4.7(c) shows boxplot of dice coefficient for hand bone segmentation with ground truth delineated by expert 3 across all male hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme. From these boxplots, it can be noticeable that the across 19 age groups, the median of DC drops also sharply from age 4 to age 5 and from age 8 to age 9 and the DC drops also consistently that from age 0 to age 8 similarly to previous result before applying contrast enhancement. After age 8, it can also be noticeable that the DC value is rather consistent. From the boxplots, it can be noticed that almost all the medians, for example, age group from age 8 onwards consistently stays above 0.75 for hand bone radiographic image after contrast enhancement, on the other hand, for hand bone x-ray image before contrast enhancement, for age group from age 8 onwards, DC value consistently stays before 0.75. From these result, we find that DC value is higher for images that have undergone contrast enhancement; or in other words, segmentation result is enhanced for hand bone radiographic image that have undergone contrast enhancement.

Table 4.5: Experimental result for segmentation using Dice Coefficient across 19 age groups for female category. The value of each cell represents the mean dice coefficient of a particular age group using the segmentation scheme against the ground truth by a particular expert.

Age Group (Female)	Before Contrast Enhancement			After Contrast Enhancement		
	Expert 1	Expert 2	Expert 3	Expert 1	Expert 2	Expert 3
0	0.9929	0.9949	0.9931	0.9923	0.9960	0.9963
1	0.9827	0.9873	0.9970	0.9930	0.9913	0.9906
2	0.9469	0.9757	0.9879	0.9848	0.9733	0.9826
3	0.9245	0.9424	0.9328	0.9722	0.9768	0.9875
4	0.8725	0.8886	0.8651	0.9135	0.9220	0.9410
5	0.8664	0.8516	0.8417	0.8970	0.9060	0.9355
6	0.8430	0.8337	0.8349	0.8853	0.8942	0.8927
7	0.8295	0.8313	0.8209	0.8801	0.8601	0.8883
8	0.7592	0.7440	0.7194	0.7869	0.7354	0.7565
9	0.7411	0.7326	0.7320	0.7703	0.7413	0.7306
10	0.7338	0.7461	0.7433	0.7557	0.7531	0.7464
11	0.7218	0.7306	0.7399	0.7399	0.7299	0.7241
12	0.7308	0.7395	0.7386	0.7502	0.7305	0.7403
13	0.7114	0.7192	0.7313	0.7299	0.7407	0.7323
14	0.7337	0.7448	0.7150	0.7570	0.7264	0.7534
15	0.7447	0.7542	0.7331	0.7427	0.7556	0.7641
16	0.7090	0.7300	0.7210	0.7401	0.7596	0.7527
17	0.7415	0.7331	0.7427	0.7488	0.7536	0.7324
18	0.7258	0.7263	0.7450	0.7443	0.7458	0.7260
Overall Mean	0.8059	0.8108	0.8071	0.8307	0.8259	0.8302
Overall Standard Deviation	0.0978	0.1000	0.1018	0.1012	0.1053	0.1105

Table 4.5 shows the experimental result for segmentation using Dice Coefficient across 19 age groups for female category. The value of each cell represents the mean dice coefficient of a particular age group using the segmentation scheme against the ground truth by a particular expert. From the table, several findings are observed. First of all, the overall mean value discrepancy between male and female result. It is noted that this overall mean value indicates the average DC value of the segmented bone across all the age group in which ranging from 0 to 18, for each of the three expert delineation of bone anatomical structure on the radiographic images. For male, without contrast enhancement, the overall mean of DC value is 0.8189, 0.8129, 0.8193 respectively for each expert ground truth, while for female, without contrast enhancement, the overall mean is 0.8059, 0.8108, 0.8071. From the results in term of DC value, it is observed that the result for male is slightly consistently higher than DC value of female hand bone radiographic images. This result is fascinating but not surprising since the female bone maturity development is faster than male and therefore the overlapping of bones of epiphyseal sites occur more frequently in the dataset. On the other hand, for male, with contrast enhancement, 0.8411, 0.8320, 0.8245 respectively for each expert ground truth, while for female, the overall mean is 0.8307, 0.8259, 0.8302. The average of the overall mean for male, with contrast enhancement is therefore 0.83525 while the average if the overall mean for female, with contrast enhancement is 0.8289. This results in turn implies that either with or without contrast enhancement, the segmentation accuracy for female hand bone radiographic image correspond to the complexity of the anatomical structure of the hand bone. The standard deviation of the DC value for both male and female is around 0.1. From the obtained results, this standard deviation value is astoundingly small according to the prominent change of anatomical structure of bone across 18 age group.

For female category, the significance difference before and after contrast enhancement for each expert is also verified using paired T-test. Table 4.6 illustrates the paired T-test sample for Expert 1.

Table 4.6: T-test: Paired two sample for means for Expert 1 for female category.

	<i>Before Enhancement</i>	<i>After Enhancement</i>
Mean	0.805852632	0.830736842
Variance	0.009574223	0.01025176
Observations	19	19
Pearson Correlation	0.989158123	
Hypothesized Mean Difference	0	
df	18	
t Stat	-7.208698259	
p(T<=t) one-tail	5.225E-07	
t Critical one-tail	1.734063607	
p(T<=t) two-tail	1.045E-06	
t Critical two-tail	2.10092204	

Table 4.6 shows the p value for Expert 1 is 1.045E-06. This value depicts strong evidence (<0.05) to reject the null hypothesis that both means for before and after enhancement is equal. In other words, the difference between before and after enhancement is significant. Similarly, the same test is performed on Expert 2. Table 4.7 depicts the paired T-test sample for Expert 2.

Table 4.7: T-test: Paired two sample for means for Expert 2 for female category.

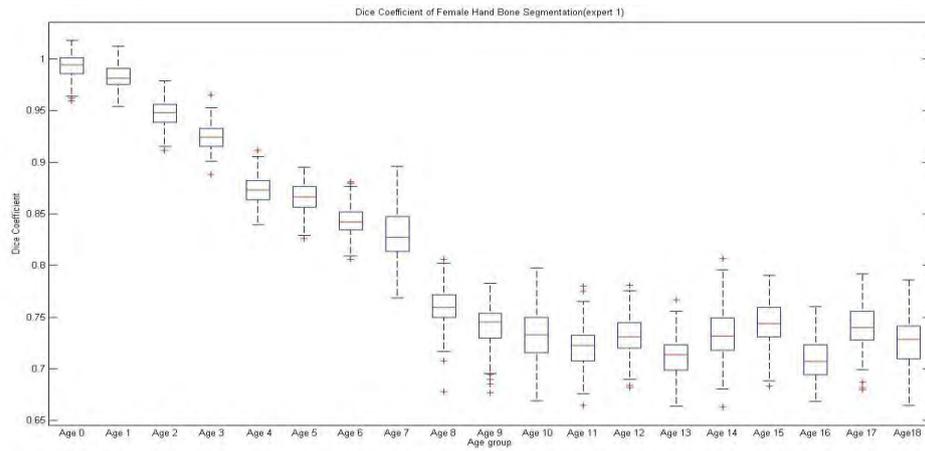
	<i>Before Enhancement</i>	<i>After Enhancement</i>
Mean	0.810836842	0.825873684
Variance	0.010008519	0.011076625
Observations	19	19
Pearson Correlation	0.979561258	
Hypothesized Mean Difference	0	
df	18	
t Stat	-3.064442606	
p(T<=t) one-tail	0.003339306	
t Critical one-tail	1.734063607	
p(T<=t) two-tail	0.006678612	
t Critical two-tail	2.10092204	

In Table 4.7, the p value obtained is 0.006678612 ($\ll 0.05$), showing strong evidence to reject the null hypothesis. For Expert 3, paired T-test is conducted too.

