AUTOMATIC ASSESSMENT OF SCOLIOSIS USING 3D ULTRASOUND IMAGING AND CONVOLUTIONAL NEURAL NETWORK

by Sunetra Banerjee

Thesis submitted in fulfilment of the requirements for
the degree of

Doctor of Philosophy

Under the supervision of

Dr. Steve Ling (Supervisor)
Dr. Steven Su (Co-Supervisor)
Professor Yong-Ping Zheng (External Supervisor)

University of Technology Sydney
Faculty of Engineering and Information Technology

March 2022, Sydney

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Sunetra Banerjee, declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Production Note:
Signature removed
prior to publication.

Sunetra Banerjee

18 March 2022

DEDICATION

This thesis is dedicated to my parents
and my husband.
Their selfless love, support and care
makes me and my life complete.

ACKNOWLEDGMENTS

गुरुर्ब्रह्मा गुरुर्विष्णु गुरुर्देवो महेश्वरः |

गुरु साक्षात परब्रह्मा तस्मै श्रीगुरवे नमः ||

- Guru Mantra

Guru creates, sustains knowledge and destroys the weeds of ignorance. I salute such a Guru.

First, I would like to express my sincere gratitude to my supervisor Dr. Steve Ling for his invaluable supervision, support, and tutelage in all the time of my academic research. His mentorship and guidance made my research experience more enriching. His bottomless patience and immense understanding have kept me motivated throughout my Ph.D.

I am extremely thankful to the extremely knowledgeable and kind-hearted Dr. Steven Su, my co-supervisor, for his valuable suggestions, thoughtful guidance, and support throughout my Ph.D. studies.

This research work would not be possible without the mentorship of Prof. Yongping Zheng, my external supervisor. Throughout my research, I was able to continuously draw from his vast ocean of knowledge. His guidance at every step of my research, from inception, data collection, and publication, was truly inspirational and I wish to express my deepest gratitude to him.

My earnest appreciation goes to Dr. Frank H.F. Leung for his valuable feedback in improving and structuring my research and publications. I wish to extend my heartfelt gratitude to my dear friends and colleagues, Ms. Juan Lyu and Mr. Zixun Huang. Their support and collaboration made my research work more complete. I wish to thank all the team members of the biomedical lab at Hong Kong Polytechnic University who contributed to this research.

Financially and facility-wise, the support I got from UTS was instrumental. I wish to express my profound gratitude to UTS for awarding me the Research Training Program Scholarship (RTPS) for the duration of my Ph.D. I also like to extend my appreciation to all the faculty members and fellow research scholars of the School of Biomedical Engineering, UTS.

I am indebted to my parents whose constant love, care, pain, and sacrifice make me who I am today. Their constant support keeps me motivated and confident. I owe a great deal to the wonderful person in my life, my husband, who lived by my side every moment

of my Ph.D. journey and without whom I would not have had the poise to walk through this path of Ph.D. I immensely respect his effort in editing and formatting my research papers as well as my thesis. He made me also realize that '*behind every successful woman, there is a man'*.

My deepest reverence goes to my grandfather and parents-in-law for their moral encouragement and support, in any way they can do through this Ph.D. journey.

Last, but not least, I express my deepest gratitude to all my family members for their endearing encouragement and love.

ABSTRACT

Scoliosis is a 3D deformation of the spine categorized by a spine curvature angle larger than 10° in the coronal plane. During the adolescent period, as the skeletal structure of a child develops, a significant rate of curve progression may occur in the spine; sometimes can be as high as up to 10° per year. This medical condition is called idiopathic scoliosis and may originate from childhood i.e. the population from age 10 to 16 years. If left untreated, adolescence idiopathic scoliosis (AIS) can have a permanent damaging effect on the skeletal structure of adolescents and can aggravate chronic back pain and thoracic insufficiency syndrome. Severe scoliosis also may affect the heart, lungs, and nerve system. Therefore, early detection of scoliosis, through periodic monitoring of spine curve progression, until the complete maturity of skeletal structure, is necessary for AIS patients. During the whole course of scoliosis treatment, a patient may be exposed for up to 25 sessions.

For monitoring the spine curve progression, a typical session involves scanning the patient in the standing posture by using a suitable imaging modality and estimating the spine curvature angle. The most widely used measurement to quantify the magnitude of spinal deformities on plain radiographs is known as the Cobb angle. The traditional Cobb angle measurement is manual and involves indicating the most tilted upper and lower end vertebrae and estimating the angle between intersecting lines drawn perpendicular to the top of the upper vertebrae and the bottom of the lower vertebrae. Several X-ray examinations are needed during the entire monitoring process. Frequent X-ray examinations pose a radiation threat and may increase the risk of breast cancer in girls, prostate cancer in boys, and the overall risk of leukemia for AIS patients.

Ultrasound imaging is a real-time, inexpensive, radiation-free, and highly portable imaging modality which makes it more safe and accessible for AIS patients than other imaging modalities. This modality works on the principle of capturing the reflected ultrasound from the cortical surface of the internal organs and mapping their topographical information. A newly developed 3D ultrasound system called the Scolioscan system has been proven reliable for scoliosis diagnosis and as efficient as the conventional radiographic Cobb method.

Adolescence idiopathic scoliosis (AIS) requires periodic monitoring of patients for continuous observation of small changes in spine curvature. The application of ultrasound

imaging in the assessment of scoliosis could reduce radiation exposure risk for young adults. Previously, spinous process angles (SPA) were measured from the Scolioscan ultrasound images, using the VPI-SP method, where the spinal column profile was taken as an anatomical reference. However, these measurement leads to an underestimation of the traditionally used Cobb angles. Ultrasound curve angle (UCA), on the other hand, is an alternative and reliable technique of scoliosis assessment, where more lateral features are used for angle measurement. The manual measurement of UCA is currently in practice which requires a lot of time and human judgment. The objective of this dissertation is to automate the UCA measurement and thereby reducing the human intervention and making the process faster.

The Scolioscan system (Model SCN801, Telefield Medical Imaging Ltd), developed in Hong Kong, is used to generate 3D volume projection image (VPI) using the conventional 3D ultrasound imaging technique. 109 images, collected from 109 patients using the Scolioscan system (82 females and 27 males), with an average age of 15.6 ± 2.7 years and different degrees of spine deformities, are used in this research retrospectively.

The overall sequence of the proposed automatic UCA measurement process is as follows: (a) Collection of spine ultrasound images, (b) Segmentation of lateral bony features & (c) Key feature selection, and (d) Finding the most tilted upper and lower lateral features and calculation of UCA. The topic of this dissertation is covered by novel techniques for automatic scoliosis assessment.

The main contributions of this research are: (a) Segmentation technique 1: Introduction of Light-Convolution Dense Selection U-Net (LDS U-Net) for ultrasound spine bony feature segmentation, (b) Segmentation technique 2: Introduction of multi-scale feature fusion Skip-Inception U-Net (SIU-Net) for ultrasound spine image segmentation, (c) Conversion of binary segmented images to RGB images with specific colour codes for key feature selection & (d) Development of Centroid Pairing and Inscribed rectangle Slope (CPI-SLO) method to find the most tilted bony features and thereby calculate the UCA.

The performance evaluation of this research shows that this automatic method can produce results comparable to manual measurements of UCA. The advantage of the automatic method is that it overcomes subjective manual annotation, reduces the human error in measurements, and improves the scalability of the system. Also, frequent monitoring of curve progression for AIS patients is safely possible by this technique as no harmful radiation is present in this imaging process. Hence, the automatic UCA measurement technique using 3D ultrasound imaging modality is robust, fast, economic, and suitable for mass diagnosis of scoliosis measurement in the adolescent population.

LIST OF PUBLICATIONS

The contents of this thesis are based on the following papers that have been published, accepted, or submitted to peer-reviewed journals and conferences.

International Journal Papers

1. **Banerjee, S.;** Lyu, J.; Huang, Z.; Leung, H.F.F.; Lee, T.T.-Y.; Yang, D.; Su, S.; Zheng, Y.; Ling, S.-H. **'Light-Convolution Dense Selection U-Net (LDS U-Net) for Ultrasound Lateral Bony Feature Segmentation'**. *Appl. Sci.* 2021, *11*, 10180.
2. **Banerjee, S.;** Lyu, J.; Huang, Z.; Leung, H.F.F.; Lee, T.T.-Y.; Yang, D.; Su, S.; Zheng, Y.; Ling, S.-H., **'Ultrasound Spine Image Segmentation Using Multi-Scale Feature Fusion Skip-Inception U-Net (SIU-Net)',** in Biocybernetics and Biomedical Engineering, Elsevier, Volume 42, Issue 1, January–March 2022, Pages 341-361.
3. Lyu, J., Bi, X., **Banerjee, S**., Huang, Z., Leung, F.H., Lee, T.T.-Y., Yang, D.-D., Zheng, Y.-P. & Ling, S.H. 2021, **'Dual-task ultrasound spine transverse vertebrae segmentation network with contour regularization'**, Computerized Medical Imaging and Graphics, vol. 89, p. 101896.
4. Lyu, J., Ling, S.H., **Banerjee, S.**, Zheng, J., Lai, K., Yang, D., Zheng, Y., Bi, X., Su, S. & Chamoli, U. 2021, **'Ultrasound volume projection image quality selection by ranking from convolutional RankNet'**, Computerized Medical Imaging and Graphics, vol. 89, p. 101847.
5. Huang, Z.; Zhao, R.; Leung, H.F.F.; **Banerjee, S**.; Lee, T.T.-Y.; Yang, D.; Lun, D.P.K.; Lam, K-M.; Ling, S.-H., Zheng, Y., **'Joint Spine Segmentation and Noise Removal from Ultrasound Volume Projection Images with Selective Feature Sharing'**, IEEE Transactions on Medical Imaging (2022).

International Conference Papers

1 **S. Banerjee**, S. H. Ling, J. Lyu, S. Su, and Y.-P. Zheng, '**Automatic Segmentation of 3D Ultrasound Spine Curvature Using Convolutional Neural**

Network', in 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020, pp. 2039-2042: IEEE.

2 **S. Banerjee**, S. H. Ling, U. Chamoli, and S. Su, '**Skip-Inception U-Net For Bony Feature Segmentation Using Ultrasound Imaging For Scoliosis Assessment'**, in Sydney Spinal Virtual Symposium 2020, Sydney, Australia, Sep. 2020.

3 J. Lyu, S.H. Ling, **S. Banerjee**, J.J.Y. Zheng, K.L. Lai, D. Yang, Y.P. Zheng, S. Su, '**3D ultrasound spine image selection using convolutional learning-to-rank algorithm'** in Engineering and Medicine Biology Society (EMBC), 2019 41st Annual International Conference of the IEEE. IEEE, 2019, pp. 4799-4802.

4 Z. Huang, L.W. Wang, F.H. Leung, S. Banerjee, D. Yang, T. Lee, J. Lyu, S.H. Ling, Y.-P. Zheng, '**Bone Feature Segmentation in Ultrasound Spine Image with Robustness to Speckle and Regular Occlusion Noise**', in 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, pp. 1566-71.

5 Z. Huang, R. Zhao, F. H. F. Leung, K.M. Lam, S.H. Ling, J. Lyu, **S. Banerjee**, T. Lee, D. Yang, Y.P. Zheng2 '**DA-GAN: Learning Structured Noise Removal In Ultrasound Volume Projection Imaging For Enhanced Spine Segmentation**', in 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), IEEE, pp. 770-4.

6 Y. An, T. Hu, J. Wang, J. Lyu, **S. Banerjee**, S.H. Ling**, 'Lung Nodule Classification using A Novel Two-stage Convolutional Neural Networks Structure'**, in Engineering and Medicine Biology Society (EMBC), 2019 41st Annual International Conference of the IEEE. IEEE, 2019, pp. 6259-6262.

TABLE OF CONTENTS

Certificate of original authorship i

Dedication ii

Acknowledgments iii

Abstract v

List of Publications viii

International Journal Papers viii

International Conference Papers viii

TABLE OF CONTENTS x

List of Figures xv

List of Tables xix

List of Acronyms xxi

CHAPTER I: INTRODUCTION 1

1.1 Background 1

1.1.1 Scoliosis 1

1.1.2 Conventional detection of scoliosis 2

1.2 Motivation 3

1.2.1 Measurement of scoliosis progression using X-ray 4

1.2.2 Measurement of scoliosis progression using Scolioscan 4

1.3 Research questions 6

1.4 Aim of the research 6

1.5 Research objectives 6

1.6 Research contributions 7

1.7 Thesis framework 8

CHAPTER II: LITERATURE REVIEW 11

2.1 Introduction 11

2.2 Imaging modalities for scoliosis detection 12

2.2.1 Radiography 12

2.2.2 Magnetic Resonance Imaging (MRI) 14

2.2.3 Computed Tomography (CT) 15

2.2.4 Ultrasound 15

2.3 Scoliosis assessment parameters 22

2.3.1 Cobb angle 22

2.3.2 Spinous Process Angle (SPA) 23

2.3.3 Ultrasound Curve Angle (UCA) 23

2.4 Scoliosis assessment methods 24

2.4.1 Manual method 24

2.4.2 Semi-automatic method 26

2.4.3 Automatic method 27

2.5 Application of Convolutional Neural Network on ultrasound image 30

CHAPTER III: GENERATION AND COLLECTION OF ULTRASOUND SPINE IMAGES 34

3.1 Introduction 34

3.1.1 B-Mode ultrasound imaging 34

3.1.2 3D B-Mode ultrasound imaging 36

3.2 Manual assessment of scoliosis using Scolioscan 36

3.2.1 The Scolioscan system 37

3.3 Moving towards automatic assessment of scoliosis 40

3.3.1 The ultrasound image dataset and pre-processing 40

3.3.2 Generation of 2D VPI images 41

3.4 Challenges in processing spine ultrasound image 42

3.4.1 Scan noise 42

3.4.2 Variability of bony features 43

3.5 Identifying ROIs from ultrasound spine image 44

3.5.1 Need for a suitable segmentation architecture 45

CHAPTER IV: LIGHT-CONVOLUTION DENSE SELECTION U-NET (LDS U-NET) FOR ULTRASOUND LATERAL BONY FEATURE SEGMENTATION 46

4.1 Introduction 46

4.1.1 Medical image segmentation using CNN 46

4.1.2 Approach to designing a suitable architecture 48

4.2 Methodology of LDS U-Net 49

4.2.1 Depthwise separable convolution 50

4.2.2 Light dense block 51

4.2.3 Multi-scale path 52

4.2.4 Selection gate 53

4.2.5 Ablation study 55

4.3 Experimental setup 55

4.3.1 Dataset 55

4.3.2 Pre- and post-processing & data augmentation 56

4.3.3 Implementation details 56

4.3.4 Evaluation metrics 57

4.4 Analysis of performance of LDS U-Net 57

4.4.1 Evaluation of performance of key features 58

4.4.2 Comparison of LDS U-Net with other contemporary models 59

4.5 Discussion on LDS U-Net 63

4.6 Key takeaways from LDS U-Net 65

4.7 Areas of improvement of LDS U-Net 66

CHAPTER V: ULTRASOUND SPINE IMAGE SEGMENTATION USING MULTI-SCALE FEATURE FUSION SKIP-INCEPTION U-NET (SIU-NET) 68

5.1 Introduction 68

5.1.1 Re-visiting the capabilities of contemporary CNN architectures to handle variations 68

5.1.2 Approach to re-designing a novel architecture to better handle variabilities and noise 70
5.2 Methodology of SIU-Net 71
5.2.1 Inception block 71
5.2.2 Dense-skip connection 73
5.2.3 Ablation study 74
5.3 Experimental setup 75
5.3.1 Dataset 75
5.3.2 Implementation details 75
5.3.3 Evaluation metrics 76
5.4 Analysis of performance of SIU-Net 76
5.4.1 SIU-Net outperforms U-Net, UNet++, and MultiResUNet in the segmentation of ultrasound spine image dataset 77
5.4.2 SIU-Net performs best for bony feature edge detection 77
5.4.3 SIU-Net performs the best identification of LBF 78
5.4.4 SIU-Net is most capable of identifying TBF 79
5.4.5 SIU-Net is more consistent than MultiResUNet and UNet++ 80
5.4.6 Further examples of qualitative and quantitative comparison of SIU-Net with other models 81
5.4.7 Comparison of accuracy of bony feature detection 81
5.5 Discussion on SIU-Net 82
5.6 Key takeaways of SIU-Net 91
CHAPTER VI: AUTOMATIC CALCULATION OF ULTRASOUND CURVE ANGLE 92
6.1 Introduction 92
6.2 Selection of the most suitable segmentation technique for subsequent angle calculation 93
6.2.1 Quantitative evaluation of various segmentation techniques 93

6.2.2 Evaluation of various segmentation techniques for usability......................95
6.3 Demarcation of Regions of Interest..96
6.3.1 Methodology: binary to RGB conversion of segmented images 97
6.4 Automatic calculation of Scoliosis angle (UCA)..99
6.4.1 Identification of vertebrae features..100
6.4.2 Measurement of main thoracic angle 100
6.4.3 Measurement of thoraco-lumbar angle...105
6.5 Experiments 106
6.5.1 Dataset..106
6.5.2 Implementation Details 107
6.5.3 Analysis of the performance of CPI-SLO with various segmentation methods..107
6.6 Results...108
6.6.1 Performance of CPI-SLO compared with conventional manual UCA measurement...108
6.7 Discussion & analysis...111
6.8 Conclusion 114
Chapter VII: CONCLUSIONS & FUTURE WORK 116
7.1 Research summary 116
7.2 Research contributions..117
7.3 Recommendations for future studies..119
References..121

LIST OF FIGURES

FIG 1.1. COBB ANGLE 2
FIG 1.2. ANGLE MEASUREMENT TECHNIQUES FOR ULTRASOUND IMAGING MODALITY- (A) SPA (B) UCA.. 5
FIG 1.3. ANATOMY OF AN ULTRASOUND SPINE IMAGE (A) INPUT IMAGE, (B) TYPES OF BONY FEATURE IDENTIFIED (C) ANGLE MEASUREMENT SCHEME 5
FIG 1.4. A COMPLETE FLOWCHART FOR AUTOMATIC UCA MEASUREMENT.. 9
FIG 2.1. CATEGORIZATION OF SCOLIOSIS 11
FIG 2.2. TREATMENT OF SCOLIOSIS.. 12
FIG 2.3. X-RAY SPINE IMAGE 13
FIG 2.4. MRI SPINE IMAGE 14
FIG 2.5. CT SPINE IMAGE 15
FIG 2.6. ULTRASOUND SPINE IMAGE 16
FIG 2.7. SPECKLE NOISE 20
FIG 2.8. SCOLIOSIS ASSESSMENT PARAMETERS 22
FIG 2.9. X-RAY COBB ANGLE MEASUREMENT 22
FIG 2.10. SPA MEASUREMENT 23
FIG 2.11. (A) UCA MEASUREMENT WITH ULTRASOUND IMAGE AND (B) COBB ANGLE MEASUREMENT WITH X-RAY IMAGE 23
FIG 3.1. TYPES OF ULTRASOUND IMAGING.. 34
FIG 3.2. B-MODE ULTRASOUND FOETUS SPINE IMAGE.. 34
FIG 3.3. FLOWCHART OF CURRENT (MANUAL) SCOLIOSIS ASSESSMENT TECHNIQUE USING SCOLIOSCAN SYSTEM 36
FIG 3.4. A SCOLIOSCAN SYSTEM.. 37
FIG 3.5. FOUR SUPPORTERS HELPING PATIENT MAINTAIN A STABLE POSTURE; THE SUPPORTS ARE ADJUSTED AGAINST THE CHEST AND HIP BOARDS, AND THEIR LENGTHS ARE ADJUSTED ACCORDINGLY 37
FIG 3.6. SCANNING PROCEDURE WITH THE SCOLIOSCAN PROBE, ULTRASOUND GEL IS APPLIED ON THE AREA TO BE SCREENED 37
FIG 3.7. SOFTWARE INTERFACES FOR (A) SCANNING AND (B) VPI IMAGE ANALYSIS AND ANGLE MEASUREMENT OF SCOLIOSCAN 39

FIG 3.8. (A) 4 SPINE VPI IMAGES SHOWING THE CORONAL PLANE WITH DIFFERENT LEVELS OF SPINE DEFORMITY WITH THE TWO LINES MANUALLY DRAWN (B,C,D) THE THREE LINES TO FORM TWO PAIRS TO MEASURE THE CURVATURES OF SPINE IN THE THORACIC AND LUMBAR REGIONS, RESPECTIVELY 40
FIG 3.9. THE SAMPLE OF NINE 2D ULTRASOUND IMAGES IN DIFFERENT DEPTHS OF A 3D SPINE IMAGE 41
FIG 3.10. NARROW-BAND, NON-PLANAR VOLUME RENDERING ALGORITHM FOR 2D IMAGE GENERATION 42
FIG 3.11. EXAMPLE OF US SCAN NOISE. (A) 3D VPI IMAGE ACQUIRED FROM MANUAL SCAN. THE RED FRAMES SHOWS MULTIPLE B-MODE IMAGES, WHO PILED UP TO FORM 3D IMAGE, THE BLUE AND YELLOW FRAMES SHOW THE CORONAL AND SAGITTAL IMAGES RESPECTIVELY. (B) GREEN DOTTED LINES SHOWS THE MISSING B-MODE IMAGE, WHICH RESULTS THE SCAN NOISE, (C) CORONAL PROJECTED IMAGE WITH SCAN NOISE (NOISY AREA IN GREEN BOX) 43
FIG 3.12. AN EXAMPLE OF ULTRASOUND SCAN NOISE. GENERATION OF VPI IMAGE BY NON-PLANAR VOLUME RENDERING FROM 3D US VOLUME DATA. (A) ORIGINAL VPI IMAGE WITH SCAN NOISE, (B) VPI IMAGE WITHOUT SCAN NOISE, (C) SCAN NOISE AND CORRESPONDING CLEAR AREAS IN BLUE AND GREY BOXES 43
FIG 3.13. (A), (B), AND (C): LATERAL BONY FEATURES ARE OF DIFFERENT SHAPES, SIZES, AND LOCATIONS, (D), (E), AND (F): SLOPES BETWEEN TWO ADJACENT BONY FEATURES (THORACIC) ARE DIFFERENT. 44
FIG 3.14. ULTRASOUND CURVE ANGLE (UCA) MEASUREMENT (A) ORIGINAL ULTRASOUND SPINE IMAGE, (B) MARKED LATERAL BONY FEATURES WITH THREE TYPES OF ANATOMICAL FEATURES, (C) & (D) ILLUSTRATION OF METHODOLOGY 45
FIG 4.1. VARIOUS NOISE IN ULTRASOUND SPINE IMAGE (A) RAW IMAGE, (B) AND (C) TYPES OF SPECKLE NOISES 46
FIG 4.2. PROPOSED ARCHITECTURE OF LDS U-NET 49
FIG 4.3. SCHEMATIC REPRESENTATION OF CONVENTIONAL VS. DEPTHWISE SEPARABLE CONVOLUTION 50
FIG 4.4. PROPOSED LIGHT DENSE BLOCK 51
FIG 4.5. MULTI-SCALE INCEPTION MODULE 53
FIG 4.6. SELECTION GATE 54
FIG 4.7. FOUR IMAGE SETS WITH EACH HAVING A RAW IMAGE PLUS A TRUTH MASK 56