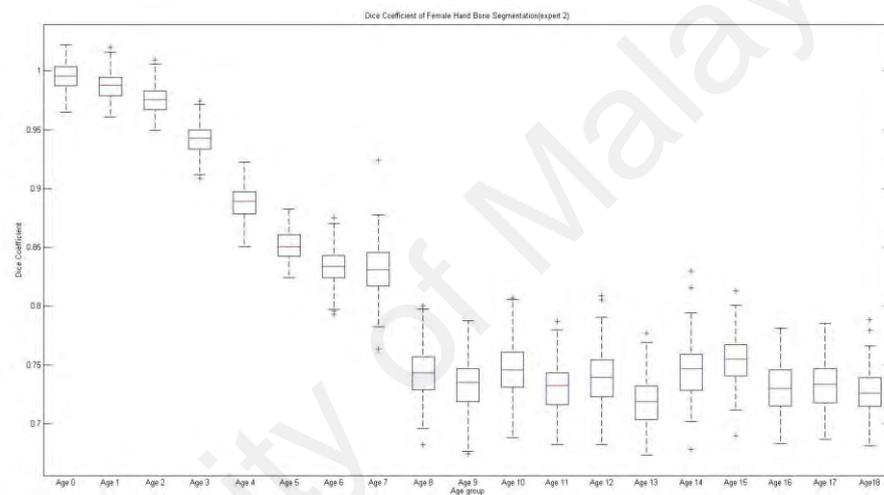
Table 4.8: T-test: Paired two sample for means for Expert 3 for female category.

	<i>Before Enhancement</i>	<i>After Enhancement</i>
Mean	0.807089474	0.830173684
Variance	0.010355691	0.012218172
Observations	19	19
Pearson Correlation	0.951713865	
Hypothesized Mean Difference	0	
df	18	
t Stat	-2.95022288	
p(T<=t) one-tail	0.004281256	
t Critical one-tail	1.734063607	
p(T<=t) two-tail	0.008562512	
t Critical two-tail	2.10092204	

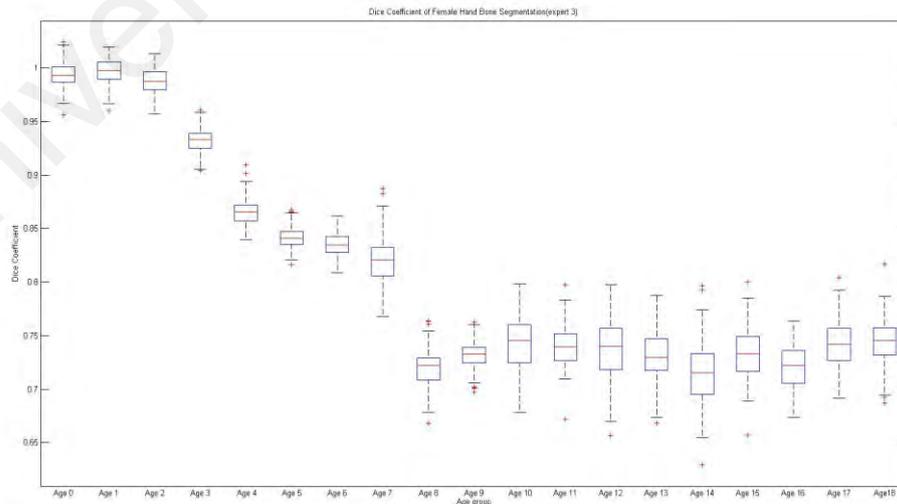
In Table 4.8, the p value obtained is 0.008562512 showing strong evidence to reject null hypothesis. From the statistical analysis, three expert's validation results show significant difference before and after enhancement. Therefore, this significant difference shows the enhancement able to improve the segmentation results of hand bone image in female category and contribute substantially in bone age assessment.



(a) Boxplot of DC for hand bone segmentation with ground truth delineated by expert 1 across all female hand bones with ages ranging from 0 to 18.



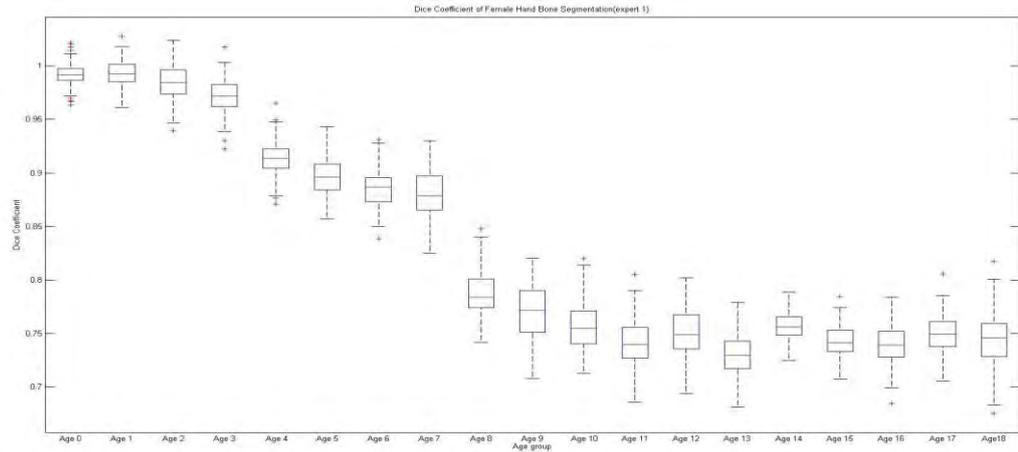
(b) Boxplot of DC for hand bone Segmentation with ground truth delineated by expert 2 across all female hand bones with ages ranging from 0 to 18.



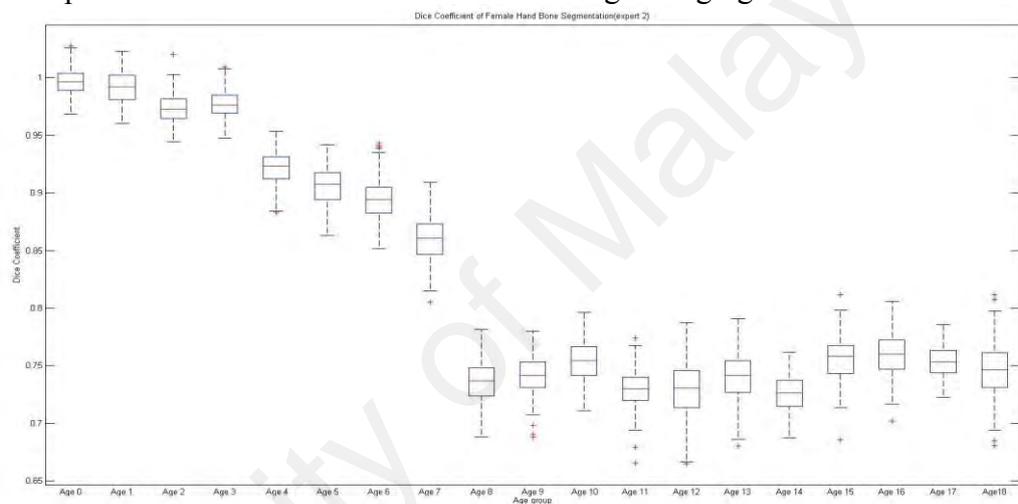
(c) Boxplot of DC for hand bone segmentation with ground truth delineated by expert 3 across all female hand bones with ages ranging from 0 to 18.

Figure 4.8 Boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1, 2 and 3 across all female hand bones with ages ranging from 0 to 18, before applying the contrast enhancement scheme.

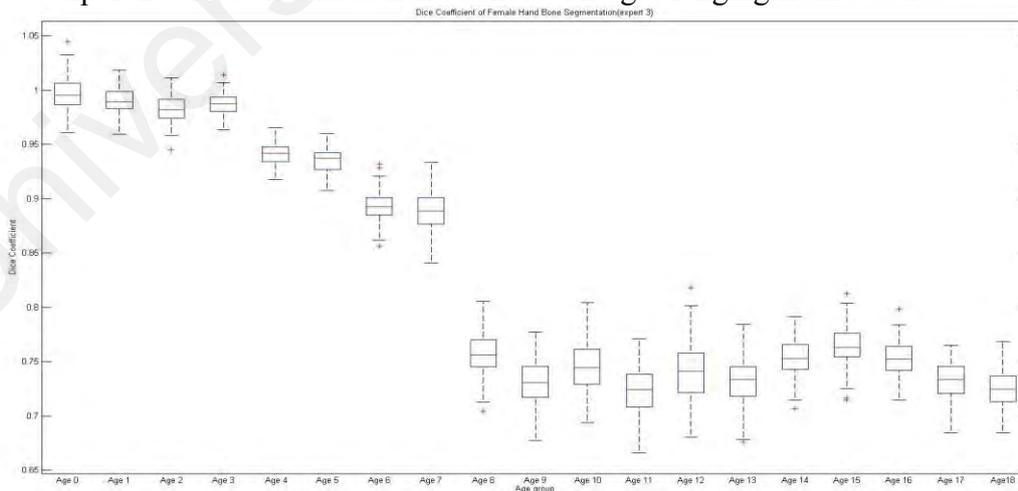
Figure 4.8 shows the boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1, 2 and 3 across all female hand bones with ages ranging from 0 to 18, before applying the contrast enhancement scheme. Figure 4.8(a) shows the boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1 across all female hand bones with ages ranging from 0 to 18, before applying the contrast enhancement scheme. Figure 4.8(b) shows boxplot of dice coefficient for hand bone segmentation with ground truth delineated by expert 2 across all female hand bones with ages ranging from 0 to 18, before applying the contrast enhancement scheme. Figure 4.8(c) shows boxplot of dice coefficient for hand bone segmentation with ground truth delineated by expert 3 across all female hand bones with ages ranging from 0 to 18, before applying the contrast enhancement scheme. From these boxplots, it can be noticeable that the across 19 age groups, the median of DC drops sharply from age 7 to age 8 as opposed to the previous result for male in which the median of DC drops sharply from age 8 to age 9. Another finding from the three plots is that the result corresponds well with the previous result of male such that before the DC value drop sharply, the mean value drops consistently in which mean drops from 0.9929 to 0.8295 for expert 1, mean drops from 0.9949 to 0.8313 for expert 2, mean drops from 0.9931 to 0.8209 for expert 3. However, the mean value drop happens earlier as compared to the male in term of the DC value. This result corresponds well with development expectation such that the bone age development of female children is faster than the male children. The segmentation result is therefore records a drop since 7 from the radiographic image. This is due to the rapid anatomical bone structure development and all the carpals, phalanges and metacarpals started to overlap with each other and pose serious issue to clear-cut segmentation.



(a) Boxplot of DC for hand bone segmentation with ground truth delineated by expert 1 across all female hand bones with ages ranging from 0 to 18.



(b) Boxplot of DC for hand bone segmentation with ground truth delineated by expert 2 across all female hand bones with ages ranging from 0 to 18.



(c) Boxplot of DC for hand bone segmentation with ground truth delineated by expert 3 across all female hand bones with ages ranging from 0 to 18.

Figure 4.9 Boxplot of Dice Coefficient for hand Bone segmentation with ground truth delineated by expert 1, 2 and 3 across all female hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme.

Figure 4.9 shows the boxplot of dice coefficient for hand bone segmentation with ground truth delineated by expert 1, 2 and 3 across all female hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme. Figure 4.9 (a) shows the boxplot of Dice Coefficient for hand bone segmentation with ground truth delineated by expert 1 across all female hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme. Figure 4.9 (b) shows boxplot of dice coefficient for hand bone segmentation with ground truth delineated by expert 2 across all male hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme. Figure 4.9 (c) shows boxplot of dice coefficient for hand bone segmentation with ground truth delineated by expert 3 across all female hand bones with ages ranging from 0 to 18, after applying the contrast enhancement scheme. From these boxplots, it can be noticeable that the across 19 age groups, the median of DC drops also sharply from age 4 to age 5 and from age 7 to age 8 and the DC drops also consistently that from age 0 to age 7 similarly to previous result before applying contrast enhancement. After age 7, it can also be noticeable that the DC value is rather consistent. From the boxplots, it can be noticed that almost all the medians, for example, age group from age 8 onwards consistently stays between 0.75 for hand bone radiographic image after contrast enhancement. On the other hand, for hand bone radiographic image before contrast enhancement, for age group from age 8 onwards, DC value although consistently stays between 0.75 as well, but in overall the DC value range is slightly higher than before contrast enhancement. Besides there are some outliers found which are believed cause by extreme conditioned radiographic images in the dataset in which the contrast is washed-out even before contrast enhancement and therefore there is a few outliers in some age group.

4.3 Quantitative Analysis using Classifier

For this quantitative analysis, five maturity stage are considered, which are stage D, stage E, stage F, stage G and stage H. These stages indicate the level of maturity for ossification sites of bones. These five stages possess quite similar features with each others and it is challenging to identify them. Therefore, these five stages are chosen in order to measure the effect of the improved visual quality from the proposed enhancement scheme onto the classification result. In this experiments, 2500 images with the cropped ossification sites are used. In this experiments, the images are balance distributed to the five stages, where each stage has 500 images. From the 2500 images, three different sets of training images are created. The first set of training images are the original images without any enhancement. The second set of training images are the enhanced images that are generated by using the bi-histogram equalization algorithm with random values of the separating point and clipped limit. By using random value on separating point and clipped limit, the output of the enhanced image includes all the possible outcome for all the approaches that is discussed in Section 2.3.1. The third set of training images are the enhanced images that are computed by using the proposed Pareto optimized bi-histogram, where the separating point and clipped limit are optimally determined. With these three set of training images, classifiers are trained with the same setting in order to get the optimum classifiers. The training parameter settings are made standardize throughout the experiments. The classification result from the classifier that is trained from the first set of training images (original image without enhancement) is tabulated in Table 4.9. The classification results from the classifier that adopts random values of separating point and clipped limit is illustrated in Table 4.10. Table 4.11 depicts the classification result that is obtained from the classifier that is trained using Pareto optimized bi-histogram images. Prior to the discussion on the three classification results, the metrics that used in the experiment that includes precision, recall and F1 score will be described and explained.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (4.5)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (4.6)$$

$$\text{F1 score} = \frac{2 (\text{Precision})(\text{Recall})}{\text{Precision} + \text{Recall}} \quad (4.7)$$

The precision value as defined by Equation (4.5) is the ratio of true positive to the total number of positive recognized result. Therefore, the precision value of the classifier quantifies the reliability of the classifier result if the result is positive. It means the higher the precision value, the more reliable the classifier result is, and vice versa. The recall value as defined by Equation (4.6) is the ratio of true positive to the total number of positive assigned images. On the other hand, the recall value, represents the confidence to recognize a test image with positive label. It means the higher the recall value, the higher the chance a particular image of positive label will be recognized correctly, and vice versa. The F1 score represents the harmonic mean between precision and recall as indicated by Equation (4.7). The F1 score indicates the classifier accuracy by considering both precision and recall.