FIG 4.8. INDIVIDUAL SEGMENTATION RESULT OF ULTRASOUND SPINE IMAGE: (A) RAW IMAGE, (B) TRUTH MASK, (C) LD MODEL, (D) LCS MODEL, (E) LDI MODEL, AND (F) PROPOSED LDS U-NET MODEL. 59
FIG 4.9. QUALITATIVE COMPARISON OF 6 DIFFERENT CASE MODELS: (A) RAW IMAGE (B) TRUTH MASK, (C) LD MODEL, (D) LCS MODEL, (E) LDI MODEL, AND (F) PROPOSED LDS U-NET MODEL 60
FIG 4.10. FOUR SPINE IMAGES SETS (A, B, C, D) CONSISTING OF THE INPUT IMAGE, TRUTH MASK, AND SEGMENTATION RESULTS USING U-NET, ATTENTION U-NET, MULTIRESUNET, AND THE PROPOSED LDS U-NET MODEL 62
FIG 4.11. LDS U-NET IS UNABLE TO DISTINGUISH T12 LEVEL AND 1ST LBF 66
FIG 5.1. OVERALL ARCHITECTURE OF PROPOSED SIU-NET 71
FIG 5.2. INCEPTION BLOCK IN PROPOSED ARCHITECTURE 72
FIG 5.3. FEATURE FUSION SCHEME USING THE DENSE CONVOLUTION OPERATION 73
FIG 5.4. PERFORMANCE COMPARISON OF MODELS USING (A) JACCARD INDEX, (B) DICE INDEX AND (C) HISTOGRAM EUCLIDEAN DISTANCE 77
FIG 5.5A. SIU-NET IS ABLE TO DISTINGUISH TBF PAIRS 78
FIG 5.5B. SIU-NET IS ABLE TO DISTINGUISH TBF AND LBF 78
FIG 5.6A. SIU-NET IS ABLE TO SEGMENT LBF 78
FIG 5.6B. SIU-NET IS ABLE TO IDENTIFY SIX INDIVIDUAL LBFS 78
FIG 5.7A. SIU-NET IS ABLE TO IDENTIFY TBFS THAT ARE INVISIBLE IN TM 79
FIG 5.7B. SIU-NET IS ABLE TO DETECT THE EDGES OF TBFS THAT ARE INDISTINGUISHABLE IN OTHER TWO METHODS 79
FIG 5.8A. SIU-NET OUTPERFORMS MULTIRESUNET IN MIDDLE BONY FEATURE SEGMENTATION 79
FIG 5.8.B. SIU-NET IS ABLE TO DISTINGUISH BOTH TBFS AND LBFS CLEARLY 79
FIG 5.9. QUALITATIVE COMPARISON OF ALL MODELS USING ULTRASOUND SPINE IMAGE . 80
FIG 5.10. PROPORTION (%) OF IMAGES WHERE SEGMENTATION OUTPUTS CAN DETECT ALL BONY FEATURES FROM RESPECTIVE TRUTH MASKS 82
FIG 6.1. PROCESS FLOW OF CENTROID PAIRING AND INSCRIBED RECTANGLE SLOPE (CPI-SLO) METHOD 94
FIG 6.2. COMPARISON OF USABILITY OF IMAGES FOR AUTOMATIC SCOLIOSIS ASSESSMENT 96

FIG 6.3. RGB CONVERSION OF BINARY-SEGMENTED ULTRASOUND SPINE IMAGES (A) SEGMENTED BINARY IMAGE (B) RGB CONVERSION WITHOUT SPECIFIC COLOUR CODE (WITHOUT ALGORITHM 1) & (C) RGB CONVERSION WITH SPECIFIC COLOUR CODE (WITH PROPOSED ALGORITHM 1) 97

FIG 6.4. IDENTIFICATION OF VERTEBRAE FEATURES (A) INPUT IMAGE, (B) SEGMENTED BONY FEATURE (C) IDENTIFIED TBFS (D) IDENTIFIED LBFS........ 100

FIG 6.5. MEASUREMENT OF MAIN THORACIC ANGLE USING CENTROID PAIRING METHOD 100

FIG 6.6. DIFFERENT CASES OF MEASUREMENT OF MAIN THORACIC ANGLE USING CENTROID PAIRING METHOD........ 101

FIG 6.7. CONJOINT LBF 105

FIG 6.8. MEASUREMENT OF THORACO-LUMBAR ANGLE USING LARGEST INSCRIBED RECTANGLE METHOD........ 105

FIG 6.9. CORRELATIONS ($\boldsymbol{R2}$) AND REGRESSION EQUATIONS BETWEEN (A) MANUAL-UNET, (B) MANUAL-LDS U-NET & (C) MANUAL-SIU-NET AND THORACO-LUMBAR ANGLES (E) MANUAL-UNET, (F) MANUAL-LDS U-NET & (G) MANUAL-SIU-NET 109

FIG 6.10. BLAND–ALTMAN PLOTS WHICH DEMONSTRATES THE DIFFERENCES BETWEEN THE (A) MANUAL-UNET, (B) MANUAL-LDS U-NET & (C) MANUAL-SIU NET FOR MAIN THORACIC ANGLES AND (E) MANUAL-UNET, (F) MANUAL-LDS U-NET & (G) MANUAL- SIU NET FOR THORACO-LUMBAR ANGLES. THE CENTRAL LINE REPRESENTS THE BIAS AND THE DOTTED LINES REPRESENT THE 95% LIMITS OF AGREEMENT..... 110

FIG 6.11. SPECIAL CASES FOR MAIN THORACIC UCA MEASUREMENT (A) UPPER RIGHT TWO BONY FEATURES ARE CONJOINT, EXTRA BONY FEATURE IS PRESENT IN RIGHT SIDE (B) $\boldsymbol{Pn - Pm = 1}$ & (C) $\boldsymbol{Pm = Pn}$........ 113

FIG 6.12. SPECIAL CASES FOR THORACO-LUMBAR UCA MEASUREMENT: (A) & (B) LOWER MOST LBF IS BROKEN AND IS COUNTED AS 7........ 114

LIST OF TABLES

TABLE 1.1 RELATIONSHIP BETWEEN COBB ANGLE AND SEVERITY OF SCOLIOSIS 2
TABLE 2.1 SUMMARY OF LITERATURE REVIEW 17
TABLE 2.2 FEATURES OF VARIOUS IMAGING MODALITIES IN SCOLIOSIS ASSESSMENT 21
TABLE 2.3 SUMMARY OF VARIOUS SCOLIOSIS ASSESSMENT METHODS 29
TABLE 3.1 APPLICATION OF ULTRASOUND IMAGING IN DIAGNOSIS OF SPINAL DEFORMITIES 35
TABLE 3.2 SUMMARY OF THE DATA SET USED IN THE RESEARCH 42
TABLE 3.3: STEPS TO MITIGATE SCAN NOISE AND THEIR DRAWBACKS 43
TABLE 4.1: QUANTITATIVE EVALUATION OF ABLATION STUDY 58
TABLE 4.2: QUANTITATIVE PERFORMANCE EVALUATION OF VARIOUS ARCHITECTURES .. 60
TABLE 4.3: COMPARISON OF NUMBER OF BONY FEATURES IDENTIFIED (AVG.) 63
TABLE 4.4: EVALUATION OF COMPUTATIONAL REQUIREMENT 63
TABLE 4.5: LDS U-NET DETECTION RATE 67
TABLE 5.1: QUANTITATIVE EVALUATION 76
TABLE 5.2: QUANTITATIVE EVALUATION OF FEW SELECTED INDIVIDUAL IMAGES 81
TABLE 5.3: COMPARISON OF SEGMENTATION PERFORMANCE OF VARIOUS ARCHITECTURES IN SPINE IMAGES USING VARIOUS IMAGING MODALITIES 83
TABLE 5.4: COMPARISON OF SCOLIOSIS CURVATURE ANGLES MEASURED USING VARIOUS TYPES OF ULTRASOUND IMAGES WITH TRADITIONAL COBB ANGLE 84
TABLE 5.5: COMPARISON OF PERFORMANCE OF VARIOUS ULTRASOUND SCOLIOSIS MEASUREMENT INDICES WITH RADIOGRAPHIC COBB ANGLE 85
TABLE 5.6: SUMMARY OF SEGMENTATION ARCHITECTURES APPLIED TO ULTRASOUND IMAGES OF SPINE AND OTHER BODY PARTS 87
TABLE 5.7: COMPARISON OF DETECTION RATES FOR VARIOUS SEGMENTATION ARCHITECTURES IN ULTRASOUND SCOLIOSIS MEASUREMENT 91
TABLE 6.1: COMPARISON OF QUANTITATIVE SEGMENTATION RESULTS FOR VARIOUS SEGMENTATION ARCHITECTURES 95
TABLE 6.2: COMPARISON OF DETECTION RATES FOR VARIOUS SEGMENTATION ARCHITECTURES 95
TABLE 6.3: MEAN RANGE OF UCA ANGLE 108

TABLE 6.4: COMPARISON OF ACCURACY BETWEEN AUTOMATIC UCAS AND MANUAL UCA 109
TABLE 6.5: CORRELATION BETWEEN AUTOMATIC UCAS AND MANUAL UCA 111
TABLE 7.1: SUMMARY OF PERFORMANCE OF PROPOSED SEGMENTATION METHODS 118

LIST OF ACRONYMS

AIS	Adolescent Idiopathic Scoliosis
MRI	Magnetic Resonance Imaging
VPI	Volume Projection Image
SPA	Spinous Process Angle
UCA	Ultrasound Curve Angle
TBF	Thoracic Bony Features
LBF	Lumbar Bony Features
LDS U-Net	Light-Convolution Dense Selection U-Net
SIU Net	Skip-Inception U-Net
CPI-SLO	Centroid Pairing and Inscribed rectangle Slope
CT	Computed Tomography
PA	Posterior-Anterior
COL	Centre Of Laminae
SP	Spinous Process
TP	Transverse Process
CNN	Convolutional Neural Network
ROI	Region of Interest
BMI	Body Mass Index
.png	Portable Graphics Format (file format)
ReLU	Rectified Linear Unit
LD	Light Dense Model
LCS	Light Convolution Selection Model
LDI	Light-Dense Inception Model
JS	Jaccard Similarity
DC	Dice coefficient
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
SCI	Spinal Cord Injury
SMA	Spinal Muscular Atrophy

RSN-U-net	U-Net With Robustness to Speckle and Regular Occlusion Noise
D-TVNet	Dual-Task Ultrasound Transverse Vertebrae Segmentation Network
ASPP	Atrous Spatial Pyramid Pooling
IB	Inception Block
BN	Batch Normalization
DSC	Dense Skip Connections
IU	Inception+ U-Net Model
IRs	Inception + Res Path Model
ISP	Inception + Encoder Skip Path Model
TM	Truth Mask
AOR	Aid of previous Radiographs
GVF	Gradient Vector Flow
US	Ultrasound
SOSORT	Scientific Society on Scoliosis Orthopaedic and Rehabilitation Treatment

CHAPTER I: INTRODUCTION

1.1 Background

1.1.1 Scoliosis

Scoliosis is a deformation of the spinal cord, usually S- or C-shaped, where a curvature, generally greater than 10°, occurs in the plane between the dorsal and ventral parts, i.e., coronal plane. In some cases, the degree of curvature is stable, while in others, it increases over time. It affects primarily during the adolescent period; as the skeletal structure of a child develops, a significant rate of curve progression may occur in the spine (Reamy & Slakey 2001). This medical condition is called idiopathic scoliosis and originates in children in the age range of 10–16 years (Cassella & Hall 1991). Adolescent idiopathic scoliosis (AIS) is found in 0.47–5.2% of the general population and is prevalent across the world, such as in Hong Kong (3.08%) (Fong et al. 2015), China (5%) (Cheng et al. 2015), USA (2–4%) (Horne, Flannery & Usman 2014) and Germany (5.2%) (Konieczny, Senyurt & Krauspe 2013). Further, it is found that girls are twice as susceptible to AIS compared to boys (Konieczny, Senyurt & Krauspe 2013). If left untreated, AIS can have a permanent damaging effect on the skeletal structure of adolescents and can aggravate chronic back pain and thoracic insufficiency syndrome (Campbell Jr et al. 2003). Severe scoliosis may also affect the heart, lungs, and nervous system (Li, Yang, et al. 2017; Liu et al. 2020; Ran, Zhi-hong & Jiang-na 2011). Therefore, early detection of scoliosis through periodic monitoring of spine curve progression until the complete maturity of skeletal structure is necessary for AIS patients (Schulte et al. 2008; Yan et al. 2020). The complete course of scoliosis treatment may involve 25 sessions (Reamy & Slakey 2001).

The exact cause of scoliosis is unknown, but doctors theorize that it may happen during birth or may be the side effect of certain nerve or muscle problems such as cerebral palsy. The doctors also attribute it to genetic factors and hormonal problems. There are no early signs for this disease. As the disease matures, patients develop some symptoms such as uneven shoulders, inflated curvature of the spine, disproportional alignment of hips, back pain and discomfort. If not treated properly, progressive scoliosis may cause constant back pain, breathing problems, or even physical disability for life (Li, Yang, et al. 2017; Liu et al. 2020; Ran, Zhi-hong & Jiang-na 2011). So, the early detection of scoliosis is critical to the preventing aggravation during the later stage of life.

1.1.2 Conventional detection of scoliosis

The conventional method of diagnosing scoliosis is using Cobb's measurement technique to measure the curvature angle of the spine on a standing radiograph (Cobb 1948). The traditional Cobb angle measurement is a manual method and involves

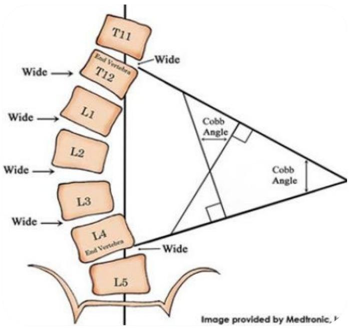


Fig 1.1. Cobb Angle
(Safari et al. 2019)

Table 1.1 Relationship between Cobb angle and severity of scoliosis

Cobb Angle	Severity of scoliosis
10°-30°	Mild
30°-45°	Moderate
>45°	High

indicating the most tilted upper and lower end vertebrae and estimating the angle between intersecting lines drawn perpendicular to the top of the upper vertebrae and the bottom of the lower vertebrae (shown in Fig 1.1). Depending on the value of Cobb angle, the severity of scoliosis can be predicted, as shown in Table 1.1.

Several X-ray examinations are needed over the course of monitoring the scoliosis curve progression (Horne, Flannery & Usman 2014; Kim et al. 2010). However, frequent X-ray examinations pose a radiation threat. Moreover, it results in an increased risk of breast cancer among women, prostate cancer in men, and risk of leukemia in AIS patients (Doody et al. 2000; Hoffman et al. 1989; Schmitz-Feuerhake & Pflugbeil 2011; Simony et al. 2016). Levy et al. demonstrated that AIS patients who are exposed to X-rays multiple times have a 2.4% higher risk of developing cancer than normal young people (Levy et al. 1996). Therefore, new imaging modalities that emit limited ionizing radiation has become the need of the hour (Presciutti, Karukanda & Lee 2014).

Other imaging modalities to detect scoliosis are summarized as follows:

(a) An advanced low-dose bi-planar radiograph system called EOS was introduced for scoliosis detection (McKenna et al. 2012). It provides detailed and high-quality images useful for medical diagnosis. Though it reduced the required radiation dose by 8–10 times that of the conventional X-ray, the high operation and installation cost of EOS prevented it from becoming a useful option for mass scoliosis detection (Rehm et al. 2017; Vergari et al. 2019).

(b) Magnetic resonance imaging (MRI), a radiation-free technology, can also be used to identify the features of the bone in the spine for the detection of scoliosis. Though MRI offers advantages of high resolution with 3D information, the limitation of a traditional MRI setup is that a patient can be scanned only in the supine position. For accurate scoliosis assessment, the patient should be scanned in a standing position, but standing MRI setups needs a dedicated space for installation and significant operating time. Further, it is currently unavailable and expensive for large-scale access (Ungi et al. 2014).

(c) Conventional ultrasound imaging is another popular technique owing to its cost-effectiveness, almost real-time measurement, and radiation-free procedure and is one of the most suitable techniques for this purpose. However, a 2D ultrasound cannot facilitate the viewing of complex spinal cord structures due to its inherent lack of good resolution. In a recent development, by adding B-mode ultrasound using some position sensors, this problem of 2D viewing was eliminated, although the process is quite tedious because it requires dozens of images (Chen, Lou & Le 2011).

A recent invention in this field is freehand 3D ultrasound in which the conventional 2D ultrasound is combined with a position sensor (Huang et al. 2005). Subsequently, Zheng et al. developed a 3D ultrasound system called Scolioscan (Zheng, Lee, et al. 2016) and demonstrated its reliability for scoliosis diagnosis (Zhou & Zheng 2015) and efficacy vis-à-vis the conventional radiographic Cobb method (Zheng, Lee, et al. 2016). Scolioscan comprises an ultrasound scanner with a built-in linear probe, a frame structure, an electromagnetic spatial sensing device, and propriety software. The procedure involves freehand scanning using the probe from the bottom to the top of the patient's back, covering the whole spine area, while the electromagnetic spatial sensing device continuously detects the probe's position and orientation. The B-mode images, collected along with their corresponding position and orientation information, are used to reconstruct into a 3D image, thereby forming the coronal view of the spine, using the volume projection image (VPI) method (Cheung, Zhou, Law, Mak, et al. 2015).

1.2 Motivation

The research aims to introduce a novel technique to monitor scoliosis in a patient and can be applied in large scale. For doing so, it is worthwhile to update on the traditional and current methods of scoliosis angle detection using X-ray and 3D ultrasound VPI imaging.

1.2.1 Measurement of scoliosis progression using X-ray

The measurement of Cobb angle using a radiograph involves imaging the anterior spinal column structures to describe the anterior spinal deformity. In this process, the measurement accuracy is largely dependent on the awareness and practice of the observer and the quality of images. It is difficult for experts to make accurate measurements because of the large anatomical variations among patients from different age groups and could result in a large number of inter- or intra-observer errors (Horng et al. 2019).

1.2.2 Measurement of scoliosis progression using Scolioscan

While the Cobb angle indicates the anterior spinal deformity, another index more suitable to ultrasound imaging, called spinous process angle (SPA), measures the posterior spinal deformity. In a nutshell, SPA is the angle formed between the lines drawn through the most tilted part of the spinous column profile of coronal ultrasound images. Several measurement techniques to calculate the SPA were developed and have been proven to be comparable to the earlier gold standard of Cobb angle obtained through the X-ray method:

1.2.2.1 Spinous process angle

a) Manual measurement of SPA: In the VPI-SP method, inflection points are derived to locate the spinous column profile in the VPI images, and then the 6^{th}-order polynomial curve fitting technique is used to calculate the spine curvature, as shown in Fig 1.2a (Cheung, Zhou, Law, Mak, et al. 2015). Since the derivation of inflection points is done manually, this process is subject to the expertise of the observer and hence time-consuming.

b) Automatic measurement of SPA: In (Zhou et al. 2017), phase congruency and the novel two-fold threshold strategy were employed to measure spine curvature automatically. Though this method overcame the limitations of manual measurement, the computational time was long as the program spent most of the time computing the phase congruencies of the images.

Also, in general, the SPA measurement approach has two limitations: (a) compared to Cobb angle measurement, the severity of curvature is underestimated, showing that it was less angulated compared to the corresponding Cobb angle (Brink et al. 2018; Zheng, Lee, et al. 2016); (b) the SPA measurement also uses more landmarks than the Cobb angle, which introduces more sources of variation for each measurement (Morrison et al. 2015).

1.2.2.2 Ultrasound curve angle

An alternative indicator to assess scoliosis, a new measure called ultrasound curve angle (UCA) (Fig 1.2b), has been demonstrated to be comparable to radiographic Cobb angle (Lee et al. 2021). In SPA, the spinous column profile is the key anatomical reference for angle measurement, whereas in UCA, more lateral features of the spine are employed for this measurement. Therefore, for UCA measurement, an alternative VPI method was explored.

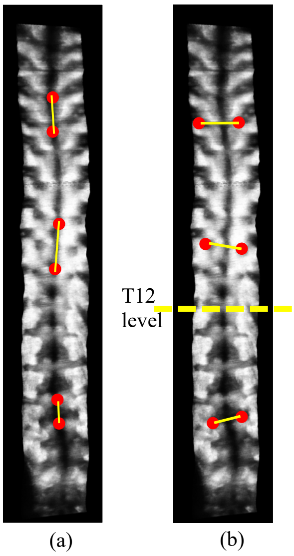


Fig 1.2. Angle measurement techniques for ultrasound imaging modality- (a) SPA (b) UCA

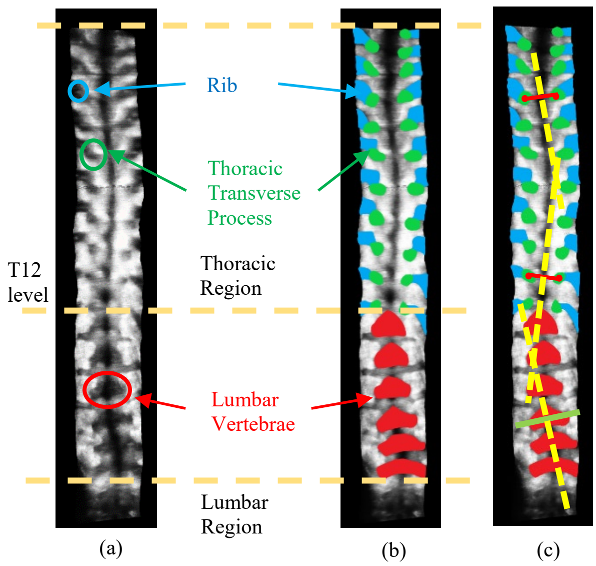


Fig 1.3. Anatomy of an ultrasound spine image (a) Input image, (b) types of bony feature identified (c) angle measurement scheme

Manual Measurement of UCA: A 3D coronal ultrasound image was split into nine 2D images with different depth profiles. Two raters were then asked to select the 2D image with the most visible lateral features among those nine images. One such 2D ultrasound spine image is shown in Fig 1.3a. Fig 1.3b illustrates the various segments of the spine – the thoracic bony features, rib, T12 level, and lumbar bony features. Fig 1.3c outlines the current methodology used by a medical expert to calculate the UCA. The most tilted thoracic bony features (TBFs) and lumbar bony features (LBFs) are identified by evaluators. Further steps involve drawing lines through the centre of the identified TBF and LBF pairs and allocating lines to measure the UCA (Lee et al. 2021).