Table 4.9: Classification Accuracy of Test Images without Enhancement

	Stage D	Stage E	Stage F	Stage G	Stage H
True Positive	378	357	312	298	305
False Positive	122	143	188	202	195
False Negative	114	141	232	216	147
Precision	0.756	0.714	0.624	0.596	0.610
Recall	0.768	0.717	0.574	0.578	0.675
F1 Score	0.762	0.715	0.598	0.587	0.641

Table 4.10: Classification Accuracy of Test Images with Bi-Histogram Equalization using Random Values of Separating Points and Clipped Limit

	Stage D	Stage E	Stage F	Stage G	Stage H
True Positive	398	384	342	321	341
False Positive	102	116	158	179	159
False Negative	86	145	173	172	138
Precision	0.796	0.768	0.684	0.642	0.682
Recall	0.822	0.726	0.689	0.651	0.712
F1 Score	0.809	0.746	0.686	0.646	0.697

Table 4.11: Classification Accuracy of Test Images with Pareto Optimized Bi-Histogram Equalization

	Stage D	Stage E	Stage F	Stage G	Stage H
True Positive	421	419	402	398	413
False Positive	79	81	98	102	87
False Negative	51	80	108	118	87
Precision	0.842	0.838	0.804	0.796	0.826
Recall	0.892	0.840	0.789	0.771	0.826
F1 Score	0.867	0.839	0.796	0.783	0.826

In order to obtain the classification result, the test image need to be assigned to respective skeletal maturity stage from TW method prior performing automatic recognition using classifier. True positive count increases by one when the classifier result of maturity stage matches with pre-assigned maturity stage. For instance in Table 4.9, the number “378” of stage D means that there is 378 result out of 500 ossification sites test image that match the assigned or labelled maturity stage which is D. In other words, the classifier agree with the pre-assigned maturity stage for 378 labelled test images out of 500 test images as stage D. On the other hand, false positive is the number of test images

that have not matched the labelled result. For instance in Table 4.9, the false positive amounts to 122 representing that 122 of them, out of 500 images of stage D, have been recognized as various maturity stage except stage D. In other words, it is noted that the number of false negative amounts to the total number of test ossification sites (which are previously assigned as any maturity stage but not stage D) that have been wrongly classified as stage D.

As shown in Table 4.9 for test images without undergoing any enhancement, the F1 score of stage D, stage E, stage F, stage G and stage H are respectively calculated as 0.762, 0.715, 0.598, 0.587 and 0.641. As shown in Table 4.10, the F1 scores for images with different values of separation point and also the clipped limit for stage D, stage E, stage F, stage G and stage H are calculated respectively as 0.809, 0.746, 0.686, 0.646, and 0.697. From these two tables, we can compute the improvement by using bi-histogram equalization as a control set. It is observed that the improvement are recorded as 6.16%, 4.33%, 1.47%, 10.05%, and 8.73%, respectively. In other words, these are the improvement values before applying our proposed Pareto optimized bi-histogram equalization. So these improvement values reflected that even solely using bi-histogram equalization, there are improvement in all stages of maturity level. Table 4.11 reviews that the higher improvement value is obtained in the classification accuracy if the classifier is trained using image that have undergone the proposed Pareto Optimized Bi-Histogram Equalization. The improvement value is shown in Table 4.11. By comparing the F1 scores between Table 4.9 and Table 4.11, the improvement for all stages is significant with improvement values are recorded as 7.16%, 12.4%, 16.03%, 21.21% and 18.51%, respectively for stage D to stage H. It can be observed that for each stage, after using the Pareto optimized framework, the improvement is further magnified.

Figure 4.10 illustrates the F1 score for five stages from Stage D to Stage H where the classification accuracy is obtained from the test images without enhancement, enhancement using random value of separating point and clipped limit and finally with enhancement using Pareto-optimized bi-histogram equalization. From Figure 4.10, the F1 score of test images with the proposed contrast enhancement framework (Pareto-optimized bi-histogram equalization) is higher as compared to other test images with enhancement using random value of separating point and clipped limit and without enhancement.

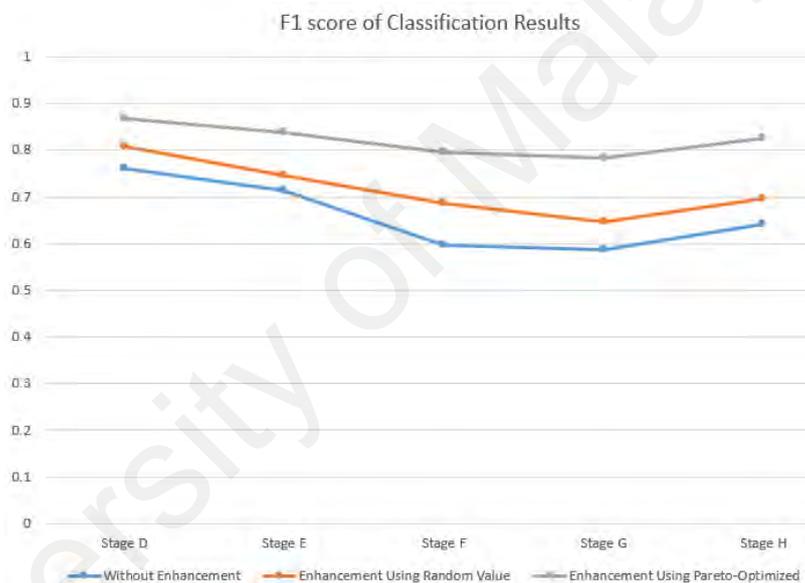


Figure 4.10 F1 score of classification accuracy without enhancement, enhancement using random value of separating point and clipped limit and enhancement using Pareto-optimized

CHAPTER 5: CONCLUSION AND FUTURE WORK

This chapter presents the contribution of this thesis, conclusion and suggestions for future research.

5.1 Thesis Contribution

The two principal contributions that devoted to the contrast enhancement and segmentation of this thesis are described as following:

- 1) The development of a contrast enhancement framework to enhance the visual quality of ossification sites of hand bone radiograph with the purpose to increase the accuracy of bone age assessment.
- 2) The development of a bone segmentation framework which adopts the Gaussian mapping method to assign different weights for each pixel according to bone texture characteristic to optimize the binary transformation operation.
- 3) The first attempt (to the best of our knowledge) of applying convolutional neural network in ossification sites delineation for bone age assessment purpose.
- 4) The design of overall framework of bone age assessment spatial based preprocessing to facilitate subsequent process of bone age assessment.

5.2 Conclusion

The present thesis presents the design and demonstration of procedures to improve the visual quality of the region of interest that is the ossification sites of hand bone radiographic image with the purpose to increase the accuracy of classifier in identifying the maturity stage. The objective of this visual quality improvement is to obtain holistic

result that take several objectives into considerations such as the contrast, brightness preservation and detail preservation, using bi-histogram equalization with two variables which are the separating points and the clipped limits. To achieve that, the selected variables values of the non-dominated solutions are obtained. These solutions are termed as Pareto solutions and the resultant images using these variables values are regarded as Pareto optimized image. Qualitative results have shown that the visual quality has been improved and the most excited part is that the important visual features that can be used to identify maturity stage has been emphasized in different images of the non-dominated solutions. Quantitative result supports the result of qualitative result in which the classifier trained by Pareto optimized images significantly outperform the classifier trained by image without enhancement and image with enhancement but without the Pareto optimization procedure. Thus we can conclude that the Pareto optimized solutions contribute in improving the performance of classifier.

5.3 Future Work

This research work can be extended to improve the bone age assessment in two main approaches. First approach is devoted to the framework design that can be adopted directly in the computer assisted bone age assessment system. The overall process of bone age assessment is enhanced by improving the segmentation performance of the hand bone prior to the bone age categorization of the segmented hand bone or ossification sites. The second approach is that the proposed enhancement framework on visual appearance of hand bone image can facilitate the inspection by doctors that are adopting GP methods in bone age assessment, in which the enhanced image contain sharper ossification sites that emphasize more on the pertinent information that matter to the bone age assessment. To further extend both capabilities in this approach, the designed approach will be tested on

different type of computer assisted system that have been applied in the hospital as well as conducting more visual tests with more pediatrics. One of the main tasks that can be performed in near future is to further complement the designed approach with the expertise knowledge and experiences from different pediatrics by modelling their thinking process and decision-making criteria into the proposed approach. The purpose is to reduce the assumptions and number of parameters.

In terms of algorithm execution in reducing the qualitative evaluation by experts, the number of manually determined parameters in the entire framework can be further minimized by using machine learning method. This can be done by collecting more data from different experts. Once the amount of data is sufficient, a modelling can be defined using current advancement of machine learning approach in order to automate the manual assignment of parameters. By the time the process is modelled by machine learning model, the qualitative evaluation by experts can be minimized.

In order to translate the current framework to real clinical environment, the algorithm should be verified with more data and the standardization of data format to the current computer assisted bone age assessment system must be performed. Incorporation of the proposed design algorithm into current computer assisted bone age assessment will be the main focus for near future attempt to accelerate the process of applying the research outputs to the real clinical environment. If the research output has been adopted by various experts in clinical settings, the further improvement that can be considered is the adaptive nature of the algorithm itself. It is crucial to equip the current algorithm with the ability to adapt to changes such as the change of the external factors that include various quality of image from new X-ray machines with different technologies. The adaptation also applied to different ethnic groups without much bias. Thus, it can be concluded that there are still many more to be done as future attempt to further improve the current approach.

However, the current performance of the approaches have shown us that the approach holds great potential to contribute in the area of bone age assessment.

University of Malaya

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Article in Academic Journals:

- L. K. Meng., H. Y. Chai, B. Pinguan-Murphy, M. I. M. Salim and K. W. Lai (2017) "Holistic contrast enhancement of carpals ossification sites for skeletal age assessment system." *The Journal of Engineering*, 2017(8): 479-494.
- L. K. Meng., K. Azira, N. Muhamad Hanif Ahmad, B. Pinguan-Murphy, H. Y. Chai, S. MahezaIrna Mohamad and K. W. Lai (2019). "Carpal bone segmentation using fully convolutional neural network." *Current Medical Imaging*, 15(10): 983 – 989.
- H. Y. Chai, L. K. Meng., H. Mohamed, H. H. Woon and K. W. Lai (2016). "Elimination of character-resembling anomalies within a detected region using density-dependent reference point construction in an automated license plate recognition system." *Journal of Electronic Imaging* 25(6): 1-11.

University of Malaya

REFERENCES

- Abadi, M., P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden and X. Zhang (2016). "TensorFlow: A system for large-scale machine learning."
- Abdullah-Al-Wadud, M., M. H. Kabir and O. Chae (2008). A spatially controlled histogram equalization for image enhancement. 23rd International Symposium on Computer and Information Sciences, ISCIS 2008.
- Abdullah-Al-Wadud, M., M. H. Kabir, M. A. A. Dewan and C. Oksam (2007). "A Dynamic Histogram Equalization for Image Contrast Enhancement." IEEE Transactions on Consumer Electronics 53(2): 593-600.
- Aicardi, G., M. Vignolo, S. Milani, A. Naselli, P. Magliano and P. Garzia (2000). "Assessment of skeletal maturity of the hand-wrist and knee: A comparison among methods." American Journal of Human Biology 12(5): 610-615.
- Alemán-Flores, M., L. Álvarez and V. Caselles (2007). "Texture-Oriented Anisotropic Filtering and Geodesic Active Contours in Breast Tumor Ultrasound Segmentation." Journal of Mathematical Imaging and Vision 28(1): 81-97.
- Arici, T., S. Dikbas and Y. Altunbasak (2009). "A Histogram Modification Framework and Its Application for Image Contrast Enhancement." IEEE Transactions on Image Processing 18(9): 1921-1935.
- Babalola, K. O., T. F. Cootes, C. J. Twining, V. Petrovic and C. Taylor (2008). 3D brain segmentation using active appearance models and local regressors. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 5241 LNCS: 401-408.
- Braunstein, E. M. C., P.; Buckwalter, K.; Bland, P.; Meyer, C.R. (1990). "Adaptive histogram equalization in digital radiography of destructive skeletal lesions." Radiography 166(3): 883-885.
- Bull, R. K., P. D. Edwards, P. M. Kemp, S. Fry and I. A. Hughes (1999). "Bone age assessment: a large scale comparison of the Greulich and Pyle, and Tanner and Whitehouse (TW2) methods." Archives of Disease in Childhood 81(2): 172-173.
- Cao, F., H. K. Huang, E. Pietka and V. Gilsanz (2000). "Digital hand atlas and web-based bone age assessment: system design and implementation." Computerized Medical Imaging and Graphics 24(5): 297-307.

- Chai, H. Y., T. T. Swee, G. H. Seng and L. K. Wee (2013). "Multipurpose contrast enhancement on epiphyseal plates and ossification centers for bone age assessment." *BioMedical Engineering OnLine* 12(1): 27.
- Chao, W. and Y. Zhongfu (2005). "Brightness preserving histogram equalization with maximum entropy: a variational perspective." *IEEE Transactions on Consumer Electronics* 51(4): 1326-1334.
- Chen Hee, O. and N. A. M. Isa (2010). "Adaptive contrast enhancement methods with brightness preserving." *IEEE Transactions on Consumer Electronics* 56(4): 2543-2551.
- Chen Hee, O. and N. A. M. Isa (2010). "Quadrants dynamic histogram equalization for contrast enhancement." *IEEE Transactions on Consumer Electronics* 56(4): 2552-2559.
- Chen Hee, O., N. S. P. Kong and H. Ibrahim (2009). "Bi-histogram equalization with a plateau limit for digital image enhancement." *IEEE Transactions on Consumer Electronics* 55(4): 2072-2080.
- Chen, L.-C., M. Collins, Y. Zhu, G. Papandreou, B. Zoph, F. Schroff, H. Adam and J. Shlens (2018). Searching for Efficient Multi-Scale Architectures for Dense Image Prediction. *Proceedings of the 32nd International Conference on Neural Information Processing Systems*: 8713-8724.
- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F. and Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. 2018 *The European Conference on Computer Vision (ECCV)*.
- Chen, S.-D. and A. Rahman Ramli (2004). "Preserving brightness in histogram equalization based contrast enhancement techniques." *Digital Signal Processing* 14(5): 413-428.
- Deb, K., A. Pratap, S. Agarwal and T. Meyarivan (2002). "A fast and elitist multiobjective genetic algorithm: NSGA-II." *IEEE Transactions on Evolutionary Computation* 6(2): 182-197.
- Faisal, A., S. Ng, S. L. Goh, J. George, E. Supriyanto and k. w. lai (2015). "Multiple LREK active contours for knee meniscus ultrasound image segmentation." *IEEE Transactions on Medical Imaging* 34.
- Falk, T., D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. MRRakchi, A. Böhm, J. Deubner, Z. Jaeckel, K. Seiwald, O. Dovzhenko, O. Tietz, C. Dal Bosco, S. Walsh, D. Saltukoglu, T. Tay, M. Prinz, K. Palme, M. Simons and O. Ronneberger