The selection of the best image, clear detection of TBFs and LBFs, identification of most tilted thoracic and lumbar features, and finally calculation of angles play a vital role in UCA measurement. All the identification and calculation processes are manual, time-consuming and fully dependent upon the expertise of the doctors and raters.

To overcome the problem associated with manual UCA measurement, a novel automatic process for calculation of UCA for scoliosis assessment is developed through this research.

1.3 Research questions

There are five major questions guiding this research:

- RQ1. Regular observation of scoliosis progression is necessary for proper treatment. But according to research, frequent X-ray examinations during the growth period increase the risk of cancer in AIS patients. Moreover, the diagnosis process should be affordable. **Is there any imaging modality that is economic and has no side effects?**
- RQ2. The manual and semi-automatic method of scoliosis detection depends on the subjective experience of radiologists. Moreover, it is time-consuming. **How can automatic detection of scoliosis be performed on a large-scale?**
- RQ3: For manual scoliosis assessment using ultrasound imaging modality, the bony feature is identified by medical experts. **How to automatically identify the bony features for scoliosis detection?**
- RQ4: For manual scoliosis assessment using ultrasound imaging modality, the spine curvature angle is calculated through human intervention. **How can the scoliosis curvature angle be measured automatically?**
- RQ5. So far, manual measurement of spine curvature is considered the best scoliosis assessment technique. **How accurately can an automatic assessment system for scoliosis assessment correlate with the manual measurement system?**

1.4 Aim of the research

Based on the research questions, the project's aims were decided. This research aims to develop a radiation-free assessment of scoliosis using 3D ultrasound and convolutional neural network. The intent is to develop an innovative technique to automatically assess the spinal deformity, eliminate the need for manual intervention of physicians and provide output with a high level of precision and scalability. This work will provide accurate, fast, and comprehensive assessment to scoliosis curvature angle.

1.5 Research objectives

The main objectives of this research are as follows:

- To develop a radiation-free and economic scoliosis detection system
- To segment the lateral bony features from noisy ultrasound spine images
- To automatically calculate the ultrasound curvature angle (UCA) without any human intervention
- To make the automatic scoliosis detection system comparable with the manual one

1.6 Research contributions

The automatic UCA measurement process (Fig 1.4) is designed and developed for early detection and mass screening of scoliosis. The contributions are as follows:

1. **Segmentation technique 1:** Introduction of Light-Convolution Dense Selection U-Net (LDS U-Net) for ultrasound spine bony feature segmentation: Identification of bony features from ultrasound spine image is a very important step for the automatic angle calculation process for scoliosis assessment. To identify the bony features without any human intervention, segmentation of bony features is done in this research. To overcome the challenges associated with ultrasound spine imaging and clear segmentation of bony features, Light-Convolution Dense Selection U-Net (LDS U-Net) is proposed. The proposed architecture has three main aspects: (a) the basic U-Net (Ronneberger, Fischer & Brox 2015) structure is adopted as the network architecture, but the conventional convolutional layers are replaced with dense depth-wise separable convolution layers to increase the computational efficiency; (b) selection gates (Oktay et al. 2018) are deployed for a smarter identification of the target bony features; and (c) the encoders–decoders are connected using multi-scale skip-pathways (Ibtehaz & Rahman 2020) to enhance feature fusion. The proposed LDS U-Net can identify the thoracic and lumbar bony features more consistently when compared to other contemporary models. The result confirms that LDS U-Net can successfully segment lateral bony features and can be included as a step prior to the automatic angle calculation.
2. **Segmentation technique 2:** Introduction of multi-scale feature fusion Skip-Inception U-Net (SIU-Net) for ultrasound spine image segmentation: Two major concerns regarding bony features segmentation in ultrasound spine images are addressed in this research: 1) choosing the appropriate kernel size for the convolution operation which can handle the large variation in locations, shapes,

and sizes of the features; and 2) designing a suitable architecture that can extract semantically rich features and fuse multi-scale features for a better segmentation output. To solve these problems, a novel deep learning architecture, Skip-Inception U-Net or SIU-Net, is proposed to suitably segment the thoracic and lumbar bony features in the ultrasound spine image dataset. The standard U-Net is adopted as the main network architecture, and the simple convolutional layers are replaced with inception blocks (Szegedy, Vanhoucke, et al. 2016). The encoders–decoders are bridged using newly designed decoder side skip-pathways (Zhou et al. 2018). The performance of the proposed network is evaluated in terms of both the pictorial quality and segmentation accuracy using ultrasound VPI images. The results show that this network gives an improved performance in both cases, which makes the network usable as a step prior to automatic angle measurement.

3. **Automatic angle calculation:** Development of centroid pairing and inscribed rectangle slope (CPI-SLO) method to find the most tilted bony features to calculate the UCA: Manual ultrasound curve angle (UCA) measurement has been established to have a good agreement with the traditional radiographic Cobb angle method (Lee et al. 2021). Therefore, the sequence of the manual UCA measurement process is encompassed in this work for automatic angle measurement. Three algorithms are proposed in this work to select key features from segmented spine images and then calculate the main thoracic and thoraco-lumbar angles separately. The proposed algorithms of automatic UCA measurement are constructed such that they can smartly adapt themselves depending on the quality of the bony feature information in the segmented outputs. The algorithms use centroid pairing and the largest inscribed rectangle to find the slope of thoracic and lumbar regions, respectively. This method is called the CPI-SLO method.

1.7 Thesis framework

The automatic UCA measurement technique comprises three steps:

1. Manual selection of image with best lateral features from 2D volume projection ultrasound spine images
2. Segmentation of lateral bony features (TBFs and LBFs)
3. Calculation of automatic UCA

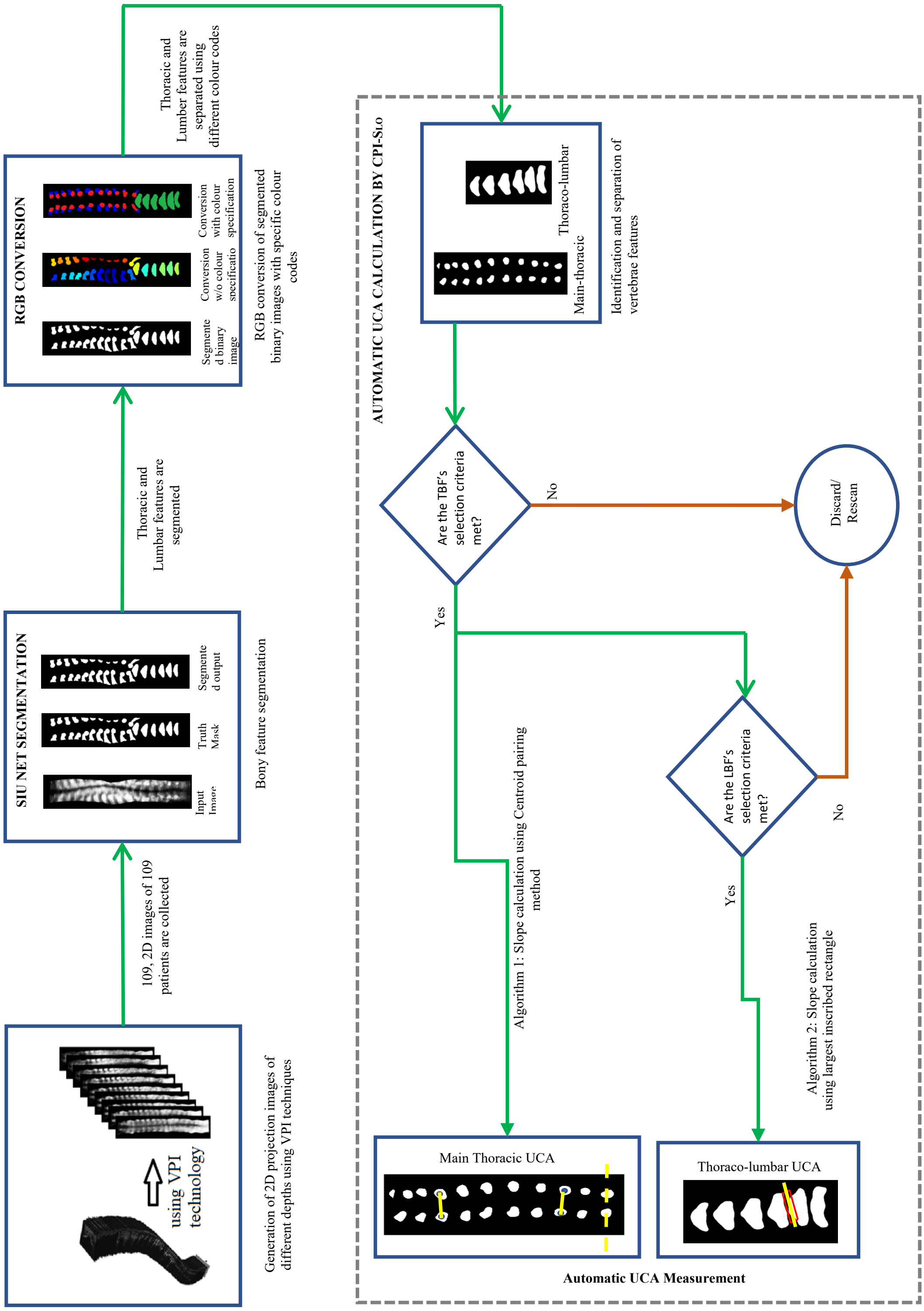


Fig 1.4. A complete flowchart for automatic UCA measurement

The thesis is structured into seven chapters:

Chapter 1 provides a brief introduction, objectives, and contributions of this research followed by an outline of the thesis.

Chapter 2 delivers a brief background of 2D and 3D ultrasound imaging techniques. This chapter also includes a detailed discussion on the various imaging modalities used for scoliosis assessment, manual, semi-automatic, and automatic scoliosis angle detection techniques using ultrasound imaging modality, baseline, and state-of-the-art image processing, and machine learning classification and segmentation techniques applied on ultrasound medical images.

Chapter 3 details the generation and collection of the ultrasound spine image dataset.

Chapter 4 focuses on the development of the first proposed segmentation technique, light-convolution dense selection U-Net (LDS U-Net) for ultrasound spine bony feature segmentation.

Chapter 5 presents a detailed description of the second proposed segmentation technique, multi-scale feature fusion skip-inception U-Net (SIU-Net) for ultrasound spine image segmentation.

Chapter 6 concentrates on the RGB conversion of segmented binary images for key feature selection for UCA calculation and introduces another proposed algorithm, centroid pairing and inscribed rectangle slope (CPI-SLO) method, to find the most tilted bony features to calculate UCA.

Chapter 7 concludes the thesis and presents the scope of future works.

CHAPTER II: LITERATURE REVIEW

2.1 Introduction

Humankind has been battling scoliosis for thousands of years. The very first documentation of the detection and cure of scoliosis was in an ancient Indian manuscript, Shrimad Bhagavatam, written around 3102 BCE. According to the text, the Hindu god Krishna corrected the scoliosis of a woman whose spine was deformed in three different places (Kumar 1996).

Scoliosis comes from the Greek word 'skoliosis', which means crooked. It is an ailment in which the spinal cord deforms progressively with time. Depending on the patient's age, cause, type of curve, and severity of the disease, scoliosis can be classified into two categories (Fig 2.1): idiopathic and non-idiopathic (Konieczny, Senyurt & Krauspe 2013). Adolescent idiopathic scoliosis (AIS) is the most common type of scoliosis, with an annual prevalence of 0.47–5.2% and a female/male ratio of 1.5:1 to 3:1. Its severity increases with age (Jada et al. 2017).

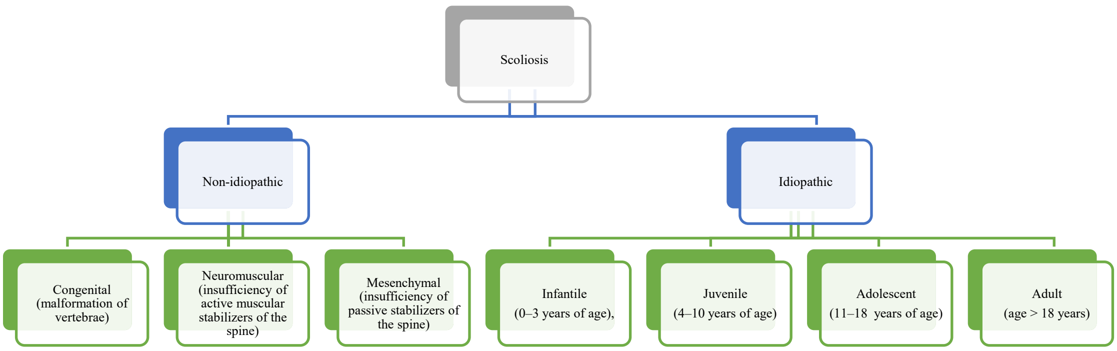


Fig 2.1. Categorization of Scoliosis

The treatment options for scoliosis vary with the age of the patient, location, severity, and risk of progression of the curvature. According to the International Scientific Society on Scoliosis Orthopaedic and Rehabilitation Treatment (SOSORT), there are three basic treatment options for various stages of scoliosis based on the degree of curvature of the spine (Dunn et al. 2018). Figure 2.2 summarizes the relationship between the treatment options and curvature of the spine i.e. severity of scoliosis.

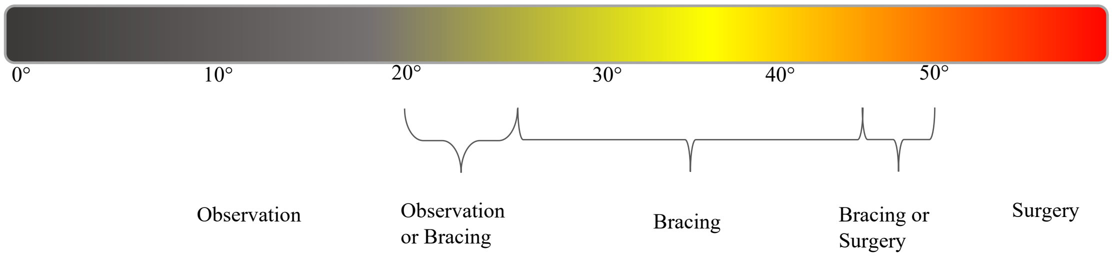


Fig 2.2. Treatment of Scoliosis

Patients with AIS undergo spinal surgery for correcting their deformity (Choudhry, Ahmad & Verma 2016). Early detection is the key to reducing the morbidity of AIS.

The popular method of detection and diagnosis of scoliosis involves (a) multiple X-ray scanning sessions and (b) measurement of Cobb angle, which is a gold standard to assess AIS (Cobb 1948). The main drawback of this method is the repeated exposure to ionizing radiation. Levy et al. established that an AIS patient undergoing frequent X-ray scans has a 2–4% higher chance of getting cancer than normal young people (Levy et al. 1996). Further, other non-ionizing radiation modalities such as MRI are very expensive and not designed for scanning in standing position.

2.2 Imaging modalities for scoliosis detection

Scoliosis may have several causes; however, it is most commonly idiopathic. Radiography, magnetic resonance imaging (MRI), computed tomography (CT), and ultrasound are the imaging modalities used for assessment of scoliosis.

2.2.1 Radiography

Plain Radiography: Plain radiography remains the backbone of the assessment of scoliosis during initial, as well as preoperative and postoperative, stages to plan surgery and follow-up. Generally, standing posterior-anterior (PA), lateral, and side-bending X-ray radiographs are the standard for scoliosis evaluation (Fig 2.3). Standing radiographs are ideal for diagnosis and periodic monitoring of patients because the spinal balance and curve magnitude change when a patient is in the supine position (Jada et al. 2017).

The conventional method of diagnosis of scoliosis is using Cobb's measurement technique to measure the curvature angle of the spine on a standing radiograph (Cobb 1948). If two lines are drawn such that one is perpendicular to the upper endplate of the most titled upper vertebra and another one is perpendicular to the lower endplate of the most tiled lower vertebra, then the angle between them is known as the Cobb angle (Cobb 1948).

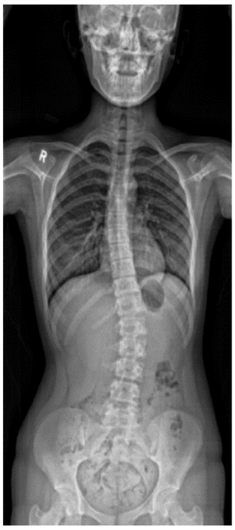


Fig 2.3. X-ray spine image (Imran et al. 2020)

The Cobb angle method is the accepted standard for measuring scoliosis using radiographs. The Cobb angle varies in the same spine radiographed in the supine and upright positions (Cassar-Pullicino & Eisenstein 2002). The clinical impression of scoliosis requires radiographic confirmation; a 10° deformity is the minimum measurement that establishes true scoliosis. In practical terms, this means that most children with scoliosis require regular review combined with radiographic assessment and measurement of the angle of curvature (Cobb angle) (Wright 2000).

There are obvious concerns that repeated radiographs result in an excessive radiation burden, especially concerning the developing breast tissue in girls. Nash Jr et al. estimated that 22 radiographic examinations are performed in the course of scoliosis management (Nash Jr et al. 1979). According to research, excessive radiation exposure can cause breast cancer in girls and may result in leukemia and prostate cancer in general (Levy et al. 1996).

EOS: Georges Charpak won the Nobel Prize in Physics, 1992, for his study involving multi-wire proportional chamber systems allowing accurate elementary particle detection. From a diagnostic perspective, Charpak's technique could be combined with vertical bi-planar slot-scanning technology (EOS) to produce high-quality diagnostic radiographic images using 50–80% less radiation than required for conventional radiography (McKenna et al. 2012), (Morvan et al. 2011), (Melhem et al. 2016). EOS imaging is capable of scanning the spine in 8–15 seconds (Deschênes et al. 2010), and using the anteroposterior and lateral image captured using the EOS technology, its software system can reconstruct a highly accurate 3D model of the spine (Le Bras et al. 2003). One of the primary advantages of the EOS imaging system over CT is that it can capture images with the patient in an upright standing or seated position (Dubousset et al. 2005).

Although when compared to traditional radiography, the radiation dosage of the EOS is much lower, it is still not negligible. Therefore, an X-ray shielding room should be used as the installation space for the EOS system. Moreover, if the EOS system is used for scoliosis longitudinal follow-up, the amount of diagnostic radiation absorbed could result in adverse effects. Further, the cost of the machine and its maintenance are exorbitant, which makes it practically impossible to adopt it in regular settings (Lai et al. 2021).

2.2.2 Magnetic Resonance Imaging (MRI)

Magnetic resonance imaging (MRI) is one of the most preferred techniques for viewing the spine curvature, and it has a revolutionary impact on imaging the neural axis (Fig 2.4). It is a clinical tool that does not use ionizing radiation and provides excellent delineation of the contents of the spinal canal, including the spinal cord. Its ability to produce detailed images with excellent tissue contrast in any plane, without the use of ionizing radiation, makes it a promising method of assessing the scoliotic spine. One of the chief advantages of this technique is that it can provide the spine's anatomical, muscular, and neurological information (Jada et al. 2017). The nature of scoliosis, with different areas of the spine moving in and out of the plane of imaging, means that interpretation might be difficult. In addition, the examination is time-consuming, and once the surgery has been performed, the presence of metalwork can make further studies with MRI suboptimal, primarily because of the magnetic susceptibility artifacts that are generated from the internal fixation. Further, young children require sedation and occasionally general anaesthesia to comply with MRI-compatible equipment (Wright 2000).

Fig 2.4. MRI spine image (Rajasekaran et al. 2010)

According to Musson et al., most MRIs performed for scoliosis are requested by orthopaedic surgeons before corrective surgery for AIS. MRI is indicated in several clinical circumstances and is being increasingly used by surgeons and paediatricians. The most common use is in the assessment of AIS, primarily to look for intraspinal abnormalities. MRI is also used for assessment and follow-up of congenital scoliosis and the developmental causes of scoliosis (Musson et al. 2010).

Compared to traditional imaging options, MRI incurs considerable costs, as well as scan time. The prolonged duration of the scan mandates the use of sedation in special cases such as paediatric patients. Moreover, the sequences that MRI uses are novel and experimental and are, thus, not compatible with standard scanners (Pasha et al. 2021).

Generally, the aim of MRI in AIS is to detect clinically occult abnormalities that may affect the surgical technique, consent, and prognosis. This imaging technique is mostly used for all preoperative cases and those with vertebrae segmental deformity and neuraxial abnormalities (Lai et al. 2021). However, one of the primary differences between traditional MRI and standing MRI lies in the patient's posture; while traditional MRI captures images with the patient in the supine position, standing MRI requires the patient to be in standing position. Moreover, it requires a dedicated space for its installation, as well as a significant amount of scanning time. Further, since it is a novel technique with considerably high demands, it is not widely accessible yet (Ungi et al. 2014).

2.2.3 Computed Tomography (CT)

Children with underlying congenital skeletal malformations and segmentation abnormalities require a slightly divergent imaging approach. MRI provides excellent detail about the nature of the spinal canal and cord, but the detail of the bony anatomy can be evaluated further with localized CT and 3D reconstructed images of the area of interest (Fig 2.5). However, in practice, often plain radiography suffices. CT plays a role in postoperative patients who cannot be imaged appropriately in the MRI environment. Localized CT might help in identifying postoperative complications. Notably, considerable artifact is common because of the metalwork present, but the images might be interpretable (Wright 2000).

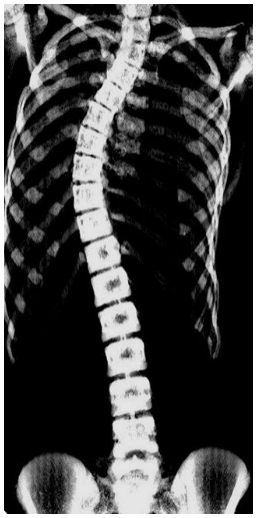


Fig 2.5. CT spine image (Keenan et al. 2014)

The radiation dose remains approximately 12 times higher than low-dose radiography. The scans are not performed in natural standing position, thus limiting the assessment of the 3D spinal alignment (Pasha et al. 2021).

2.2.4 Ultrasound

Ultrasonography as a technique, for estimating the Cobb angles in patients with scoliosis, has been studied by several researchers in recent times. Notably, the ultrasound angles were 15–37% smaller in magnitude than standard Cobb angles. Nevertheless, excellent correlations between the ultrasound angles and Cobb angles were identified; the coefficient of determination was ≥ 0.970. Further, the ultrasound angles were reliable, with an intra-class correlation coefficient of ≥ 0.84 (Brink et al. 2018).