- (2019). "U-Net: deep learning for cell counting, detection, and morphometry." *Nature Methods* 16: 67-70.
- Gerstenfeld, L. C. (2003). "Endocrine Control of Skeletal Maturation: Annotations to Bone Age Readings " *Trends in Endocrinology & Metabolism* 15(1): 5-5.
- Gertych, A., E. Piętko and B. Liu (2007). "Segmentation of regions of interest and post-segmentation edge location improvement in computer-aided bone age assessment." *Pattern Analysis & Applications* 10(2): 115-123.
- Giordano, D., C. Spampinato, G. Scarciofalo and R. Leonardi (2010). "An automatic system for skeletal bone age measurement by robust processing of carpal and epiphysal/metaphysal bones." *IEEE Transactions on Instrumentation and Measurement* 59(10): 2539-2553.
- Girshick, R., J. Donahue, T. Darrell and J. Malik (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*: 580-587.
- Goldstein, T. and S. Osher (2009). "The Split Bregman Method for L1-Regularized Problems." *Siam Journal on Imaging Sciences* 2(2): 323-343.
- Gonzalez, R. and R. Woods (2002). *Digital Image Processing (2nd Edition)*, Prentice Hall.
- Gonzalez, R. and R. Woods (2007). *Digital Image Processing (3rd Edition)*, Prentice Hall.
- Gu, K., D. Tao, J. Qiao and W. Lin (2018). "Learning a No-Reference Quality Assessment Model of Enhanced Images With Big Data." *IEEE Transactions on Neural Networks and Learning Systems* 29(4): 1301-1313.
- Gu, K., W. Lin, G. Zhai, X. Yang, W. Zhang and C. W. Chen (2017). "No-Reference Quality Metric of Contrast-Distorted Images Based on Information Maximization." *IEEE Transactions on Cybernetics* 47(12): 4559-4565.
- Gu, Z., J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao and J. Liu (2019). CE-Net: Context Encoder Network for 2D Medical Image Segmentation. In *IEEE Transactions on Medical Imaging* 38(10): 2281-2292.
- He, K., G. Gkioxari, P. Dollár and R. Girshick (2017). Mask R-CNN. 2017 *IEEE International Conference on Computer Vision (ICCV)*.

- He, K., X. Zhang, S. Ren and J. Sun (2016). Deep Residual Learning for Image Recognition. In IEEE Conference on Computer Vision and Pattern Recognition: 770-778.
- Horé, A. and D. Ziou (2010). Image Quality Metrics: PSNR vs. SSIM. 2010 20th International Conference on Pattern Recognition.
- Hossain, F. and M. R. Alsharif (2007). Image Enhancement Based on Logarithmic Transform Coefficient and Adaptive Histogram Equalization. In International Conference on Convergence Information Technology: 1439-1444.
- Hsieh, C. W., T. C. Liu, T. L. Jong and C. M. Tiu (2010). "A fuzzy-based growth model with principle component analysis selection for carpal bone-age assessment." Medical and Biological Engineering and Computing 48(6): 579-588.
- Hsieh, C. W., T. L. Jong, Y. H. Chou and C. M. Tiu (2007). "Computerized geometric features of carpal bone for bone age estimation." Chinese Medical Journal 120(9): 767-770.
- Ibrahim, H. and N. S. P. Kong (2007). "Brightness Preserving Dynamic Histogram Equalization for Image Contrast Enhancement." IEEE Transactions on Consumer Electronics 53(4): 1752-1758.
- Jia, Y., E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama and T. Darrell (2014). "Caffe: Convolutional Architecture for Fast Feature Embedding." MM 2014 - Proceedings of the 2014 ACM Conference on Multimedia.
- Joseph, L. (2005). Computer-aided bone age assessment of children using a digital hand atlas. Retrieved from <https://ipilab.usc.edu/computer-aided-bone-age-assessment-of-children-using-a-digital-hand-atlas-2/>.
- Josephus, C. S. and S. Remya (2011). Multilayered contrast limited adaptive histogram equalization using frost filter. 2011 IEEE Recent Advances in Intelligent Computational Systems, RAICS 2011.
- Khan, K. M., B. S. Miller, E. Hoggard, A. Somani and K. Sarafoglou (2009). "Application of Ultrasound for Bone Age Estimation in Clinical Practice." Journal of Pediatrics 154(2): 243-247.
- Kim, M. and C. Min (2008). "Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement." IEEE Transactions on Consumer Electronics 54(3): 1389-1397.

- Kim, T. and J. Paik (2008). "Adaptive contrast enhancement using gain-controllable clipped histogram equalization." *IEEE Transactions on Consumer Electronics* 54(4): 1803-1810.
- Konak, A., D. W. Coit and A. E. Smith (2006). "Multi-objective optimization using genetic algorithms: A tutorial." *Reliability Engineering & System Safety* 91(9): 992-1007.
- Kong, N. S. P., H. Ibrahim, O. Chen Hee and D. C. J. Chieh (2009). Enhancement of Microscopic Images Using Modified Self-Adaptive Plateau Histogram Equalization. *International Conference on Computer Technology and Development, 2009. ICCTD '09.*
- Krizhevsky, A., I. Sutskever and G. Hinton (2012). "ImageNet Classification with Deep Convolutional Neural Networks." *Neural Information Processing Systems* 25.
- Land, E. H. and J. J. McCann (1971). "Lightness and Retinex Theory." *Journal of the Optical Society of America* 61(1): 1-11.
- Landry, D. J., D. E. RaSIDE and J. J. Vanhoutte (1979). "On the classification of congenital abnormalities from hand radiographs." *Pattern Recognition* 11(4): 289-296.
- Levitt, T. S., M. W. Hedgcock Jr, J. W. Dye, S. E. Johnston, V. M. Shadle and D. Vosky (1993). "Bayesian inference for model-based segmentation of computed radiographs of the hand." *Artificial Intelligence in Medicine* 5(4): 365-387.
- Lidong, H., Z. Wei, W. Jun and S. Zebin (2015). "Combination of contrast limited adaptive histogram equalisation and discrete wavelet transform for image enhancement." *IET Image Processing* 9(10): 908-915.
- Liu, C., L.-C. Chen, F. Schroff, H. Adam, W. Hua, A. L. Yuille and F.-F. Li (2019). Auto-DeepLab: Hierarchical Neural Architecture Search for Semantic Image Segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*: 82-92.
- Liu, J., J. Qi, Z. Liu, Q. Ning and X. Luo (2008). "Automatic bone age assessment based on intelligent algorithms and comparison with TW3 method." *Computerized Medical Imaging and Graphics* 32(8): 678-684.
- Liu, X., G. Cheung, X. Ji, D. Zhao and W. Gao (2019). "Graph-Based Joint Dequantization and Contrast Enhancement of Poorly Lit JPEG Images." *IEEE Transactions on Image Processing* 28(3): 1205-1219.

- Long, J., E. Shelhamer and T. Darrell (2015). Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- Mahmoodi, S., B. S. Sharif, E. Graeme Chester, J. P. Owen and R. Lee (2000). "Skeletal growth estimation using radiographic image processing and analysis." IEEE Transactions on Information Technology in Biomedicine 4(4): 292-297.
- Manos, G. K., A. Y. Cairns, I. W. Ricketts and D. Sinclair (1994). "Segmenting radiographs of the hand and wrist." Computer Methods and Programs in Biomedicine 43(3-4): 227-237.
- Manos, G., A. Y. Cairns, I. W. Ricketts and D. Sinclair (1993). "Automatic segmentation of hand-wrist radiographs." Image and Vision Computing 11(2): 100-111.
- Martin, D. D., D. Deusch, R. Schweizer, G. Binder, H. H. Thodberg and M. B. Ranke (2009). "Clinical application of automated Greulich-Pyle bone age determination in children with short stature." Pediatric Radiology 39(6): 598-607.
- Mentzel, H. J., C. Vilser, M. Eulenstein, T. Schwartz, S. Vogt, J. Böttcher, I. Yaniv, L. Tsoref, E. Kauf and W. A. Kaiser (2005). "Assessment of skeletal age at the wrist in children with a new ultrasound device." Pediatric Radiology 35(4): 429-433.
- Michael, D. J. and A. C. Nelson (1989). "HANDX: A model-based system for automatic segmentation of bones from digital hand radiographs." IEEE transactions on medical imaging 8(1): 64-69.
- Milletari, F., N. Navab and S.-A. Ahmadi (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In Fourth International Conference on 3D Vision: 565-571.
- Mohan, S. and M. Ravishankar (2013). Modified contrast limited adaptive histogram equalization based on local contrast enhancement for mammogram images. Communications in Computer and Information Science. 296 CCIS: 397-403.
- Morris, D. T. and C. F. Walshaw (1994). "Segmentation of the finger bones as a prerequisite for the determination of bone age." Image and Vision Computing 12(4): 239-245.
- Pietka, E., A. Gertych, S. Pospiech, F. Cao, H. K. Huang and V. Gilsanz (2001). "Computer-assisted bone age assessment: Image preprocessing and epiphyseal/metaphyseal ROI extraction." IEEE transactions on medical imaging 20(8): 715-729.

- Pizer, S. M., E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman and K. Zuiderveld (1987). "Adaptive histogram equalization and its variations." *Computer Vision, Graphics, and Image Processing* 39(3): 355-368.
- Prasoon, A., K. Petersen, C. Igel, F. Lauze, E. Dam and M. Nielsen (2013). Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolutional Neural Network. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*: 246-253.
- Przelaskowski, A. (2008). *Computer-Aided Diagnosis: From Image Understanding to Integrated Assistance*. Information Technologies in Biomedicine. E. Pietka and J. Kawa, Springer Berlin / Heidelberg. 47: 44-54.
- Qian, X., J. Wang, S. Guo and Q. Li (2013). "An active contour model for medical image segmentation with application to brain CT image." *Medical physics* 40: 021911.
- Qing, W. and R. K. Ward (2007). "Fast Image/Video Contrast Enhancement Based on Weighted Thresholded Histogram Equalization." *IEEE Transactions on Consumer Electronics* 53(2): 757-764.
- Rajavel, P. (2010). "Image dependent brightness preserving histogram equalization." *IEEE Transactions on Consumer Electronics* 56(2): 756-763.
- Ren, S., K. He, R. Girshick and J. Sun (2016). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39(6): 1137-1149.
- Rucci, M., G. Coppini, I. Nicoletti, D. Cheli and G. Valli (1995). "Automatic analysis of hand radiographs for the assessment of skeletal age: A subsymbolic approach." *Computers and Biomedical Research* 28(3): 239-256.
- Sang-Yeon, K., H. Dongil, C. Seung-Jong and P. Jong-Seok (1999). "Image contrast enhancement based on the piecewise-linear approximation of CDF." *IEEE Transactions on Consumer Electronics* 45(3): 828-834.
- Sato, K., K. Ashizawa, M. Anzo, F. Otsuki, S. Kaneko, T. Tanaka, K. Tsukagoshi, A. Nimura, H. Matsuoka, N. Matsuo, H. Mitani and M. Murata (1999). "Setting up an automated system for evaluation of bone age." *Endocrine Journal* 46(SUPPL.): S97-S100.
- Satoh, M. (2015). "Bone age: assessment methods and clinical applications." *Clinical Pediatric Endocrinology* 24(4): 143-152.

- Sazonova, N. and S. Schuckers (2010). "Fast and efficient iris image enhancement using logarithmic image processing." Proc SPIE 7667.
- Sebastian, T. B., H. Tek, J. J. Crisco and B. B. Kimia (2003). "Segmentation of carpal bones from CT images using skeletally coupled deformable models." Medical Image Analysis 7(1): 21-45.
- Se-Hwan, Y., K. Jin Heon and K. Suki (2010). "Image enhancement using a fusion framework of histogram equalization and laplacian pyramid." IEEE Transactions on Consumer Electronics 56(4): 2763-2771.
- Se-Hwan, Y., K. Jin Heon and K. Suki (2011). Contrast enhancement using a weighted Histogram Equalization. IEEE International Conference on Consumer Electronics (ICCE).
- Sengee, N. and C. Heung (2008). "Brightness preserving weight clustering histogram equalization." IEEE Transactions on Consumer Electronics 54(3): 1329-1337.
- Sengee, N., A. Sengee and C. Heung-Kook (2010). "Image contrast enhancement using bi-histogram equalization with neighborhood metrics." IEEE Transactions on Consumer Electronics 56(4): 2727-2734.
- Sengee, N., B. Bazarragchaa, K. Tae Yun and C. Heung Kook (2009). Weight Clustering Histogram Equalization for Medical Image Enhancement. IEEE International Conference on Communications Workshops.
- Shaziya, H., K. Shyamala and R. Zaheer (2018). Automatic Lung Segmentation on Thoracic CT Scans Using U-Net Convolutional Network. 2018 International Conference on Communication and Signal Processing (ICCSP).
- Sim, K. S., C. P. Tso and Y. Y. Tan (2007). "Recursive sub-image histogram equalization applied to gray scale images." Pattern Recognition Letters 28(10): 1209-1221.
- Simonyan, K. and A. Zisserman (2014). "Very Deep Convolutional Networks for Large-Scale Image Recognition." arXiv 1409.1556.
- Singh, T. and C. S. Vikram (2002). "Histogram equalization of vibration fringes in holography." Optik - International Journal for Light and Electron Optics 113(11): 499-503.
- Soong-Der, C. and A. R. Ramli (2003). "Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation." IEEE Transactions on Consumer Electronics 49(4): 1301-1309.