2.2.4.1 3D Ultrasound for Scoliosis detection

The use of ultrasound imaging modality (Fig 2.6) for the assessment of spinal health has been a popular subject of research in recent times (Hwang et al. 2021). Tawfik et al. (Tawfik et al. 2020) established that, in the identification of main spinal abnormalities in infants, the diagnostic value of spinal ultrasonography is equivalent to that of MRI. In another research, Zhang et al. (Zhang et al. 2021) showed that, for long-term spinal deformity treatment, such as cases of severe scoliosis, ultrasound imaging is safe and gives patient comfort along with producing a clear intrathecal structure for guided treatment. Ultrasound is also applied in pre-procedural imaging of central neuraxial blockade of the spine to identify, in real-time, the ideal trajectory (best angle, direction of approach, and depth) and to optimize the subsequent invasive treatment with fewer needle passes and skin punctures (Kalagara et al. 2021).

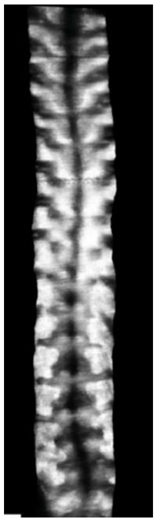


Fig 2.6. Ultrasound spine image (Lee et al. 2021)

Ultrasound imaging is a radiation-free imaging modality that is used extensively in the medical field due to its low cost. However, visibility issues might be encountered when trying to detect bony structures from B-mode ultrasound images. Therefore, a 2D B-mode ultrasound image is not feasible to detect spine deformation. An invention in this field is freehand 3D ultrasound, in which the conventional 2D ultrasound is combined with position sensors to help the system overcome the drawbacks of normal 2D ultrasound imaging technique. However, 3D ultrasound imaging systems are generally experimental prototypes, and they are not ready enough to use for large clinical applications. Using this system, the measurement of spine curvature angle and scanning is done manually. Therefore, there are chances of human errors and variation in the measurement. So, more research is needed to make a 3D ultrasound system a perfect and reliable clinical diagnosis tool.

Among all available techniques, a 3D ultrasound volume projection image is the most suitable to visualize spine anatomy. Zheng et al. illustrated that Scolioscan is a reliable and radiation-free system to measure spine deformity using 3D ultrasound imaging for AIS patients (Zheng, Lee, et al. 2016). The result is very satisfactory when compared with the conventional radiographic Cobb method. Moderate to strong correlations were found between the Scolioscan angles and Cobb angles (r2 > 0.72). The intra- and inter-observer reliability were also very good, with an intra-class correlation of >0.87. Therefore, it can be concluded that Scolioscan is a promising 3D ultrasound imaging system for scoliosis screening and monitoring curve progression (Zheng, Lee, et al. 2016).

Scolioscan, developed in Hong Kong, is used to generate 3D volume projection image (VPI) using conventional 3D ultrasound imaging techniques. The instrument has a rigid frame with two movable supporting boards and four supporters to support patients to maintain a stable standing posture during imaging. It also has two LCD screens and one touch screen that is used by the operator for uploading patient information, parameter setting during the scan, generating VPI images, conducting measurements, and producing reports.

Ultrasound is one of the most easily accessible and applicable diagnostic techniques. It has thus become an important part of clinical practice. Its primary benefits and distinguishing features include a low cost and zero radiation emission. However, the chances of human error are high, while those of detection of low magnitude curves are low. Nevertheless, it must be preferred over repeated radiographic assessments to safely monitor curve progression over time (Girdler et al. 2020).

In Table 2.1, a summary of all imaging modalities used for scoliosis assessment is given.

Table 2.1 Summary of literature review

Year & Author	Study	Radiography	MRI	CT	Ultrasound
(Wright 2000)	Imaging in scoliosis	It confirms the diagnosis, identifies the underlying cause and monitors the degree of curvature.	Any child for whom surgery is anticipated.	CT is used in a postoperative patient who cannot be imaged appropriately in the MRI environment.	N/A
(Cassar-Pullicino & Eisenstein 2002)	Imaging in scoliosis	Plain radiography is the standard imaging modality for scoliosis diagnosis.	1) MRI should be the first choice when further imaging is needed. 2) Cost of 'screening MRI' is a matter of concern.	N/A	N/A
(Inoue et al. 2005)	Pre-operative MRI study of AIS patients	N/A	250 patients with AIS were examined	N/A	N/A

			using MRI before spinal surgery		
(Musson et al. 2010)	Imaging for childhood scoliosis	Children with scoliosis routinely undergo a plain radiographic assessment of the spine, which allows quantification of scoliosis.	MRI is mostly performed before corrective surgery for AIS.	Though radiation is high, CT is most often used in complicated cases for preoperative planning.	N/A
(Shi et al. 2015)	Correlation between supine MRI and standing Radiograph	Cobb angle measurement using X-ray is the standard method of scoliosis assessment.	Supine MRI is a reliable alternative to radiographs when Cobb angle >40°	N/A	N/A
(Cheung, Zhou, Law, Mak, et al. 2015)	Ultrasound volume projection imaging for assessment of scoliosis	90% of AIS patients with progressing curves receive unnecessary radiographic intervention.	1) Expensive and time-consuming. 2) Scanning is done in supine position and spine angle calculation is less accurate.	CT is usually utilized for assessing the vertebra rotation in the supine position, which is still disputable because of using supine position in scanning	The 3D ultrasound imaging system can successfully assess scoliosis
(Zheng, Lee, et al. 2016)	Validation study for a radiation-free scoliosis assessment system	1) Exposure to X-ray during childhood significantly increases the chances of leukemia and prostate cancer. 2) EOS is costly and not portable.	N/A	N/A	3D ultrasound imaging is an affordable, radiation-free diagnostic method for assessing scoliosis.
(Zhou et al. 2017)	Automatic measurement of spine curvature using 3D ultrasound	Frequent X-ray examinations can have harmful effects on the human body, especially children.	MRI examinations are performed in the supine position, as well as are time-consuming and expensive.	N/A	US imaging is a real-time, cost-effective and radiation-free technique and comparable to X-ray for scoliosis assessment.
(Jada et al. 2017)	Evaluation and management of AIS	Standing radiographs are used for observation and diagnosis of AIS because spinal balance and curve magnitude change when a patient is in supine position.	MRI is done for patients who need further investigation and surgical treatment.	CT imaging is used for patients with critical conditions.	N/A
(Brink et al. 2018)	Investigate the reliability and	Involves ionizing radiation, and the radiation doses are	MRI is non-ionizing but does not	N/A	Excellent correlations

	the validity of different ultrasound measurement techniques for coronal curve severity	cumulative, resulting in a 9–10 times higher radiation exposure in these patients.	visualize cortical bone very well, is time-consuming, expensive and cannot be performed in upright position.		between the ultrasound and radiographic Cobb measurement
(Vo, Le & Lou 2019)	Semi-automatic 3D ultrasound reconstruction for studying the severity of AIS	X-ray Cobb angle method underestimates spinal deformity, as well as subjects patients to ionizing radiation, thereby making them susceptible to cancer	Traditional MRI is prone to significant errors since it requires patients to be in the supine position which may change the spinal curvature.	Exposure to ionizing radiation makes patients vulnerable to cancer.	3D US can be used safely to detect scoliosis because it does not produce radiation, is affordable, portable and produces real-time results
(Girdler et al. 2020)	Emerging imaging techniques for diagnosing AIS	Although radiograph systems such as EOS imaging are expensive, they are safe to use because they emit low radiation doses. The image quality is comparable with that of conventional radiography.	N/A	CT has the disadvantage of greater radiation exposure, cost, and scanning in a supine position.	US angles show excellent correlation with standard radiographic Cobb angles, as well as help in assessing curve growth
(Yeung et al. 2020)	Comparing scoliotic curvature between prone (CT) and upright positions (EOS) in preoperative AIS patients.	EOS is compatible with CT in clinical settings and emits much less radiation. Thus, it can be safely used with young patients during puberty.	N/A	Although CT is an excellent tool for diagnosing adolescent idiopathic scoliosis (AIS) preoperatively, its chief disadvantage lies in its high radiation production and the need for scanning to be performed in the upright position.	N/A
(Pasha et al. 2021)	Quantitative imaging of the spine in AIS patients	Multiple radiographic follow-ups increase the radiation dose.	MRI's scan time and costs are very high compared to other imaging modalities.	1) Radiation dose is approximately 12 times that needed for a low-dose radiograph.	1) Radiation-free imaging technique in a weight-bearing position.

				2) Supine posture scanning limits 3D spinal assessment.	2) Future applications of AI to reduce the image processing time for 3D measurement can make ultrasound an affordable low-cost portable alternative for 3D spinal imaging.
(Lai et al. 2021)	3D ultrasound spine angle measurement	Though an X-ray radiograph is a gold standard for scoliosis evaluation, it poses a radiation risk.	Traditional MRI scanning is performed in supine position, and standing MRI scanning is not widely available.	N/A	US imaging is cheap, radiation-free, portable, and more accessible compared to other modalities for spine monitoring.

2.2.4.2 Problems with Ultrasound imaging

There are a few problems associated with an ultrasound imaging modality:

i. **Speckle Noise:** Ultrasound images are affected by speckle noise (Fig. 2.7), which decreases the quality of the image. The noise makes the image very difficult to be

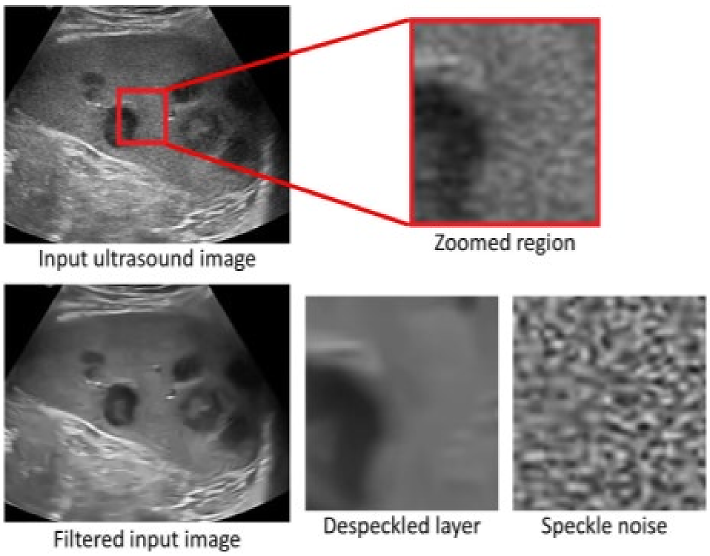


Fig 2.7. Speckle noise

analysed by doctors. As speckle noise behaves as information, it makes the clinical information hard to differentiate (Loupas, McDicken & Allan 1989). So, removal of speckle noise is necessary for better diagnosis of scoliosis.

ii. **The low contrast of ultrasound image:** Ultrasound images are low contrast images because the speed of sound varies depending on the tissues; it becomes difficult to differentiate between fat- and water-based tissues on these images. Therefore, contrast enhancement is very important to extract clinical information from ultrasound images.

Table 2.2 compares all the imaging modalities used for scoliosis assessment.

Table 2.2 Features of various imaging modalities in scoliosis assessment

Imaging Modality	Exposure to Radiation	Scanning posture	Cost	Image Quality	Scan time	Application
Radiograph	Yes	Standing	Moderate	Good	Fast	Periodic assessment of scoliosis
MRI	No	Supine	High	Good	Slow	Surgery cases of scoliosis
CT	Yes	Supine	High	Good	Slow	Critical and surgery cases of scoliosis
3D Ultrasound	No	Standing	Low	Moderate	Fast	Periodic assessment of scoliosis

Key Takeaway:

- MRI and CT imaging modalities are used for critical, pre- or post-operative spine deformation cases and are costly for periodic monitoring of AIS. Therefore, these two imaging modalities are not suitable for early scoliosis detection.
- Radiographic Cobb angle measurement is the gold standard for scoliosis detection. But the ionizing radiation of radiographic imaging modality makes it unsafe for periodic assessment of curve progression in adolescents.
- Ultrasound imaging modality is radiation-free, real-time, portable, and cost-effective. It is suitable for early scoliosis detection, as the ultrasound and X-ray measurements of spine curvature angles show very good correlations and agreement. But ultrasound images are associated with low contrast and speckle noise, which make it challenging to be used for spine curvature assessment.

2.3 Scoliosis assessment parameters

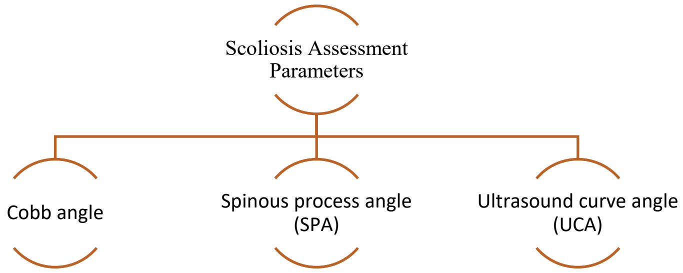


Fig 2.8. Scoliosis Assessment Parameters

2.3.1 Cobb angle

The standard clinical measurement for adolescent idiopathic scoliosis is the coronal Cobb angle (Cobb 1948), measured from the end-plates of the end vertebral bodies in a standing radiograph (Fig. 2.9). This measurement of posterior-anterior (PA) column structures describes the PA spinal deformity (Herzenberg et al. 1990). A Cobb angle greater than 10° indicates scoliosis, and the progression of a curve is indicated by an increase of at least 5° between radiographs from consecutive clinical visits indicates (Lonstein & Carlson 1984).

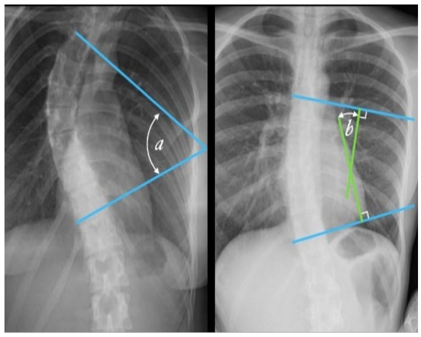


Fig 2.9. X-ray Cobb angle measurement (https://radiopaedia.org/articles/cobb-angle)

Though the Cobb angle is the gold standard measurement technique for assessing spinal curvature on radiographs and is used as the basis for clinical treatment decisions, it has some major drawbacks. Firstly, the intra-observer and inter-observer errors in Cobb angle measurements are typically 3–5° and 5–7°, respectively (Morrissy et al. 1990), (Pruijs et al. 1994). Secondly, a relatively large difference in accuracy was found between the anterior and the posterior vertebral regions, since several anatomical landmarks on the posterior arch, such as the transverse and/or spinous processes, maybe barely visible on the X-ray images, which causes reconstruction error that leads to results discrepancies (Mitulescu et al. 2002). Thirdly, for the manual Cobb angle measurement process, the measurement accuracy is mostly dependent on the awareness and practice of the observer and the quality of images.

2.3.2 Spinous Process Angle (SPA)

The spinous process angle (SPA) (Herzenberg et al. 1990) is a spinal measurement technique that links the tips of the spinous processes to construct a continuous curve (Fig 2.10 (Zhou et al. 2020)), which can be used to obtain a curvature measurement. The posterior deformity of the spine can be defined using this measurement. Some researchers have confirmed the high correlation between the SPA and the Cobb angle for moderate curves (20° to 40°) (Li et al. 2010), (Li et al. 2012).

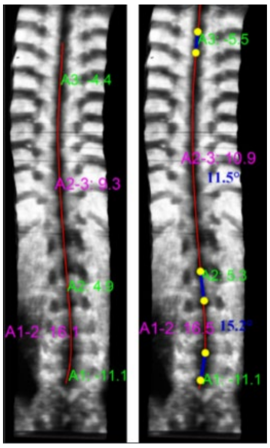
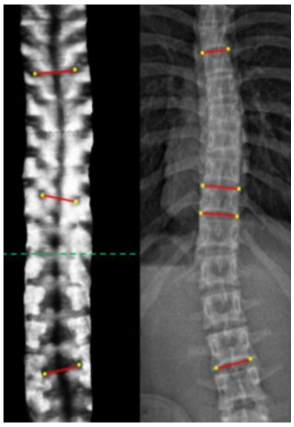


Fig 2.10. SPA measurement (Zhou et al. 2017)

However, the severity of scoliosis with large curvatures can be underestimated by the SPA value as compared to traditional Cobb angles (Brink et al. 2018), (Zheng, Lee, et al. 2016). Also, SPA needs more landmarks than the Cobb angle for measurement and therefore introduces more sources of variation (Morrison et al. 2015).

2.3.3 Ultrasound Curve Angle (UCA)

Lee et al. proposed a new measurement parameter for scoliosis assessment using ultrasound imaging modality and named as ultrasound curve angle (UCA) (Fig. 2.11) (Lee et al. 2021). Lateral bony features are used as the landmarks for UCA measurement. UCA was proved to be highly correlated with X-ray Cobb angle and thus a reliable ultrasound parameter for scoliosis measurement.



(a) (b)

Fig 2.11. (a) UCA measurement with ultrasound image and (b) Cobb angle measurement with X-ray image (Lee et al. 2021)

Key Takeaway:

- Though X-ray Cobb angle is the traditional technique for scoliosis assessment, it has three major drawbacks, i.e. intra and inter-observer error, reconstruction error as transverse/spinous processes are hardly visible on X-ray images and human error in manual measurement.
- SPA is highly correlated with Cobb angle but it has two major drawbacks, i.e. it underestimates the severity of scoliosis with large curvature compared to Cobb angle and it introduces more sources of variation, as it needs more landmarks for measurement.

- UCA is a newly developed scoliosis measurement parameter and establishes a good agreement with X-ray Cobb angle measurement. Unlike SPA, it uses lateral features as landmarks for measurement. UCA gives reliable spine curve measurements.

2.4 Scoliosis assessment methods

Scoliosis is assessed in two steps:

1. Detection of spinous profile column or lateral bony features
2. Calculation of spine curvature angle

In the literature review, I have classified all the scoliosis assessment methods in three categories: (a) Manual method, where all the two steps of scoliosis assessment were done manually, (b) Semi-automatic method, where the land-marking on vertebrae for detection of the spinous profile was done manually and (c) Automatic method, where both the two steps of scoliosis assessment were done automatically without any observer intervention.

The three methods are then further categorized in two sub-groups depending on the imaging modality used for scoliosis assessment: (a) Measurement using X-ray (Parameter: Cobb angle) & (b) Measurement using Ultrasound (Parameter: Cobb angle, Spinous Process Angle, Ultrasound Curve Angle).

2.4.1 Manual method

2.4.1.1 X-Ray Cobb angle

A methodological survey was carried out to determine the reliability of the Cobb angle measurement (Pruijs et al. 1994). In the first part of the research, the variation in radiograph image quality was calculated using Cobb angle measurements made by one investigator on serial radiographs of patients with a fixed spinal curvature following spinal fusion for scoliosis. The second part involved investigating the accuracy of Cobb angle measurement by comparing measurements on the same radiographs of 46 scoliosis patients obtained by three investigators, two orthopedic surgeons, and an orthopedic fellow who was assigned to a school screening project. The study showed that, although different investigators proved the reproducibility of the Cobb angle measurement, the variation in production of a spinal radiograph is an important source of error.

Another study was done to determine the intra-observer and inter-observer reliability of end vertebra definition and Cobb angle measurement using printed and digital radiographs of 48 patients with scoliosis (Gstoettner et al. 2007). The result showed that for the Cobb angle measurement the definition of end vertebrae introduced the main source of error. Therefore, digital radiography could not improve the measurement accuracy for the Cobb angle.

A systemic review was done on measuring procedures to determine the Cobb angle in AIS (Langensiepen et al. 2013). The study showed that reproducibility of manual measurement of X-ray Cobb angle produced high variability in intra and inter-observer agreement. According to the study, automatic measurement procedures were better than manual ones.

2.4.1.2 Ultrasound Spinous Process Angle (SPA) & Ultrasound Curve Angle (UCA)

3D ultrasound is a valuable diagnostic method for scoliosis. Zheng et al. demonstrated that a radiation-free Scolioscan system is reliable for further research and clinical applications for scoliosis (Zheng, Lee, et al. 2016). Coronal plane image is formed by Volume projection images (VPI) method which gives volumetric image within 10 mm depth along the anteroposterior direction. The location of the spinal cord can be detected near the mid-line of VPI. Eventually, it helps to measure the spinal deformity angle, by calculating SPA. Though the angle measured is comparable to the Cobb angle, the feasibility of this method, when dealing with a larger number of scoliosis cases and with different curvatures, is still unknown.

Wang et al. evaluated the reliability and validity of clinical 3D ultrasound imaging on lateral curvature measurement of AIS using the center of laminae (COL) method, with their corresponding MRI measurement using Cobb method (Wang et al. 2015). Though the result showed a good agreement between the angles generated in two imaging modalities, there was no experiment for the applicability of the COL method for a larger clinical trial.

A freehand 3D ultrasound system was developed to eliminate three problems of X-ray imaging: (a) radiation hazard and (b) 2D viewing of 3D anatomy of spine & c) large inter-observer variation (Cheung, Zhou, Law, Lai, et al. 2015). Though the result showed that the ultrasound volume projection imaging method can be a promising approach for the assessment of scoliosis, the manual bony landmark detection using 3D ultrasound underestimated the deformity of the spine compared to the traditional X-ray Cobb angle.

Spinous process (SP) and transverse process (TP) ultrasound images were used to measure the spinal curvature angle and proved to be comparable with the traditional Cobb angle (Brink et al. 2018). The ultrasound angles were 15%–37% smaller as compared with the Cobb angles and therefore, good linear correlations were seen between all ultrasound angles and the Cobb angle.

Coronal 3D-ultrasound images with more lateral features for angle measurement were created using an alternate 3D ultrasound image reconstruction method (Lee et al. 2021). The angle was named ultrasound curve angle (UCA) and unlike using the spinous process as an anatomical reference, thoracic and lumbar bony features were used as a reference for angle measurement. The ultrasound and X-ray measurements of spine curvature showed good correlations and agreement. Though the outcome of this research was very promising, vertebrae structure selection for angle measurement was done by raters and prone to have manual error.

2.4.2 Semi-automatic method

2.4.2.1 X-Ray Cobb angle

A mask-based segmentation algorithm was proposed for semi-automatic X-ray Cobb angle measurement to diagnose scoliosis without human error (Samuvel, Thomas & Mini 2012). In this method, as a first step, the landmark on the centre of each vertebra was identified manually on the X-ray image. In the next step, segmentation of each vertebra was done by placing optimized masks of landmark points. Finally, the Cobb angle was calculated automatically from lower and upper-end vertebrae. As the landmarks were identified manually, this method could not fully eliminate the human factor for Cobb angle measurement.

Another semi-automatic Cobb angle measurement technique using frontal radiographic images demonstrated that this technique was more accurate than the manual technique (Sardjono et al. 2013). Three curve-fitting techniques, piece-wise linear, splines, and polynomials, each with three variants were used to calculate Cobb angle automatically, and the result was compared with manually measured Cobb angle.