- Soong-Der, C. and A. R. Ramli (2003). "Minimum mean brightness error bi-histogram equalization in contrast enhancement." *IEEE Transactions on Consumer Electronics* 49(4): 1310-1319.
- Srinivas, N. and K. Deb (1994). "Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms." *Evolutionary Computation* 2(3): 221-248.
- Sun, X., Q. Xu and L. Zhu (2019). "An Effective Gaussian Fitting Approach for Image Contrast Enhancement." *IEEE Access* 7: 31946-31958.
- Sundaram, M., K. Ramar, N. Arumugam and G. Prabin "Histogram Modified Local Contrast Enhancement for Mammogram Images." *Applied Soft Computing* 11(8): 5809-5816
- Tanner, J. M. (1994). "Automatic Bone Age Measurement Using Computerized Image Analysis." *Journal of Pediatric Endocrinology and Metabolism* 7(2): 141-146.
- Tanner, J. M., D. Oshman, G. Lindgren, J. A. Grunbaum, R. Elsouki and D. Labarthe (1994). "Reliability and validity of computer-assisted estimates of tanner-whitehouse skeletal maturity (CASAS): Comparison with the manual method." *Hormone Research in Paediatrics* 42(6): 288-294.
- Tanner, J. M., M.J.R. Healy, H. Goldstein, and N. Cameron (2001). *Assessment of skeletal maturity and prediction of adult height (TW3 method) (3rd ed.)*. London: W.B. Saunders.
- Thodberg, H. H. (2009). "An Automated Method for Determination of Bone Age." *J Clin Endocrinol Metab* 94(7): 2239-2244.
- Thodberg, H. H. and D. D. Martin (2016). "Comment on 'The evidential value of developmental age imaging for assessing age of majority' by Cole, *Annals of Human Biology*, 2015." *Annals of Human Biology* 43(6): 577-578.
- Thodberg, H. H. and L. Sävendahl (2010). "Validation and reference values of automated bone age determination for four ethnicities." *Academic Radiology* 17(11): 1425-1432.
- Thodberg, H. H., J. Böttcher, J. Lomholt, S. Kreiborg, G. Wolf and A. Pfeil (2016). "A new implementation of digital X-ray radiogrammetry and reference curves of four indices of cortical bone for healthy European adults." *Archives of Osteoporosis* 11(1).
- Thodberg, H. H., R. R. van Rijn, O. G. Jenni and D. D. Martin (2017). "Automated determination of bone age from hand X-rays at the end of puberty and its

applicability for age estimation." *International Journal of Legal Medicine* 131(3): 771-780.

Thodberg, H. H., S. Kreiborg, A. Juul and K. D. Pedersen (2009). "The BoneXpert method for automated determination of skeletal maturity." *IEEE Transactions on Medical Imaging* 28(1): 52-66.

Thodberg, K., L. U. Sørensen, J. W. Christensen, P. H. Poulsen, B. Houbak, V. Damgaard, I. Keseler, D. Edwards and P. B. Videbech (2016). "Therapeutic effects of dog visits in nursing homes for the elderly." *Psychogeriatrics* 16(5): 289-297.

Thodberg, K., L. U. Sørensen, P. B. Videbech, P. H. Poulsen, B. Houbak, V. Damgaard, I. Keseler, D. Edwards and J. W. Christensen (2016). "Behavioral responses of nursing home residents to visits from a person with a dog, a robot seal or a toy cat." *Anthrozoos* 29(1): 107-121.

Thomas, G., D. Flores-Tapia and S. Pistorius (2010). Fast digital image contrast enhancement. *IEEE Instrumentation and Measurement Technology Conference (I2MTC)*.

Thomas, G., D. Flores-Tapia and S. Pistorius (2011). "Histogram Specification: A Fast and Flexible Method to Process Digital Images." *IEEE Transactions on Instrumentation and Measurement* 60(5): 1565-1578.

Tian, Y., Q. Wan and F. Wu (2007). Local Histogram Equalization Based on the Minimum Brightness Error. *Fourth International Conference on Image and Graphics (ICIG 2007)*.

Tsai, C.-M. (2013). "Adaptive Local Power-Law Transformation for Color Image Enhancement." *Applied Mathematics & Information Sciences* 7: 2019-2026.

Verma, D., T. Peltomäki and A. Jäger (2009). "Reliability of growth prediction with hand-wrist radiographs." *The European Journal of Orthodontics* 31(4): 438-442.

Waktola, S., K. Grudzien and L. Babout (2019). Stagnant Zone Segmentation with U-Net. *2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*.

Wang, S. and G. Luo (2018). "Naturalness Preserved Image Enhancement Using a Priori Multi-Layer Lightness Statistics." *IEEE Transactions on Image Processing* 27(2): 938-948.

- Wang, S., K. Ma, H. Yeganeh, Z. Wang and W. Lin (2015). "A Patch-Structure Representation Method for Quality Assessment of Contrast Changed Images." *IEEE Signal Processing Letters* 22(12): 2387-2390.
- William W Greulich, S. I. P. (1959). "Radiographic Atlas of Skeletal Development of the Hand and Wrist." Oxford University Press, London.
- Wongsritong, K., K. Kittayaruasiriwat, F. Cheevasuvit, K. Dejhan and A. Somboonkaew (1998). Contrast enhancement using multipeak histogram equalization with brightness preserving. *IEEE Asia-Pacific Conference on Circuits and Systems*.
- Yang, T. and J. Song (2018). An Automatic Brain Tumor Image Segmentation Method Based on the U-net. 2018 IEEE 4th International Conference on Computer and Communications (ICCC).
- Yeong-Taeg, K. (1997). "Contrast enhancement using brightness preserving bi-histogram equalization." *IEEE Transactions on Consumer Electronics* 43(1): 1-8.
- Yeong-Taeg, K. (1997). Quantized bi-histogram equalization. *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997*
- Yeong-Taeg, K. (1997). Quantized bi-histogram equalization. *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997*.
- Yu, W., C. Qian and Z. Baeomin (1999). "Image enhancement based on equal area dualistic sub-image histogram equalization method." *IEEE Transactions on Consumer Electronics* 45(1): 68-75.
- Yue, H., J. Yang, X. Sun, F. Wu and C. Hou (2017). "Contrast Enhancement Based on Intrinsic Image Decomposition." *IEEE Transactions on Image Processing* 26(8): 3981-3994.
- Zhang, L. and L. Xu (2018). An Automatic Liver Segmentation Algorithm for CT Images U-Net with Separated Paths of Feature Extraction. 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC).
- Zhang, Y., W. Chen, Y. Chen and X. Tang (2018). A post-processing method to improve the white matter hyperintensity segmentation accuracy for randomly-initialized U-net. 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP).

Zhou, W. and A. C. Bovik (2002). "A universal image quality index." *IEEE Signal Processing Letters* 9(3): 81-84.

Zhou, W., A. C. Bovik, H. R. Sheikh and E. P. Simoncelli (2004). "Image quality assessment: from error visibility to structural similarity." *IEEE Transactions on Image Processing* 13(4): 600-612.

Zitzler, E. and L. Thiele (1999). "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach." *IEEE Transactions on Evolutionary Computation* 3(4): 257-271.

Zoph, B., V. Vasudevan, J. Shlens and Q. Le (2017). "Learning Transferable Architectures for Scalable Image Recognition." arXiv 1707.07012.

Zuo, C., Q. Chen and X. Sui (2013). "Range Limited Bi-Histogram Equalization for image contrast enhancement." *Optik* 124(5): 425-431.

University of Malaya