Safari et al. proposed a semi-automatic algorithm to measure X-ray Cobb angle to avoid operator-dependent errors (Safari et al. 2019). The semi-automatic algorithm determined the overall structural curvature of the spine, and the Cobb angle was assessed by calculating the angle between two normal lines to the spinal curve at the inflection points of the curve. Though this computer-assisted scoliosis assessment system is reliable and accurate, the imaging modality is not free of radiation and, therefore, not safe for periodic monitoring of AIS.

2.4.2.2 Ultrasound Spinous Process Angle (SPA) & Cobb Angle

According to recent research, semi-automatic measurement of spine angle can be done using the image processing method. In the semi-automatic measurement of SPA, the 6^{th} order polynomial curve fitting method was used to measure the spine curve equation. The main aim of the work was to determine the two main inflection points to determine SPA (Zhou et al. 2016). Since in this method the tangent lines were manually input, human errors were associated with it.

To measure the severity of AIS, a semi-automatic, non-ionizing 3D ultrasound reconstruction method was proposed (Vo, Le & Lou 2019). This method uses 3D ultrasound using a voxel-based reconstruction technique with bilinear interpolation to reconstruct a 3D spinal image and assess the true spinal curvature on the plane of maximal curvature (PMC). The PMC Cobb angles were approximately greater than their corresponding posteroanterior X-ray Cobb angles by 7°. However, this method has some limitations due to ultrasound imaging characteristics. Firstly, the vertebral body could not be reconstructed due to the acquisition configuration and the lack of ultrasound energy penetrating through bone. This prevented the Cobb method from being applied to measure the traditional Cobb angle. Secondly, when the axial vertebral rotation for measuring PMC was large, the areas behind the spinal process did not receive an ultrasound signal and, therefore, could not be imaged.

2.4.3 Automatic method

2.4.3.1 X-Ray Cobb angle

Manual or semi-automatic Cobb angle measurement is sensitive to observer expertise level or experience. A computer-aided semi-automatic Cobb angle measurement system was developed using X-ray imaging modality (Zhang et al. 2010). In this system, a fussy Hough transform technique was used to detect upper- and lower-end vertebrae automatically. The variability of Cobb angle measured by this technique was much less than observer-measured angles. The problem of inter- and intra-class observers is also mitigated in this technique.

An automatic Cobb angle measurement technique was proposed to discard human intervention by automatic slope selection from vertebrae X-ray images (Kundu, Chakrabarti & Lenka 2012). Introduction of an improved version of Non-local Means for image denoisation and Otsu's automatic threshold selection for canny edge detection in this system reduced intra- (44.85%) and inter-observer (72.29%) variability for Cobb angle measurement.

Measurement of Cobb angle from X-ray images using a deep neural network produced satisfactory results (Zhang et al. 2017). Deep neural networks trained by vertebral patches were able to determine the slopes of vertebrae automatically, and the Cobb angle was calculated from these slope values. The Cobb angles measured by this method were compared with those measured manually (≤3°). This system showed an intra-class correlation coefficient greater than 0.98, which indicates that it had high repeatability for measurements of Cobb angle. The vertebral patches were assigned by the user, which is the source of error in this research.

Another deep learning-based X-ray Cobb angle measurement algorithm was proposed where the network segments the spine contour and discards unrelated regions. After that, the Cobb angle was measured from the segmented spine region by the curve fitting method (Tu et al. 2019).

An automatic method was developed to calculate the Cobb angle and classify the type of deformation of the spine using X-ray imaging modality (Janumala & Ramesh 2020). In this method, segments of the spinal cord were detected using a novel speeded-up robust features algorithm, and support vector machine was used to identify the vertebrae for Cobb angle calculation.

Convolutional neural network (CNN) showed a very promising outcome in the automatic calculation of Cobb angle from spine X-ray images (Pan et al. 2019). In the most recent and advanced work, the Cobb angle was measured using the minimum boundary rectangle method (Horng et al. 2019). In this work, vertebrae were segmented using three CNN architectures, i.e. U-Net, Residual U-Net and Dense U-Net, of which Residual U-net gave the best results. The accuracy of Residual U-net was as high as 97%, but this method is useful only with X-ray images. An efficient and accurate AIS detection system was proposed where multi-scale CNN (Deeplab V3+) was used to segment vertebrae and calculate Cobb angle successfully (Liu et al. 2021).

2.4.3.2 Ultrasound Spinous Process Angle (SPA)

Ultrasound and artificial intelligence were integrated to deliver a safer and more accessible alternative to X-ray imaging for scoliosis detection. Zhou et al. worked on automatic measurement of spine curvature using 3D ultrasound image pre-processing with phase congruency and newly developed two-fold threshold strategy (Zhou et al. 2017). This method overcame the limitations of manual measurement of SPA. e.g.

variations in measurements. However, the computational time of this method is significant, as the program spends most of the time computing the phase congruencies of the images. In conclusion, it is robust enough to handle a large group of images, but it cannot help with faster processing and detection.

Automatic detection of the location of vertebral landmarks is very crucial for the diagnosis and periodic monitoring of scoliosis. Deep learning-based end-to-end object detection algorithm, Single Shot MultiBox detector, was applied for vertebrae landmark detection using ultrasound images (Deng & Huang 2019). The preliminary experiment results on the phantom showed high accuracy in vertebra landmark detection. CNN has been applied to segment the spine from ultrasound scan images and calculate ultrasound-based angles (Ungi et al. 2020). This method was efficient and accurate, as it could complete the measurement process in less than 1 minute with only 2.2° maximum error compared to X-ray-based measurement.

The essential part of the calculation of SPA is detecting the spinous process (SP). Recent research has developed CNN-based stacked hourglass network for automatic segmentation of SP from coronal spine ultrasound images (Zeng, Ge, Gao, Zhou, Zhou, He & Zheng 2021). The measured SPA in this method was comparable to the gold standard radiographic Cobb angle. But there are a few factors, such as the different standing posture of patients and poor image quality of ultrasound, which affected the SPA measurement process.

Table 1.3 Summary of various scoliosis assessment methods

	Manual	Semi-automatic	Automatic
Cobb Angle	✓	✓	✓
SPA	✓	✓	✓
UCA	✓	×	×

Table 2.3 shows a summary of various scoliosis measurement techniques with respect to the three measuring parameters.

Key Takeaway:

- The scoliosis assessment technique consists of two main parts: (a) detection of spinous profile column or lateral bony features and (b) calculation of spine curvature angle.
- The manual method of scoliosis assessment has a few major disadvantages: (a) unable to deal with a large number of patients' data, (b) intra- and inter-observer error in angle measurement, and (c) human error in vertebrae landmark detection.
- The semi-automatic method of scoliosis detection is not free from human error in vertebrae landmark identification.

- Automatic scoliosis assessment is comparatively fast and free from human error, but this method is only available for X-ray Cobb angle and ultrasound SPA measurement. There is no evidence of research on automatic UCA measurement.

2.5 Application of Convolutional Neural Network on ultrasound image

The biggest challenge in clinical diagnosis using ultrasound imaging is high inter- and intra-operator variability. As a result of high inter-operator variability, many cancer drug trials do not accept ultrasound-derived tumour measurements. Using machine learning techniques with the ultrasound imaging can help in successfully diagnosing several diseases. However, low resolution and speckle noise are the two major problems associated with ultrasound imaging. Thus, pre-processing should be done on the acquired ultrasound image. Recently, deep learning techniques have been used to extract numerical features from ultrasound images. Vedula et al proposed a convolutional neural network to transform speckled, blurry ultrasound images into better quality images (Vedula et al. 2017). Automated ultrasound image analysis promises to play a crucial role in addressing some of these challenges.

The 3D ultrasound images of different depths have different definitions, especially for the middle spine line. All the images with varying depths do not give useful and detailed information about the spinal cord. To select the best image, a convolutional neural network (CNN), a unique kind of deep neural network, is used. The use of a CNN for analysis of medical images is the recent trend. CNN architecture is the most commonly used deep learning technique for analysing medical images. Generally, a CNN consists of several convolutional layers (Zeiler & Fergus 2014), max-pooling layers (Szegedy et al. 2015) and fully connected layers (Krizhenvshky, Sutskever & Hinton). The activation methods used in a CNN include the rectified linear unit, sigmoid and tanh (Jarrett, Kavukcuoglu & LeCun 2009).

In the past few years, deep learning, especially the CNNs (LeCun, Kavukcuoglu & Farabet 2010), has proven to be a reliable image classification technique. The CNN does not require the manual design of features because of its end-to-end characteristics. Its performance in terms of image feature extraction is excellent (Zeiler & Fergus 2014). It uses self-learning features of the convolution kernel to directly input the original image and obtain the classification result. Moreover, by increasing the number of layers, the

CNNs can extract considerably complex and detailed image features. Generally, the lower layers of the CNNs are found capable of extracting some general features of images, whereas the higher layers extract more specific features. Evan et al. demonstrated that the performance of the CNN can be significantly improved using cascading features of the different convolution layers in the CNN (Long, Shelhamer & Darrell 2015). Moreover, transfer learning and fine-tuning are used to train the CNN requiring large amounts of data, a technique lacking in the field of medical images (Phan et al. 2016). Now CNNs and transfer learning are being used for imaging breast cancer (Wahab, Khan & Lee 2017), cancer recognition(Wahab, Khan & Lee 2017), and blood cell classification (Wahab, Khan & Lee 2017). Transfer learning involves pre-training the CNN on a large dataset to learn general weights, after which these weights are transferred by fine-tuning the learning rate from pre-trained CNN for training smaller data. CNNs perform well in terms of classification. Each of the three CNNs can outperform others or present defects under different cases. Moreover, a difference in their performance can be observed owing to the difference in the sizes of their convolutional kernels and network structures.

CNNs have different architectures for various applications. Classic architectures – LeNet-5, AlexNet, VGG-16 – are made of simply weighted convolution layers, while the modern ones – ResNet, DenseNet – are more complicated and capable of efficient learning. The CNN models give remarkable enactment on ImageNet, a large visual database (Krizhenvshky, Sutskever & Hinton). Zeiler et al. highlighted that the ImageNet model works best when the softmax classifier is retrained (Zeiler & Fergus 2014).

A few popular CNN architectures are as follows:

1. AlexNet (Krizhenvshky, Sutskever & Hinton), a convolutional network, was introduced as a competitor in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012. It comprises eight layers that can extract numerous features, as well as five convolutional layers, some of which are followed by max-pooling layers, and three fully connected layers with a final 1000-way softmax. The fully connected layer is used for mapping the extracted distributed feature representations to the sample space and weight features; these layers only have 4096 neurons each.
2. GoogleNet (Szegedy et al. 2015) won the 2014 ILSVRC challenge. It succeeded in reducing ImageNet's top5-error rate to 6.67%. GooleNet employs an inception architecture to identify an optimal local sparse structure in a convolutional neural

network. An inception architecture can use different receptive fields of convolution kernel sizes of 1×1, 3×3 and 5×5 to capture scale-invariant features.

3. ResNet50 (He et al. 2016) is another efficient CNN proposed for image recognition. ResNet50 uses residual blocks to realize deep residual learning through feed-forward neural networks with shortcut connections. Shortcut connections skip one or more layers, where the output of short connections is added to the output of stacked layers. Residual blocks deepen the number of network layers; this results in the stark distinction of the extracted features of each layer, which is especially helpful for capturing high-level details.

The image features are extracted by different convolution kernels situated in various layers of the CNN; these differ in terms of their hierarchical nature. The convolution operation of the output of the previous layer and the specific convolution kernel obtained from learning generate each feature map. The featured activity in intermediate layers must be interpreted and the performance contributions of different layers must be understood in order to improve the CNN performance and adjust network parameters. However, training the CNNs on small quantities of data is extremely difficult. Bengio used weights to transfer knowledge in networks and improve performance (Bengio 2012). In general, the learning rate influences the speed of updating the CNN weights. Extremely high learning rate destabilizes the CNN weights, following which the loss function fails to converge. On the other hand, extremely low learning rate slows down the updating of weights, thereby reducing the efficiency.

Litjens et al. observed that deep learning is being increasingly used in almost every aspect of medical image analysis and that the convolutional neural network is one of the most accurate and succinct among all the deep learning methods (Litjens et al. 2017). Many previous research works have been done on cardiovascular (Yu et al. 2017), thyroid nodule (Liu et al. 2017), kidney (Q Zheng October 2018), foetal skull segmentation (Cerrolaza et al. 2018) ultrasound image analysis using CNN. The application of CNN on spinal cord image analysis is a new approach.

The inherent disadvantage of an ultrasound image is its high speckle noise and low contrast, making feature extraction from ultrasound images quite challenging. Deep learning methods, especially CNN, can extract underlying features from the images, but a large dataset is needed for the training process. Liu (2017) revealed that deep transfer learning features give very good results in thyroid nodule classification (93% accuracy). According to Zheng (2016), integrating deep learning techniques into kidney ultrasound image data can increase the performance of image classification and, consequently, the detection of kidney diseases. It is quite ideal to apply deep learning to ultrasound medical imaging.

A two-stage CNN was used to detect tumours from breast ultrasound images (Huang, Han, et al. 2019). The two-stage CNN and the refined process were robust enough to remove inherent speckle noise and low image contrast of ultrasound image. Refined region of interest (ROI)-CNN+Grading CNN gave 99% accuracy in the classification of Category 3 breast tumour, but again there is no such experiment with spine ultrasound images.

Key Takeaway:

- Automatic scoliosis assessment techniques can eliminate the measurement errors caused by human intervention.
- Convolution neural network (CNN) has been proven superior in the field of biomedical image processing.
- CNN can be applied for automatic feature extraction from ultrasound images with low contrast and speckle noise.
- CNN showed successful outcomes in the segmentation and detection of various organs from noisy ultrasound images.

CHAPTER III: GENERATION AND COLLECTION OF ULTRASOUND SPINE IMAGES

3.1 Introduction

The overall objective of this research is to develop an automatic scoliosis assessment system using ultrasound spine images and artificial intelligence. The first step in the automatic diagnosis of scoliosis, as described in Fig 1.4 (Chapter I), is the generation and collection of ultrasound spine images. This chapter aims to describe in detail the generation and collection of ultrasound spine image dataset used for this research.

Ultrasound is a cost-effective, radiation-free, and real-time imaging modality which better diagnoses scoliosis than X-ray or magnetic resonance imaging (MRI). Various modes of ultrasound are used in medical imaging (shown in Fig. 3.1).

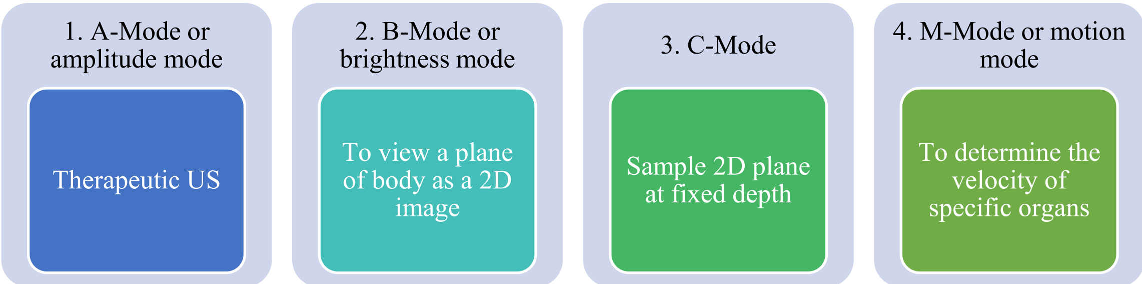


Fig 3.1. Types of Ultrasound Imaging

3.1.1 B-Mode ultrasound imaging

B-mode ultrasound imaging is a popular diagnostic tool in the field of medicine. It is capable of obtaining images in real-time using non-ionizing radiation, poses no risk to patients, and is reasonable compared to other medical imaging modalities (Szabo 2004).

Various physical occurrences can result in an ultrasound B-mode image. The most common cause of this is when the probe emits a short ultrasound pulse, some waves get

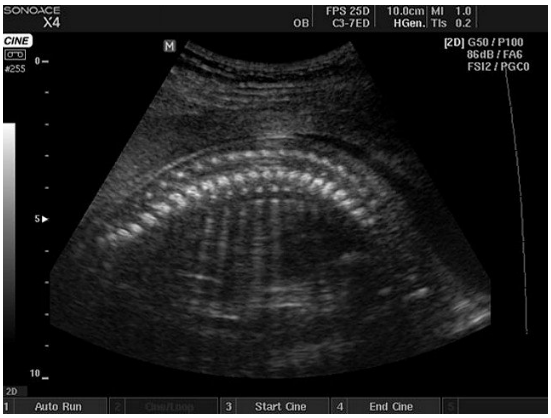


Fig 3.2. B-Mode ultrasound foetus spine image (https://www.medison.ru/ultrasound/gal414.htm)

reflected to the transducer when a tissue with different acoustic impedance is encountered; the remaining waves penetrate into the tissue deeply. The probe then detects the backscattered ultrasound waves electronically, resulting in an image if the intensity is proportional to the strength of the return echo (Szabo 2004). In short, B-Mode is a two-dimensional ultrasound image display in which the ultrasound echoes are represented by bright dots. The amplitude of the returned echo signal determines the brightness of each dot. Fig 3.2 shows a foetal spine image captured using B-mode ultrasound.

Table 3.1 summarizes the various applications of ultrasound imaging in the diagnosis of spinal deformation. According to Hwang et al., the application of ultrasound imaging modality in the assessment of spinal health has been a popular subject of research in recent times (Hwang et al. 2021). Tawfik et al. established that, in identifying main spinal abnormalities in infants, the diagnostic value of spinal ultrasonography was equivalent to that of MRI (Tawfik et al. 2020). In another research, Zhang et al. showed that, for long-term spinal deformity treatment, such as cases of severe scoliosis, ultrasound imaging is safe and gives patient comfort along with a clear intrathecal structure for guided treatment (Zhang et al. 2021). Ultrasound is also applied in pre-procedural imaging of central neuraxial blockade of the spine to identify, in real-time, the ideal trajectory (best angle, direction of approach, and depth) and to optimize the subsequent invasive treatment with fewer needle passes and skin punctures (Kalagara et al. 2021).

Table 3.1 Application of ultrasound imaging in diagnosis of spinal deformities

Author	Field of work	Region of interest (ROI)	Comments
Hwang et al. 2021	Application of ultrasound in traumatic spinal cord injury (SCI)	Spinal cord	Ultrasound imaging can serve as a powerful adjunct to various developing therapies for SCI
Tawfik et al. 2020	Comparison of spinal ultrasound with MRI for the diagnosis of spinal deviation in infants	Spinal cord and bony elements	Spinal ultrasound can be used as a first-line screening investigation for infants with spinal deviation
Zhang et al. 2021	Nusinersen through lumbar puncture with real-time ultrasound guidance in spinal muscular atrophy (SMA) patients with severe scoliosis.	Lumbar area	100% success rate (no major complications) has been achieved in radiation-free and real-time ultrasound-guided lumbar intrathecal administration of nusinersen in SMA patients with severe scoliosis
Kalagara et al. 2021	Central neuraxial blockade using ultrasound imaging	Mid-spine line, vertebral level, interlaminar space, epidural and intrathecal spaces	1) Ultrasound increases the success rate and ease of neuraxial block performance. 2) Ultrasound usage for neuraxial procedures reduces the risk of traumatic procedures and, thus, may increase safety

3.1.2 3D B-Mode ultrasound imaging

A radiation-free, freehand 3D ultrasound system named Scolioscan was invented by Professor Zheng and their team to detect scoliosis using volume projection imaging (VPI) (Zhou & Zheng 2015); (Zhou et al. 2017). Scanning of patients is done using an ultrasound probe, and the B-mode images, as well as the information pertaining to their corresponding position and orientation, are collected. This information is used for reconstructing a 3D image. On the other hand, the images presenting the coronal view of the spine are recorded using volume projection imaging (Zheng, Lee, et al. 2016).

For further processing, a single 3D VPI ultrasound image is sliced into nine 2D coronal images of different depths and qualities. Though the nine 2D images are of different qualities, they look very similar, especially the adjacent images. For further processing of the images for scoliosis assessment, selecting the best image among the lot is necessary.

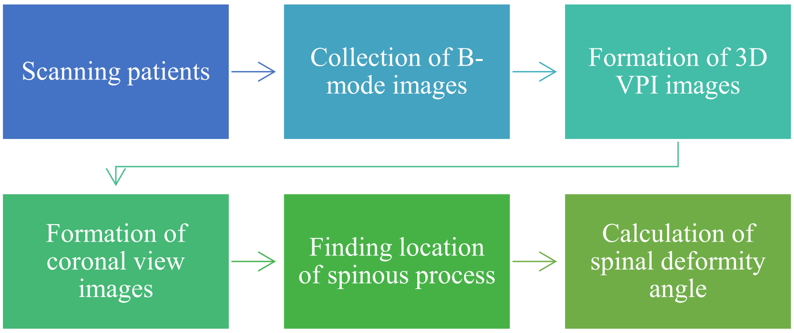


Fig 3.3. Flowchart of current (manual) scoliosis assessment technique using Scolioscan system

3.2 Manual assessment of scoliosis using Scolioscan

Among all available techniques, the 3D ultrasound volume projection imaging is the best one to visualize spine anatomy. Zheng et al. illustrated that Scolioscan is a reliable and radiation-free system to measure spine deformity using 3D ultrasound imaging for adolescent idiopathic scoliosis patients (Zheng, Lee, et al. 2016). The result is very satisfactory when compared with the conventional radiographic Cobb method. Fig 3.3

Fig 3.4. A Scolioscan System (Zheng, Lee, et al. 2016)

presents a complete flowchart of the existing manual scoliosis assessment procedure using Scolioscan.

3.2.1 The Scolioscan system

The Scolioscan system (Model SCN801, Telefield Medical Imaging Ltd), shown in Fig 3.4, developed in Hong Kong, is used to generate 3D volume projection images (VPI) depending on conventional 3D ultrasound imaging technique. The frame of the instrument is rigid, but it has two movable supporting boards and four supporters to help

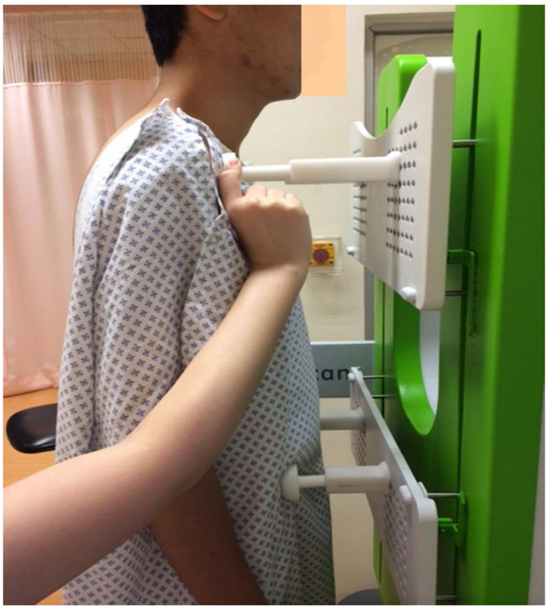


Fig 3.5. Four supporters helping patient maintain a stable posture; the supports are adjusted against the chest and hip boards, and their lengths are adjusted accordingly (Zheng, Lee, et al. 2016)

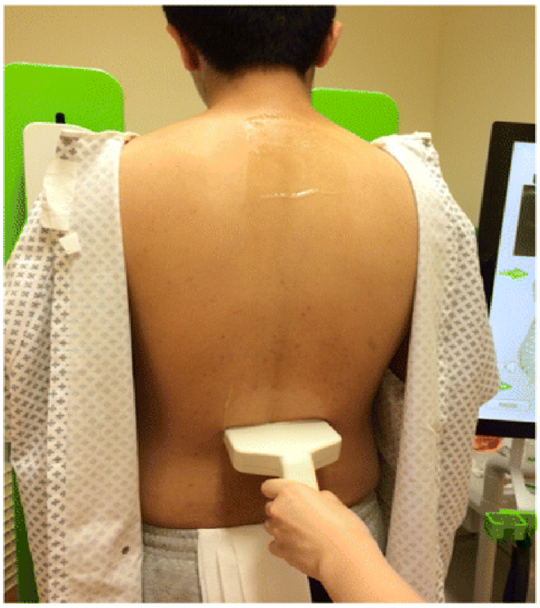


Fig 3.6. Scanning procedure with the Scolioscan probe, Ultrasound gel is applied on the area to be screened (Zheng, Lee, et al. 2016)

the patients maintain a stable standing posture during a diagnosis, as shown in Fig 3.5. The chest and hip boards are adjustable vertically to fit patients with various heights. The four supporters with adjustable lengths are fixed on the boards by implanting to the fixation holes and sealed by turning the supporter by 90°. For follow-up assessment of a single patient, the board location (vertical), support position (vertical and horizontal), and supporter length is recorded.

3.2.1.1 Scanning procedure

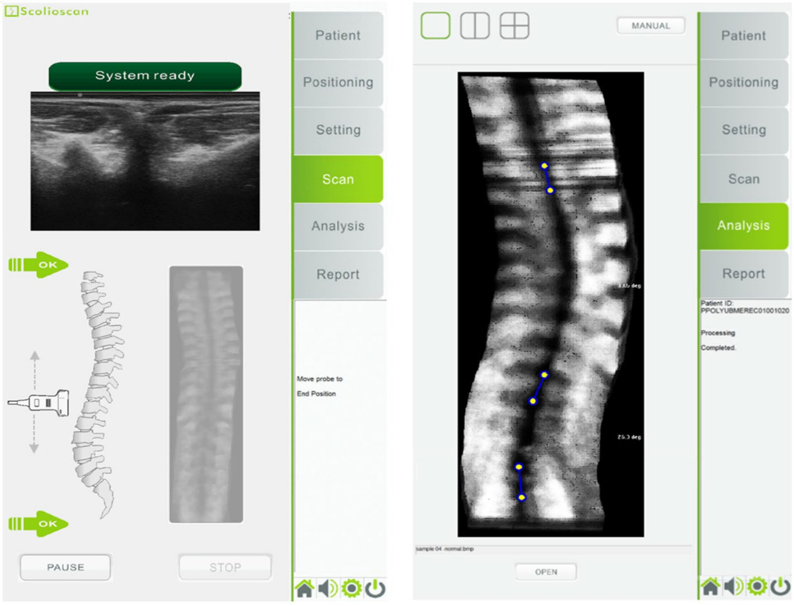
A custom-designed linear probe (frequency: 4–10 MHz), 10-cm wide, is used as the ultrasound probe for freehand scanning and generation of a 3D ultrasound image of the spine. Inside the ultrasound probe, an electromagnetic spatial sensor is fitted to sense the position and orientation of the probe. The electromagnetic transmitter is placed inside the transmitter box, as shown in Fig 3.4. Fig 3.6 shows the scanning procedure of a patient's spine. During scanning, the probe is progressed from bottom to top of the patient's back to scan the whole spine. The system also has two LCD screens and one touch screen that is used by the operator for uploading patient information, setting parameter during the scan, generating VPI images, conducting measurements and producing reports (Zheng, Lee, et al. 2016).

There is another screen at the back to provide information to the patient. The screen also has a green eye spot with location set according to the height of the patient to assist the them to keep their head and neck stable during scanning. Not only that, the screen also provides information regarding the different steps of assessment procedures to keep the patient aware of the system, thereby making the whole system more accommodating. It is attached with the propriety software for scanning, VPI image analysis and angle measurement of the spine.

The overall evaluation process involves the following:

1. Patient information registration
2. Adjusting supporters
3. Setting up the ultrasound scanner
4. Reporting

Before starting a scan, it is necessary to determine the upper and lower boundary of the scanning area. The ultrasound probe, therefore, is placed at the bottom and top of the scanning area to record the lower and upper boundaries of the scan, respectively. The location of the moving probe with respect to the upper and lower boundaries is shown in real-time on the interface to guide scanning.

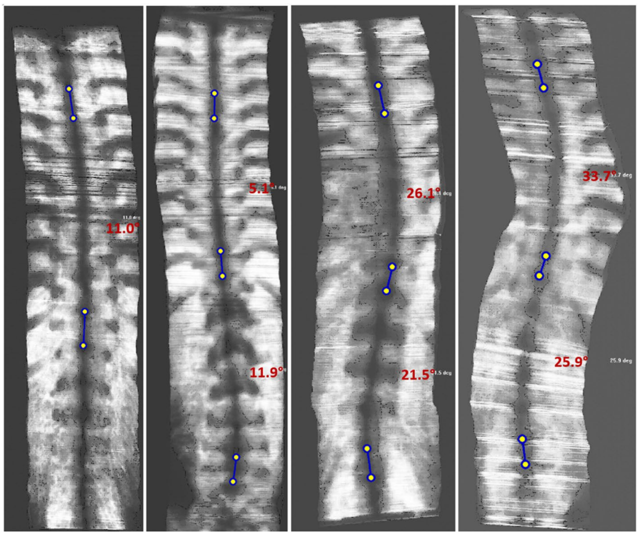


(a) (b)

Fig 3.7. Software interfaces for (a) Scanning and (b) VPI image analysis and angle measurement of Scolioscan (Zheng, Lee, et al. 2016)

3.2.1.2 Image processing and manual assessment of scoliosis angle

For collecting the B-mode images, the corresponding position and orientation information are used. All the images together construct 3D images, and volume projection imaging (VPI) is used to form coronal view images for further analysis of the spine (Cheung, Zhou, Law, Mak, et al. 2015). A volume projection image is a volumetric image that uses the average intensity of all voxels within a selected depth (approximately 10 mm) along the anteroposterior direction to form an image in the coronal plane. Fig 3.7 shows Scolioscan’s interfaces used for scanning, as well as the VPI image analysis and angle measurement, during measurement. The curve near the middle line of the 3D VPI image, which represents the location of the spinous process, is considered the spine. This spine curve is used to measure the spinal deformity, and the deformity angle is called the



(a) (b) (c) (d)

Fig 3.8. (a) 4 spine VPI images showing the coronal plane with different levels of spine deformity with the two lines manually drawn (b,c,d) the three lines to form two pairs to measure the curvatures of spine in the thoracic and lumbar regions, respectively (Zheng, Lee, et al. 2016)

Scolioscan angle (Fig 3.8), which is similar to the Cobb angle of each patient. The two most skewed portions of the spine curve should be identified as the most tilted vertebrae for scoliotic angle measurement. The local slant of the spine curve is measured by drawing two small lines (blue lines in Fig 3.8) manually from the middle of the curve on the coronal image. The orientation of these two manually drawn lines is used to estimate the Scolioscan angle.

3.3 Moving towards automatic assessment of scoliosis

3.3.1 The ultrasound image dataset and pre-processing

The 2D volume projection images were collected using Scolioscan. During the scanning, the frame rate of the system was 60 frames/sec and the average scanning speed of the ultrasound probe from the bottom to the top along the spine was 2.0 cm/sec. The experimental procedures involving human subjects were approved by the Institutional Review Board. The subjects gave informed consent for their inclusion in this study as required, and the work adheres to the Declaration of Helsinki.

A total of 109 images collected from 109 patients (82 females and 27 males) with an age range of 15.6 ± 2.7 years and having different degrees of spine deformity were used retrospectively. The mean body mass index (BMI) of the subjects was 18.3 ± 2.1 kg/m^2. None of the patients had neuromuscular or congenital problems at the time and received standard diagnosis and treatment. Nine 2D coronal images of different depths were extracted from a single 3D ultrasound VPI image (Cheung, Zhou, Law, Mak, et al. 2015). Since the quality of individual images greatly varies, human experts were employed to manually assess the clarity of the lateral bony features visible on an image and select the best image for each patient. Subsequently, 109 2D vertebral coronal images formed the input dataset for this work, and each 2D image was then resized uniformly to 2574 × 640-pixel and converted to the '.png' format.

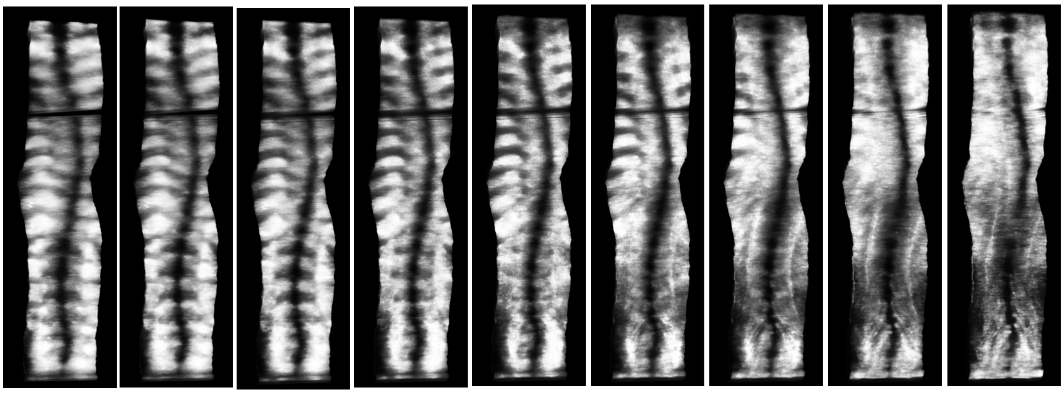


Fig 3.9. The sample of nine 2D ultrasound images in different depths of a 3D spine image

3.3.2 Generation of 2D VPI images

A patient's 3D volume image generated by Scolioscan is cut into nine 2D coronal images from different depths of the cut plane. As shown in Fig 3.9, though the 2D images are of different imaging qualities depending on the depth of the respective image, they look very similar, especially the adjacent images. Moreover, all nine images do not possess a good quality of information. So, for identification of bony features and assessment of the Scolioscan angles, good quality images should be selected from them. This image selection process is done manually by human experts.

The 3D spine ultrasound image is sliced into nine 2D vertebral anatomical images based on imaging depths (Fig 3.9). The images of different depths are of different imaging qualities. Cheung et al. introduced the method of generation of 2D coronal projection images from Scolioscan, using a narrow-band, non-planar volume rendering algorithm and is known as volume projection imaging (VPI) (Cheung, Zhou, Law, Mak, et al. 2015) (Fig 3.10). The 3D ultrasound images have different sizes according to the patients. The details of the dataset used for this research are summarized in Table 3.2.

An experienced operator from the Hong Kong Polytechnic University marks the best images manually. The best images (first 3 images from left, in Fig 3.9) are selected based on the criteria of a clear, dark line in the middle representing the spine profile, as well as other spinal features, as clearly as possible in the image, including transverse processes and ribs.

Table 3.2 Summary of the data set used in the research

Particulars	Details
Medical device	Scolioscan system which gives 3D ultrasound images as output
Patient profile	109 patients (82 females and 27 males) with an age range of 15.6 ± 2.7 years
Pre-processing	Conversion process of 3D to 2D image using Volume projection imaging technique
Image profile used as input for the research	
Number of images	109
Size in pixel	2574 × 640
Image file format	*.png (portable graphics format)

3.4 Challenges in processing spine ultrasound image

The popularly used imaging modalities for scoliosis detection are CT, MRI, X-ray and ultrasound. Among them, CT, MRI and X-ray give high-quality outputs that are easier to process. Though ultrasound imaging modality has its advantages (safe, scalable and economic), it has some drawbacks associated. These drawbacks are compounded when the spine is to be imaged.

3.4.1 Scan noise

During the acquisition of ultrasound VPI images, mostly, the operator moves the ultrasound probe on the back of the patients efficiently and gradually at a constant speed. As discussed in Section 3.2.1.1, the probe scans from bottom to top, touching the skin along the spine. The ultrasound gel acts as a lubricant between the probe and skin surface. The probe resistance at certain parts of the spine, such as protruded spinous process, abruptly change due to several reasons. The reasons may include the resistance and

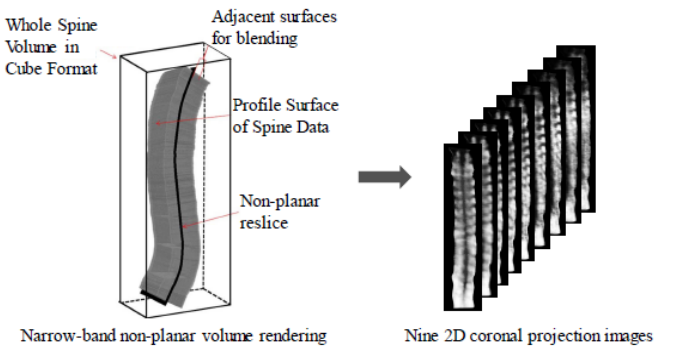


Fig 3.10. Narrow-band, non-planar volume rendering algorithm for 2D image generation

fluidity of the skin. Because of this resistance change, the speed of the probe might rise suddenly and generate fewer ultrasound frames in some particular areas.

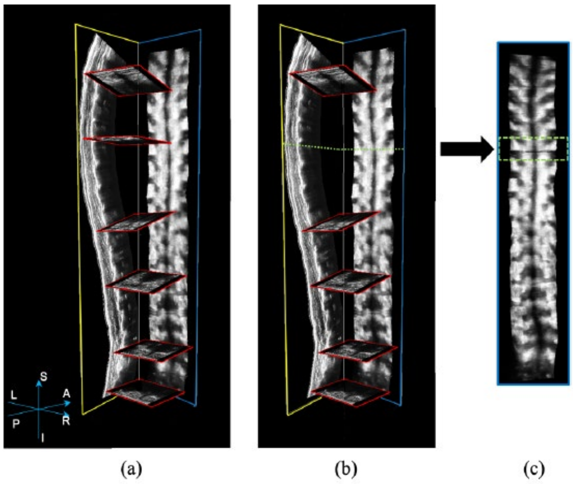


Fig 3.11. Example of US Scan noise. (a) 3D VPI image acquired from manual scan. The red frames shows multiple B-mode images, who piled up to form 3D image, the blue and yellow frames show the coronal and sagittal images respectively. (b) green dotted lines shows the missing B-mode image, which results the scan noise, (c) coronal projected image with scan noise (noisy area in green box) (Huang et al. 2022)

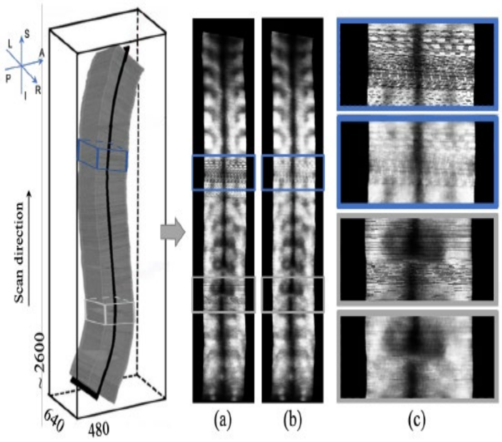


Fig 3.12. An example of ultrasound scan noise. Generation of VPI image by non-planar volume rendering from 3D US volume data (Cheung, Zhou, Law, Mak, et al. 2015). (a) original VPI image with scan noise, (b) VPI image without scan noise, (c) Scan noise and corresponding clear areas in blue and grey boxes (Huang et al. 2022)

As a result, the recreated 2D coronal images face strong scan noise, as shown in Fig 3.12a. Nevertheless, scan noise in VPI images may be caused by various factors such as ultrasound operators, probing and VPI settings. Fig 3.11 demonstrates the generation of the noise. The scan noise makes scoliosis assessment further challenging. Some steps were undertaken to mitigate the impact of ultrasound scan noise. However, they had certain drawbacks, as enumerated in Table 3.3.

Table 3.3: Steps to mitigate scan noise and their drawbacks

Steps taken to mitigate US scan noise impact	Drawbacks
Increasing the sampling frequency during scanning	High sampling frequency causes a higher requirement of scanners and thereby increases the budget in clinical applications
Algorithm-based approach to recover the VPI image from scan noise	High computational complexity makes this approach unfeasible in real-time diagnosis

Therefore, the more efficient and effective way to deal with the noise in 2D VPI images is not during the scanning process but through applying effective segmentation techniques.

3.4.2 Variability of bony features

The locations, shapes, and sizes of the bony features are different (Fig 3.13a, b & c) as are the UCA (Fig 3.13d, e & f). The slopes between two endpoints of two adjacent bony features are also different (Fig. 3.13d, e, & f). The variability of LBFs and TBFs presents a unique challenge.

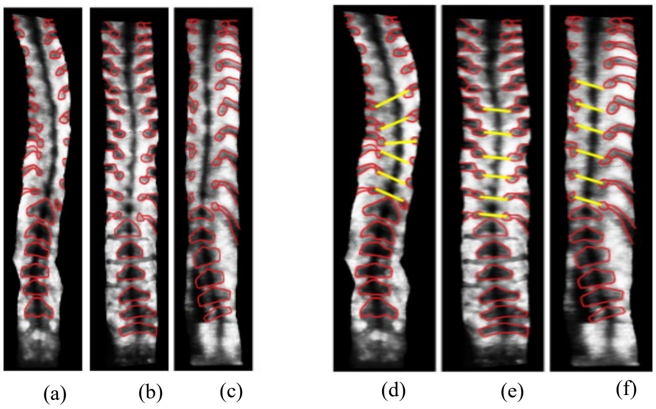


Fig 3.13. (a), (b), and (c): Lateral bony Features are of different shapes, sizes, and locations, (d), (e), and (f): Slopes between two adjacent bony features (thoracic) are different.

3.5 Identifying ROIs from ultrasound spine image

The overall flowchart for the automatic assessment of scoliosis is illustrated in Fig. 1.4 in Chapter I. The next step in the process is to identify suitable ROIs from the ultrasound spine image dataset. Image segmentation using deep learning is researched to help demarcate the regions of interest.

Ultrasound imaging continues to grow, and advances are continuously being made to the transducer design, digital systems, spatial/temporal resolution and portability. Alongside these developments, ultrasound image processing and segmentation also continue to find their place in providing an important tool in computer-aided diagnosis, therapy and image-guided interventions (Meiburger, Acharya & Molinari 2018). More recently, the importance of providing fully automated localization and segmentation techniques has become a hot topic in the field of research as the amount of data to be analysed continues to grow, and the capability of processing large multi-institutional databases proves to be an asset for large projects and studies. Clinically, these techniques can also aid in the diagnosis and treatment of patients, providing both a tool for locating potential areas of interest and for quantitatively measuring important clinical information. Automated localization techniques refer to those methods that provide an initial segmentation and localization of an object of interest within which a fine segmentation technique is used to obtain the final quantitative information. Depending on the specific application, fine segmentation techniques can be employed within a determined region of interest or on the entire ultrasound frame without the use of a localization method to first isolate the general area (Meiburger, Acharya & Molinari 2018).

In this research, a 3D volume projection image is sliced into nine 2D coronal plane images of different depths for scoliosis detection. One such 2D ultrasound spine image is shown in Fig. 3.14a. Fig. 3.14b illustrates the various segments of the spine – the thoracic bony features, rib, T12 level, and lumbar bony features. Figs. 3.14c and 3.14d outline the current methodology used by a medical expert to calculate the ultrasound curve angle (UCA). The most tilted thoracic bony features (TBFs) and lumbar bony features (LBFs) are identified by experts. Further steps involve drawing lines through the centre of the identified TBF and LBF pairs (Fig 3.14c) and allocating lines to measure the UCA (Fig 3.14d) (Lee et al. 2021). Hence, the clear segmentation of bony features is an important step for UCA measurement.

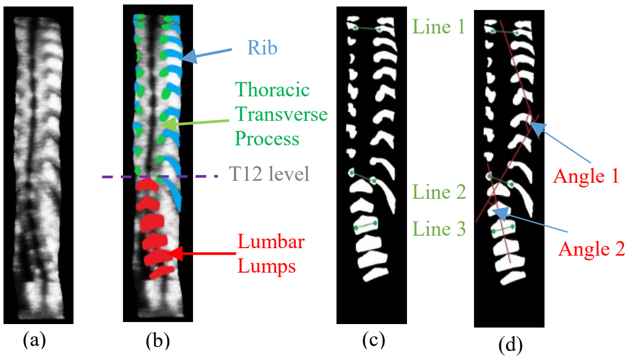


Fig 3.14. Ultrasound curve angle (UCA) measurement (a) original ultrasound spine image, (b) marked lateral bony features with three types of anatomical features, (c) & (d) Illustration of methodology

3.5.1 Need for a suitable segmentation architecture

Since ultrasound spine images suffer from low contrast, high scan noise and a high degree of variability, conventional segmentation architectures such as U-Net, Attention U-Net, MultiResUNet are not specialized to deal with the inherent drawbacks or to successfully segment these images. Hence, as a part of the research, two new segmentation algorithms were developed and are detailed in Chapter IV and Chapter V.

CHAPTER IV: LIGHT-CONVOLUTION DENSE SELECTION U-NET (LDS U-NET) FOR ULTRASOUND LATERAL BONY FEATURE SEGMENTATION

4.1 Introduction

As described in Chapter III, there are two serious challenges associated with any ultrasound image that degrade the quality of output images. They are (i) speckle noise,

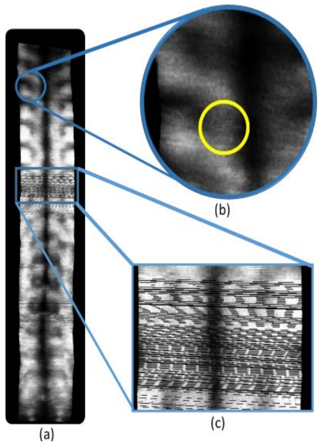


Fig 4.1. Various noise in ultrasound spine image (a) raw image, (b) and (c) types of speckle noises

which is caused by reasons such as air gap between the device and patient's body (Fig 4.1b), inconsistencies of ultrasound gel layer during scanning, improper scanning speed (Fig 4.1c), or other system losses; and (ii) low image contrast produced because the speed of sound varies for various tissues, and it is often difficult to separate fat and water-based tissues. As speckle noise appears as information, it makes the clinical data hard to differentiate (Loupas, McDicken & Allan 1989), and the low contrast of the image brings unnecessary distractions for medical practitioners.

4.1.1 Medical image segmentation using CNN

Machine learning techniques have been successfully applied in the field of medical diagnosis using ultrasound. Vedula et al proposed the transformation of speckled, blurry ultrasound images into better quality images using a convolutional neural network (CNN) (Vedula et al. 2017). In the past few years, CNNs have been successfully used in various biomedical image processing tasks, including image classification (Krizhevsky, Sutskever & Hinton 2012),(Szegedy et al. 2015),(Simonyan & Zisserman 2014), feature extraction (Zeiler & Fergus 2014), and segmentation (Ronneberger, Fischer & Brox

2015), (Long, Shelhamer & Darrell 2015), (Badrinarayanan, Kendall & Cipolla 2017), (Hu et al. 2017). Semantic segmentation has become the prior interest area in medical imaging (Long, Shelhamer & Darrell 2015),(Badrinarayanan, Kendall & Cipolla 2017),(Girshick et al. 2014),(Li, Wu, et al. 2017),(Sun & Wang 2018). However, automatic semantic segmentation of biomedical images could be difficult when there is the variability of shapes and sizes of the anatomy between patients as well as low contrast of surrounding tissues (Roth et al. 2018).

Among many machine learning techniques, the U-Net architecture was particularly successful in biomedical image segmentation (Ronneberger, Fischer & Brox 2015). U-Net can efficiently segment images with a very limited number of the annotated training dataset. A U-Net consists of a multi-layer deep encoder network that extracts spatial features from the image, and a corresponding multi-layer deep decoder network that up-samples the feature maps to predict the segmentation masks. It uses the self-learning property of the convolution kernel to input the original image and obtain the classification result. By increasing the number of layers, the U-Net can extract considerably complex and detailed image features. Generally, the shallower layers of U-Net were found capable of extracting some general features of images, whereas the deeper layers could extract more specific features. U-Net and U-Net like models have been used successfully in segmenting 2D or 3D ultrasound images of breast lesion (Amiri et al. 2020), human placenta (Han et al. 2019), liver (Thomson et al. 2019), spine anatomy (Banerjee et al. 2020; Huang et al. 2020; Lyu et al. 2021), kidney (Ravishankar et al. 2017), etc.

MultiResUNet (Ibtehaz & Rahman 2020), as an improved version of U-Net, was introduced to segment very challenging images that are not possibly done by basic U-Net, such as images having irregular shapes with multi-scaled features. In its architecture, Inception blocks (Szegedy et al. 2015) were introduced, in place of the conventional convolution layers of U-Net, to help with multilayer feature extraction. Also, to reduce the semantic gap between encoder-decoder features, the normal convolution layers were replaced with residual connections that made training easier (Szegedy, Ioffe, et al. 2016).

Attention U-Net (Oktay et al. 2018) improves the conventional U-Net with an attention gate for medical imaging that automatically learns to focus on target structures of varying shapes and sizes. Because of the small number of training data, dense connectivity is needed in biomedical image processing. Dense connectivity can be successfully incorporated within the encoder and decoder path (Guan et al. 2019),(Jin et al. 2017). Another semantic segmentation using Atrous convolutions was introduced using the

DeepLab family for better segmentation (Chen, Papandreou, Kokkinos, et al. 2017),(Chen, Papandreou, Schroff, et al. 2017),(Chen, Zhu, et al. 2018).

Recently, there is an increasing need for implementing deep learning techniques for medical problems on mobile phones, embedded systems, or any PC with high diagnostic accuracy and with a low computational requirement. Most of the CNN models are over-parameterized and need high computing power and memory for training and inferencing (Denil et al. 2013). Depthwise separable convolution layers are the solution for this problem (Sifre & Mallat 2014). Depthwise separable convolution layers are successful in forming the image classification models in two ways: (a) They give better models (e.g., Xception model (Chollet 2017)) than conventional convolution layers, with a considerably smaller number of parameters, and (b) they reduce the memory space requirement when reducing the number of parameters (e.g., the MoblieNets family of architectures (Howard et al. 2017)). Also, it has been found that the regular convolution operations can be considered to be equivalent to the depthwise separable convolution operations (Guo et al. 2018). Hence, comparing to conventional convolutional layers, using depthwise convolution layers will have similar performance in terms of accuracy but require a smaller number of training parameters.

4.1.2 Approach to designing a suitable architecture

Inspired by the above work, we aim to develop a low-computation novel hybridized deep learning architecture to suitably segment the bony features in ultrasound spine images. The approach to designing the architecture is to focus on handling the drawbacks of ultrasound spine image as efficiently as possible. The proposed architecture has three main aspects:

a) the basic U-Net structure is adopted as the network architecture, but the conventional convolutional layers are replaced with dense depthwise separable convolution layers to increase the computational efficiency
b) selection gates (Oktay et al. 2018) are deployed for a smarter identification of the target bony features
c) the encoders-decoders are connected using multi-scale skip-pathways (Ibtehaz & Rahman 2020) to enhance feature fusion.

The segmentation result of the proposed architecture is compared quantitatively and qualitatively with the basic U-Net (Ronneberger, Fischer & Brox 2015), Attention U-Net (Oktay et al. 2018), and MultiResUNet (Ibtehaz & Rahman 2020) models.

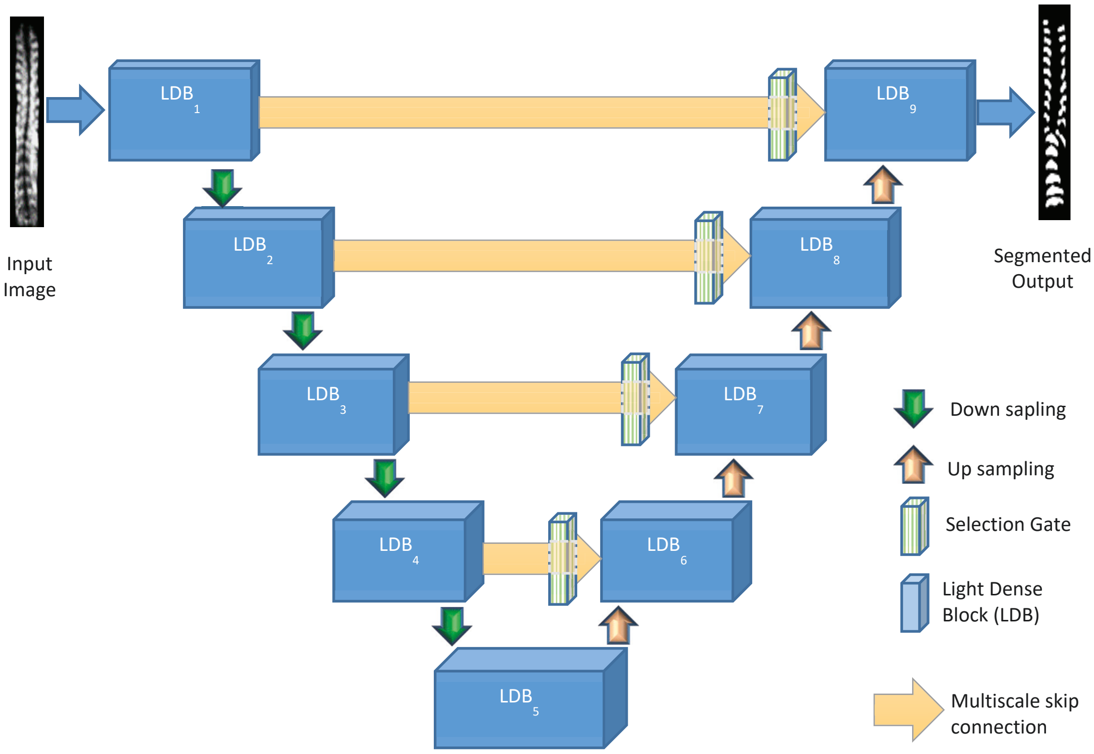


Fig 4.2. Proposed Architecture of LDS U-Net

4.2 Methodology of LDS U-Net

Our objective is to automatically segment, using a suitable CNN architecture, the lateral bony features (LBFs and TBFs) in an ultrasound spine image, which is plagued with speckle noise and low contrast. This chapter has two main contributions:

a) Development of a suitable architecture that can produce better segmentation results vis-à-vis contemporary models while handling the inherent drawbacks of ultrasound images.

b) Enhancement of the computational efficacy for a lightweight architecture.

In a nutshell, a lightweight version of U-Net that contains densely connected depthwise separable convolution followed by pointwise convolution, multiscale skip connection, and selection gates is proposed. It is inspired by several salient features used in other models such as the U-Net, MultiResUNet, and Attention U-Net. Fig. 4.2 illustrates the architecture of the proposed network of Light-Convolution Dense Selection U-Net (LDS U-Net). This network model is built on the concept of depthwise separable convolution and has three main features: (a) Novel Light Dense blocks, (b) improvised multi-scaled skip connections, and (c) selection gates. The details of depthwise separable convolution and the three features are given as follows.

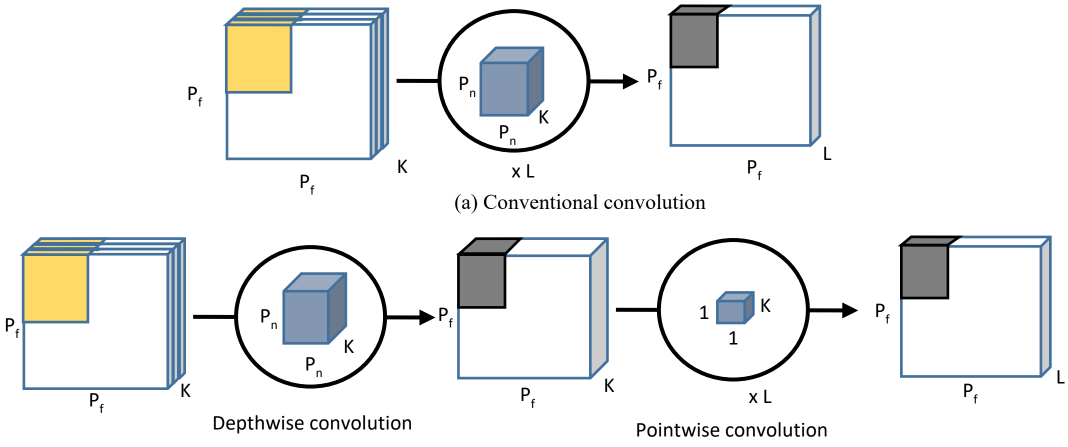


Fig 4.3. Schematic representation of Conventional vs. Depthwise Separable Convolution

4.2.1 Depthwise separable convolution

Depthwise separable convolutions lessen the number of parameters and computation used in convolutional operations while increasing representational efficacy (Kaiser, Gomez & Chollet 2017). This kind of convolution can be applied to information such as spatial, depth dimensions, and the number of channels. While a normal convolution deals with a single convolution operation, a depthwise separable convolution separates a kernel into two different kernels that do two convolution operations, namely, the depthwise convolution and the pointwise convolution. In the depthwise convolution, a spatial convolution is done independently over each channel of the input. It is followed by a pointwise convolution, where a 1×1 convolution is done to map the depthwise channel output into a new channel space (Chollet 2017).

A standard convolution layer works by applying a convolution kernel to all channels of the input image and takes a weighted sum of the input pixels covered by the kernel sliding across all input channels of the image. This means that for a standard convolution, no matter how many input channels are available, the number of output channel is one. However, in depthwise separable convolutions, features are only learned from the input channels. So, the output layer has the same number of channels as the input.

Suppose in a convolution operation, the input is of size $P_f \times P_f \times K$ with the feature map f and generates an output of size $P_g \times P_g \times L$ with the feature map g, where P_f is the spatial width and height of the input feature map, K is the number of input channels, P_g is the spatial width and height of the output feature map, L is the number of output channels. Then, for a conventional convolution operation (Fig. 4.3a.) with the convolution kernel N of size $P_n \times P_n \times K \times L$, where P_n is the spatial dimension of the kernel, the computation cost is given by the equation:

$$G_c = P_n \,.\, P_n \,.\, K \,.\, L \,.\, P_f \,.\, P_f \tag{4.1}$$

On using the depthwise separable convolution (Fig. 4.3b.) the computation cost will be the aggregation of depthwise and pointwise convolutions and is given by the equation:

$$G_d = P_n \,.\, P_n \,.\, K \,.\, P_f \,.\, P_f + K \,.\, L \,.\, P_f \,.\, P_f \tag{4.2}$$

Combining equations (4.1) and (4.2), the reduction of computation can be represented by the equation:

$$G = \frac{G_d}{G_c} = \frac{P_n \,.\, P_n \,.\, K \,.\, P_f \,.\, P_f + K \,.\, L \,.\, P_f . P_f}{P_n \,.\, P_n \,.\, K \,.\, L \,.\, P_f \,.\, P_f}$$

$$= \frac{1}{L} + \frac{1}{P_n^2} \tag{4.3}$$

Equation 4.3 represents the reduction in computation requirements of the depthwise separable convolution compared to the conventional convolution; resulting in considerably lower computing and parameter cost of the network.

As conventional U-Net, with its few layers, is not deep enough to perform this particular segmentation task, adding more layers directly and making it deeper, may solve the segmentation problem. But deeper neural network tends to develop gradient vanishing and redundant computation in network training (Bi et al. 2017). To overcome those associated problems, few modifications are required to enhance the learning process of the network.

4.2.2 Light dense block

The first modification is related to the conventional Dense network which uses regular convolutional layers and has the advantage of parameter simplicity, vanishing-gradient minimization, and feature reuse (Pleiss et al. 2017). In this chapter, a proposed Light Dense block is designed as the main building block of Light-Convolution Dense Selection U-Net (LDS U-Net). The basic structure of the Light Dense block is shown in Fig.4.4 and

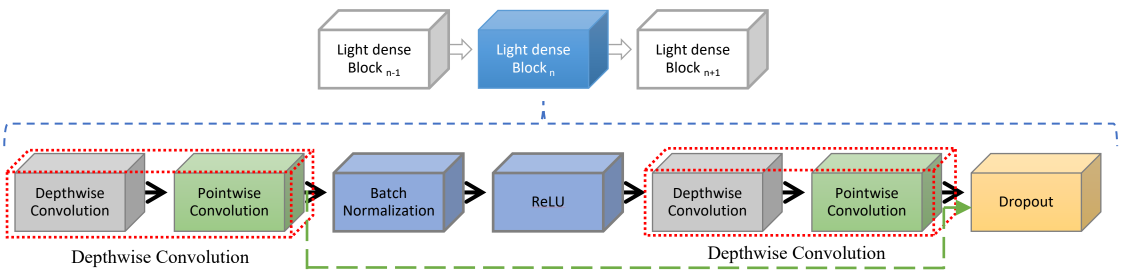


Fig 4.4. Proposed Light Dense

here, unlike a conventional Dense network, all the convolutional layers are depthwise separable convolution layers.

The first layer of the Light Dense block is a depthwise convolution unit which consists of a depthwise convolution block followed by a pointwise convolution block. The next layers are the batch normalization, Rectified linear unit (ReLU) activation function, another depthwise convolution unit, and a dropout layer. The first depthwise convolution unit is also connected densely to the dropout layer, as shown by the green dotted arrow in Fig.4.4. Through this new design, the Light Dense block delivers the same advantages as a conventional Dense block but with a smaller number of parameters.

4.2.3 Multi-scale path

In the basic U-Net architecture, there are skip pathways between the respective layers of the encoding and decoding side, and shortcut paths before the max-pooling layers in the encoder side and after the deconvolution layers in the decoder side. Often spatial information gets lost during the max-pooling operation, and the skip connections help the network to propagate information from the encoder side to the decoder side. However, the skip pathways often come up with a problem of the semantic gap during the feature fusion because the first layer of the encoder, which extracts the low-level features, is connected to the terminal layer of the decoder, which deals with more high-level features. Also, because of the added complexities of variabilities in sizes, shapes, and positions of bony features, both the low and high level features would have to be retained for detailed segmentation.

To reduce the discrepancy between the encoder-decoder features and to enhance the feature fusion, a multi-scale skip path is proposed in this chapter. Multi-scale inception (Ibtehaz & Rahman 2020) module is incorporated between the encoder and decoder layers to enhance the low-level features extracted in the encoder side. By this connection, the low-level features will undergo further processing before merging with the high-level features on the decoder side. Moreover, instead of the usual convolutional layers, multi-scale inception layers are used. They improve the utilization rate of computing resources by increasing the depth and width of the network while keeping the computational budget constant (Chen, Bentley, et al. 2018).

Inception modules have been proved to be very promising in enlarging receptive fields and capturing more context information (Liu & Huang 2018). It enhances the depiction capability of low-level features. The inception module adopts multiple branches with different kernel sizes to capture multi-scale information. This methodology is key to deal with the problem of handling the high variability of shapes, sizes, and positions of bony features in the spine ultrasound images. However, as the inception module is very computationally demanding, the normal convolution operation in the traditional inception module is replaced by the depthwise convolution in this proposed model.

In the proposed Light Dense block, a sequence of two 3×3 depthwise convolutional layers is used for doing feature extraction. In the skip path, a sequence of 3×3 convolution blocks is used, as shown in Fig 4.5, instead of bigger 5×5 and 7×7 blocks. Therefore, the outputs from the three 3×3 convolution blocks are concatenated to enhance the receptive field and reduce the semantic gap between the encoder and the decoder. A residual connection of 1×1 convolution block is also presented in the skip path to make the learning procedure stable.

In the proposed inception module, the technique to control the number of filters of the convolution layers inside the block was adapted (Ibtehaz & Rahman 2020). A parameter P is assigned to control the number of filters as given by equation (4.4),

$$P = \lambda \times F \tag{4.4}$$

where, F is the number of filters in the corresponding layers like the basic U-Net, and λ is a scaler coefficient. The numbers of filters are set to F = [32, 64, 128, 256, 512] along with the 5 layers of the LDS U-Net architecture respectively, and λ is chosen as 1.67 to ensure that the model structure is similar to the basic U-Net. The numbers of filters are set to $\frac{P}{6}$, $\frac{P}{3}$, and $\frac{P}{2}$ in the three corresponding convolutional layers for extracting multiscale features.

4.2.4 Selection gate

In an ultrasound spine image, the noise in the extraneous regions can appear as information and can hinder the segmentation process. To tackle this problem smartly and

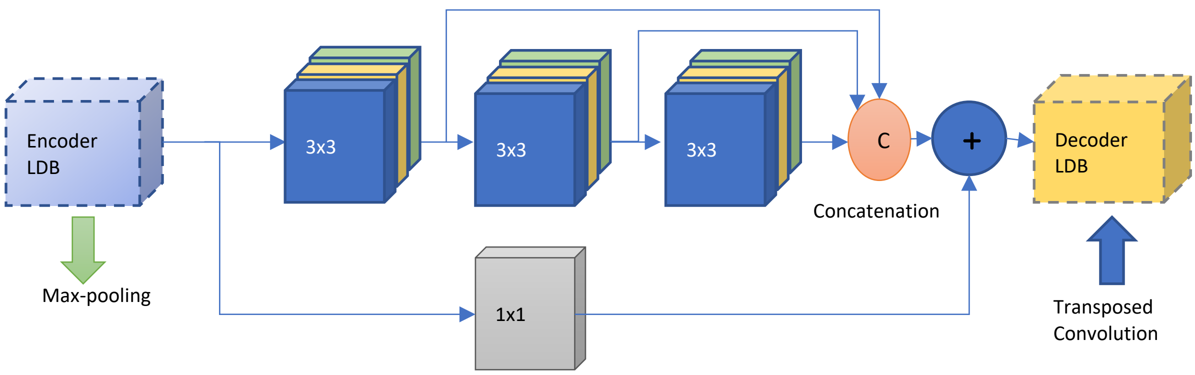


Fig 4.5. Multi-scale inception module

efficiently, selection gates are applied in the proposed network. The selection gate (Oktay et al. 2018) discards extraneous regions in the input spine ultrasound image to ignore the associated noise and allows for preferential attention to the target bony features to select the relevant features of importance. A conventional selection gate, improvised with depthwise separable convolution layers, is integrated into the proposed architecture as shown in Fig 4.6. It aims to reduce the computational overhead while increasing the segmentation accuracy.

The Light Dense convolutional layers extract more deep features when it processes more deep layers gradually. Suppose after processing the n-th layer, the generated feature map is x^n. As shown in Fig. 4.6, a is the attention coefficient that identifies and extracts only the desired parts of the image for better segmentation, F_n is the number of feature

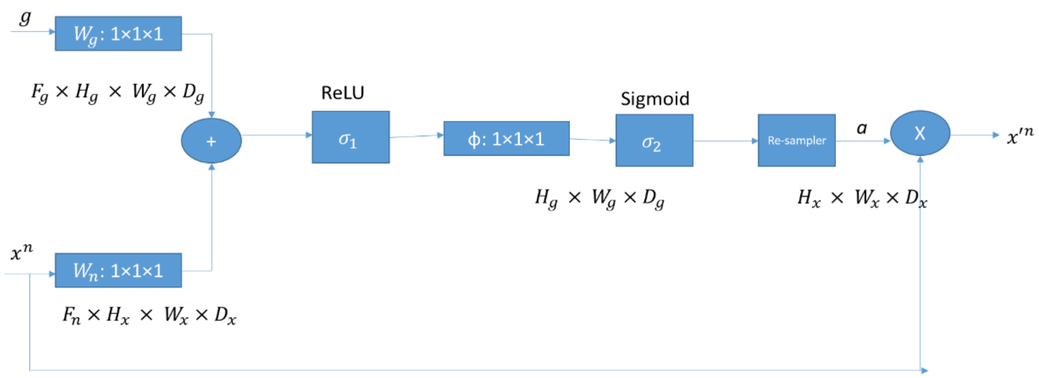


Fig 4.6. Selection Gate

maps in layer n; H_x, W_x, D_x are the height, width, and dimension of the n-th layer respectively; g is a gating vector which is booked from the preceding layer of the network, which is a coarser layer; F_g is the number of feature maps in the layer g; H_g, W_g, D_g are the height, width, and dimension of g respectively.

The selection gate works with two input vectors: x^n and g. g has a lower dimension and better feature representation as it comes from a deeper layer compared to x^n. x^n , after processing by a stride convolution, and g, after processing through a $1 \times 1 \times 1$ convolution, merge elementwise. After that, a ReLU activation layer σ_1 and a $1 \times 1 \times 1$ convolution operation take place that reduces the dimension of the resultant vector. Also, $\sigma_1\left(x_{i,c}^n\right) = \max\left(0, x_{i,c}^n\right)$, where i and c stand for the spatial and channel dimensions respectively. Then, a sigmoid layer σ_2 scales the vector to the range [0, 1] to generate the attention coefficient a, where $\sigma_2\left(x_{i,c}\right) = \frac{1}{1+\exp(-x_{i,c})}$ is the sigmoid activation function. Grid resampling of attention coefficients is done using trilinear interpolation. A value of near to 1 indicates more significant features. The attention coefficient is multiplied element-wise to x^n after resampling. The output of the selection gate is $x'^n_{i,c} = x^n_{i,c} . a^n_i$.

4.2.5 Ablation study

Ablation experiments were carried out to gauge the contribution of each feature of the LDS U-Net. Three models were successively developed before arriving at the final segmentation architecture. These intermittent models were independently evaluated using the available dataset.

The first intermitted model was a modified U-Net with light dense block and named as Light Dense (LD) model. In this model, the basic convolution layers of U-Net were replaced by the newly developed light-dense blocks. In the second intermitted model, trainable selection gates were introduced, which was trained to isolate the relevant areas of interest, amidst noisy areas and facilitated flow of only the relevant information within the network. This is named as Light Convolution Selection (LCS) model. The third intermitted model was developed from the LD model by introducing multiscale skip-paths which would allow seamless propagation of richer information through the network. This model is named as Light-Dense Inception (LDI) model. In the final version, i.e. LDS model, both the selection gate and multi-scale path are included for the segmentation of bony features with various shapes and sizes.

4.3 Experimental setup

4.3.1 Dataset

The input images used in this research are collected using the Scolioscan system (Model SCN801, Telefield Medical Imaging Ltd, developed in Hong Kong) and is described in Chapter III in detail.

The truth mask for each 2D coronal image is made by experts from Hong Kong Polytechnic University. Some sample input 2D coronal images along with their expert generated truth masks are shown in Fig 4.7. The labelling of the truth mask is done based on some key features. 1) There should be six lumbar bony features (LBF). 2) The thoracic bony features (TBFs) in the lumbar region are not labelled as they are not generally visible in ultrasound images. 3) If the last pair TBFs i.e. the T12 levels is not visible, it is labelled according to the judgment of the experts. The input raw datasets of 109 spine ultrasound images are randomly split into one training set of 79 images and one testing set of 30 images.

4.3.2 Pre- and post-processing & data augmentation

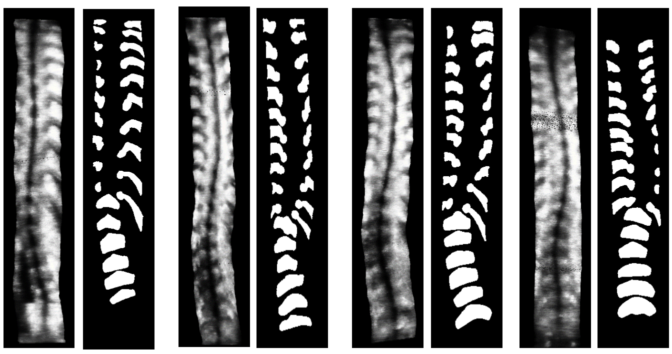
The size of each raw ultrasound image is 2574 × 640 pixels and is resized to 256 × 64 pixels through image pre-processing, maintaining the aspect ratio of the original image. The resized ultrasound images are randomly flipped and rotated for data augmentation. The objective of the experiment is to develop a segmentation architecture and to benchmark its performance against the original U-Net, Attention U-Net, and MultiResUNet. This requires no specific pre-processing except that the input images are resized to fit into the GPU memory, and the pixel values are divided by 255 to bring them into the [0,1] range. During post-processing, the image is resized back to the original size of the raw images to ensure that the image size does not impact the ultrasound curve angle (UCA) calculation for scoliosis.

4.3.3 Implementation details

Anaconda, or more specifically, Spyder software, is used to conduct the experiments. The network models are implemented using Keras (Chollet 2015) with Tensorflow backend (Abadi et al. 2016). The experiments are conducted in a GPU laptop with NVIDIA GeForce RTX 2060.

The general working principle of any semantic segmentation algorithm is to investigate each pixel and anticipate whether it represents a point of interest or merely a part of the background. Alternatively, this principle can also be treated as a pixel-wise binary classification problem with the objective of the segmentation algorithm to minimize the binary cross-entropy loss function.

For an image A, let the corresponding Truth Mask (TM) be B, and the predicted segmentation output be B'. For a pixel ma, the TM value is b_{ma} and the network predicted output is b'_{ma}. The binary cross-entropy loss for that image is defined as:



(a) (b) (c) (d)

Fig 4.7. Four Image sets with each having a raw image plus a Truth Mask

$$Cross\ Entropy\ (A, B, B') = \sum_{ma \epsilon A} (-(b_{ma}) \log(b'_{ma}) + (1 - b_{ma}) \log(1 - b'_{ma})) \quad (4.5)$$

For a batch containing p images, the loss function L becomes,

$$L = \frac{1}{p} \sum_{i=1}^{p} Cross\ Entropy\ (A_i, B_i, B'_i) \quad (4.6)$$

The goal of the model is to minimize the binary cross-entropy loss and the model is trained using the Adam optimizer (Kingma & Ba 2014). The Adam optimizer adaptively computes different learning rates for different parameters from estimates of the first and second moments of the gradients. All the models, used in this research, are trained up to 120 epochs since 120 epochs are found to be the saturation point for model accuracy, and no further progress is observed beyond this point. Finally, 3-fold cross-validation of the dataset is used to validate the consistency of the model.

4.3.4 Evaluation metrics

For quantitative performance evaluation, four very popular evaluation indices are employed - Jaccard similarity (JS) (Jaccard 1912), DICE coefficient (DC) (Dice 1945), F1 Score (Huang et al. 2015), and Pixel Accuracy:

$$Jaccard\ Similarity = \frac{J \cap \hat{J}}{J \cup \hat{J}} \quad (4.7)$$

$$Dice\ coefficient = 2 \left(\frac{J \cap \hat{J}}{J + \hat{J}} \right) \quad (4.8)$$

$$F1\ Score = \frac{2TP}{2TP + FP + FN} \quad (4.9)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FP} \quad (4.10)$$

where $\hat{J}$ is the predicted segmentation output from the method to be evaluated, and J is the expert suggested Truth Mask. The J contours are references for further segmentation analysis (Abbasian Ardakani et al. 2019). *TP, TN, FP*, and *FN* are truly positive, true negative, false positive, and false negative, respectively. True positive denotes the pixels present in both truth mask and predicted segmented region. True negative denotes the pixels present in neither segmented truth mask nor predicted segmented region. False-positive signifies the pixels present only in the predicted segmented region. False-negative signifies the pixels present in the suggested truth mask only.

4.4 Analysis of performance of LDS U-Net

This section is presented in two parts. In the first part, quantitative and qualitative comparisons of the LD model (Light Dense model), LDI model (Light-Dense Inception model), and LCS model (Light Convolution Selection model) are shown together with Light-Convolution Dense Selection U-Net (LDS U-Net) model to evaluate the importance of key features used in the segmentation of spine ultrasound images with variable shapes and sizes of bony features and to assess the overall effectiveness of the proposed network. In the second part, an extensive analysis for the newly proposed model and comparisons with basic U-Net, Attention U-Net, and MultiResUNet are made to evaluate the performance of LDS U-Net to other contemporary models.

4.4.1 Evaluation of performance of key features

Three key features are used in the proposed LDS U-Net model – Light Dense blocks, multi-scale paths, and selection gates. At the onset, it is important to assess if these features play an important role in achieving the objectives. Hence, during the ablation study, models are made by isolating the desired key features and comparing the semantic segmentation performance. At first, both the multi-scale paths and selection gates are removed from the main model. The model is named as Light Dense (LD) model and it successfully segments the thoracic and lumbar bony features. Next, the selection gates are added with the LD model, and it is named as Light Convolution Selection (LCS) model. As an alternate improvement of the LD model, the conventional shortcut connection is replaced with the inception skip path to form the Light Dense Inception (LDI) model. Finally, all the above models are implemented together as the Light Dense Selection U-Net (LDS U-Net) model.

Table 4.1: Quantitative evaluation of ablation study

Method	Avg. Jaccard Index (Std dev)	Avg. Dice Score (Std dev)	Avg. F1 Score (Std dev)	Avg. Accuracy (Std dev)
LD	0.7123 (±0.039)	0.8204 (±0.036)	0.8412 (±0.030)	0.9108 (±0.025)
LCS	0.7280 (±0.037)	0.8292 (±0.033)	0.8532 (±0.031)	0.9203 (±0.021)
LDI	0.7339 (±0.034)	0.8419 (±0.032)	0.8704 (±0.029)	0.9367 (±0.020)
Proposed LDS U-Net	**0.7415 (±0.03)**	**0.8694 (±0.028)**	**0.8885 (±0.025)**	**0.9592 (±0.020)**

4.4.1.1 Quantitative Evaluation (Ablation Study)

Table 4.1 outlines the segmentation performances for the ablation study. The LCS model generates better segmentation than the LD model (1.07% better Dice Score). It implies that the model with a selection gate can detect the target lateral bony features better than the LD model. The LDI model further improves the performance by introducing the multi-scale path with a 1.53% higher Dice Score than the LCS model. Finally, both the selection gate and multi-scale paths are included in the LDS U-Net model, and it generates a Dice Score of 3.26% higher than that of the LDI model. In terms of Jaccard Index, F1 score, and accuracy, the proposed LDS U-Net performs the best. Hence, the LDS U-Net model is chosen as the proposed model. This model also gives the smallest standard deviations in the four evaluation indices.

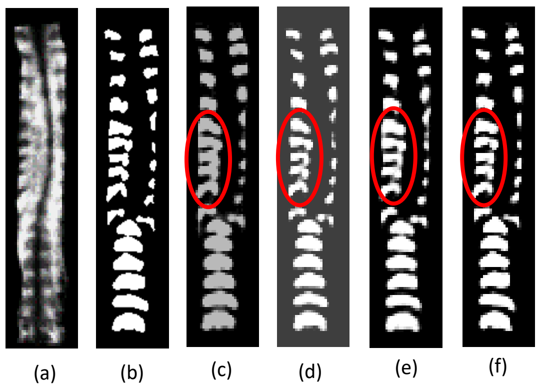


Fig 4.8. Individual segmentation result of ultrasound spine image: (a) Raw image, (b) Truth mask, (c) LD model, (d) LCS model, (e) LDI Model, and (f) proposed LDS U-Net model.

4.4.1.2 Qualitative Evaluation (Ablation Study)

The aim is to segment and identify the bony features from the ultrasound spine image for automatic scoliosis detection. The LD model successfully segments the thoracic and lumbar bony features as shown in Fig. 4.8c. Fig 4.8d shows the LCS model that gives better results in the segmentation of the left thoracic bony features (Nos. 5, 6, 7, 8 & 9) as compared to Fig. 4.8c. It implies that the model with a selection gate detects the target bony features better than the LD model. To further improve the LCS model, the conventional shortcut connection is replaced with the inception skip path to form the LDI model. From Fig. 4.8e, it is clear that the LDI model produces better segmentation of the left thoracic bony features as compared to the LD and LDC models. Finally, the proposed LDS U-Net model, as shown in Fig 4.8f, generates more accurate boundary segmentation of the left thoracic bony features than the LD, LCS, or LDI models. Fig. 4.9 shows several more qualitative visual comparisons of bony feature segmentation using the LD, LCS, LDI, and proposed LDS U-Net models.

4.4.2 Comparison of LDS U-Net with other contemporary models

The LDS U-Net is compared with three architectures, namely U-Net, Attention U-Net and MultiResUNet. Each of these models has its unique advantages. However, when it

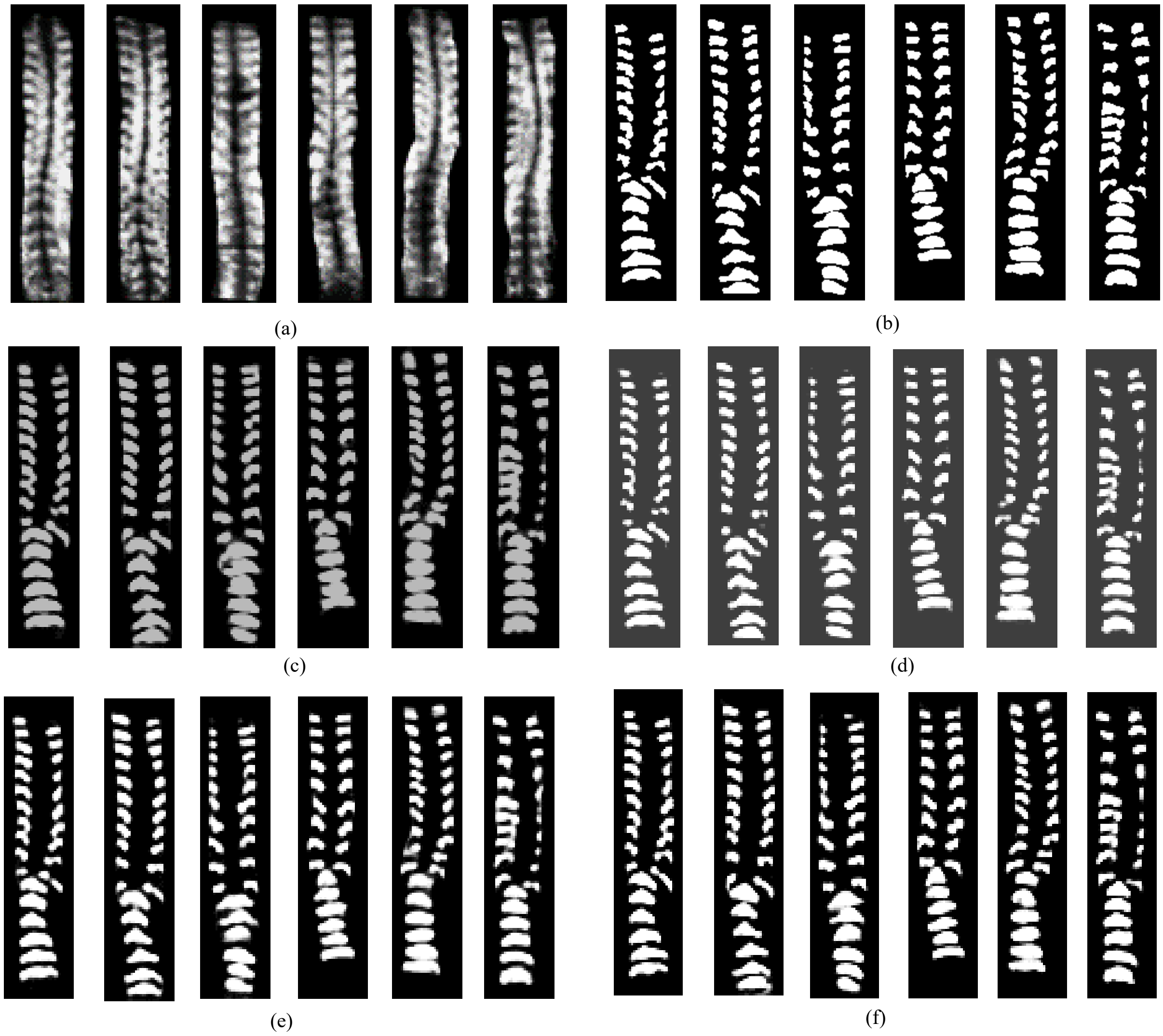


Fig 4.9. Qualitative comparison of 6 different case models: (a) Raw image (b) Truth mask, (c) LD model, (d) LCS model, (e) LDI Model, and (f) proposed LDS U-Net model

comes to the segmentation of a noisy ultrasound image, these architectures might not support deep-level data extraction and high image variability.

4.4.2.1 LDS U-Net outperforms U-Net, Attention U-Net and MultiResUNet in the segmentation of ultrasound spine image dataset

Table 4.2: Quantitative performance evaluation of various architectures

Method	Avg. Jaccard Index (Std dev)	Avg. Dice Score (Std dev)	Avg. F1 Score (Std dev)	Avg. Accuracy (Std dev)
U-Net	0.7015 (±0.035)	0.8133 (±0.039)	0.8327 (±0.0296)	0.8919 (±0.0271)
Attention U-Net	0.7189 (±0.033)	0.8297 (±0.037)	0.8401 (±0.0296)	0.9195 (±0.0269)
MultiResUNet	0.7264 (±0.032)	0.8458 (±0.033)	0.8658 (±0.0289)	0.9398 (±0.0258)
Proposed LDS U-Net	**0.7415 (±0.03)**	**0.8694 (±0.0285)**	**0.8885 (±0.0256)**	**0.9592 (±0.020)**

In Table 4.2, the segmentation output of the proposed LDS U-Net is compared with other models using four segmentation evaluation indices, i.e., Jaccard index, Dice score, F1 score, and pixel accuracy. Dice Score is the most direct evaluation index. The average dice score of LDS U-Net is 0.8964, which is 2.79%, 4.57% and 6.89%, higher than MultiResUNet, Attention U-Net, and U-Net, respectively. In terms of the Jaccard index, F1-score, and pixel accuracy, the results of LDS U-Net are much better than the other mentioned networks. The standard deviations of the results of the proposed LDS U-Net also show a smaller spread. The proposed LDS U-Net is more capable of learning almost all the features from the training dataset.

4.4.2.2 LDS U-net gives the best identification of bony features

The detection of exact locations of lateral bony features (thoracic and lumbar) is very crucial for UCA calculation. However, the associated speckle noise makes the identification process more challenging because it suppresses many important features. Four sets of images are shown in Fig. 4.10. Each set consists of an input image, a truth mask, and the segmentation results using U-Net, Attention U-Net and MultiResUNet and the proposed LDS U-Net model.

From the set shown in Fig. 4.10a, it is clear that the right side upper TBFs of the spine image are missing in the U-Net segmentation. For Attention U-Net, these are just visible, and for MultiResUNet, and LDS U-Net, they are more prominent. In the case of LBFs, identification is not so straightforward. The T12 region and six LBFs are visible in the truth mask. U-Net can identify the T12 region and five LBFs. The Attention U-Net and MultiResUNet can identify only four of the LBFs. The proposed LDS U-Net can identify all LBFs and the T12 region, which indicates that this model is capable in segmenting LBFs correctly.

From Fig 4.10b, it is clear that U-Net is not capable of segmenting all the bony features in the six different segments of LBFs. Attention U-Net and MultiResUNet are somehow able to distinguish the six features, but some features are unclear and give misleading information. On the other hand, LDS U-Net gives clearer identification of the T12 region and six LBFs, although the last two features are still conjoined.

Similar observations about TBFs and LBFs can be made in Fig. 4.10c. In U-Net, the two uppermost LBFs are not segmented correctly but appear to be fragmented, while in Attention U-Net and MultiResUNet, they appear to be elongated. LDS U-Net clearly distinguishes all the thoracic and lumbar bony features.

Fig. 4.10d shows that U-Net and MultiResUNet are unable to segment the T12 region and all LBFs from the ultrasound spine image. On the other hand, Attention U-Net can identify all the six LBFs. However, in both cases, the T12 region is conjoined with the uppermost lumber bony feature. The proposed LDS U-Net can segment the T12 region and all the lumbar bony features consistently.

4.4.2.3 Evaluation of the number of bony features identified

Identification of distinct bony features is a key criterion for evaluating the performance of various segmentation models. In a truth mask of an ultrasound spine, depending on the scanning coverage area, there are 9 pairs of TBFs, a T12 region (one pair of bones), and 6 LBFs, making a total of 26 bony features. The total number of distinguishable bony features in the outputs of each model is manually recorded. In many cases, the bony features are either conjoined or fragmented and they are not recorded as meaningful features. Table 4.3 shows that LDS U-Net provides a larger number of meaningful bony features when compared to other models. The proposed LDS U-Net can give the highest percentage of segmented images with all 26 features fully detected (76.59%).

4.4.2.4. Evaluation of computational requirement

In Table 4.4, the computational requirement of the contemporary models and the

Fig 4.10. Four Spine images sets (a, b, c, d) consisting of the input image, truth mask, and segmentation results using U-Net, Attention U-Net, MultiResUNet, and the proposed LDS U-Net model

proposed model are compared in terms of a total number of parameters and program storage requirements. LDS U-Net architecture offers advantages such as (i) low computational requirements due to the usage of depthwise separable convolution layers, (ii) feature reuse enabled through the application of novel Light Dense block

Table 4.3: Comparison of number of bony features identified (avg.)

TRUTH MASK	U-NET	ATTENTION U-NET	MULTIRESUNET	LDS U-NET
9 pairs of TBF, T12 region, and 6 LBFS – Total: 26 bony features	25.11	25.23	25.30	**25.45**
% of Images where all 26 features were fully detected	69.72 %	71.56 %	73.23 %	**76.59 %**

Table 4.4: Evaluation of computational requirement

Method	Total Number of Parameters	SIZE IN MEGABYTES (MB)
U-Net	31 Million	355 MB
Attention U-Net	37.1 Million	433.49 MB
MultiResUNet	7.2 Million	146.01 MB
Proposed LDS U-Net	**9.1 Million**	**127.5 MB**

, and (iii) enhancement of the depth and width of the network, with no incremental computational budget, through the redesigned multi-scale inception layer. As shown in Table 4.4, these advantages of the proposed model translate to a lower parameter requirement and smaller memory footprint when compared to U-Net and Attention U-Net. Although MultiResUNet requires a bit lesser number of parameters than LDS U-Net, the overall segmentation performance of MultiResUNet is lower as discussed in the previous sections.

4.5 Discussion on LDS U-Net

Being a radiation-free and economic imaging modality, the 3D ultrasound imaging system has the potential to become a very popular diagnosis technique to detect scoliosis (Zhou & Zheng 2015). However, this necessitated the establishment of new measurement indices as, in the X-ray and Cobb Angle method, scanning is done to assess anterior spinal deformity. A new index called Spinous Process Angle (SPA) was developed exclusively for ultrasound scanning to measure the posterior spinal deformity. In a nutshell, if lines are drawn through the most tilted parts of the spinous column profile of coronal ultrasound images, then the angle that is formed is known as SPA. SPA measurement focuses on the middle dark line of spine ultrasound image as the main region of interest.

Subsequently, several scoliosis assessment techniques, both manual and automatic, were researched to measure SPA and to demonstrate its equivalence to the gold-standard Cobb angle (Cheung, Zhou, Law, Mak, et al. 2015; Zhou et al. 2017). But there was an overall drawback in the SPA measurement approach that there was a high inherent tendency of the technique to underestimate the severity of curvature of scoliosis compared to Cobb angle (Zheng, Lee, et al. 2016; Zhou & Zheng 2015). Subsequently, an alternate index was developed called Ultrasound Curvature Angle (UCA). Unlike SPA, in UCA technique, the lateral bony features are the main regions of interest and further research was undertaken to demonstrate that it is equivalent to X-ray Cobb angle (Lee et al. 2021). Though manual method of UCA measurement is established (Lee et al. 2021), there is a need to find out a suitable method to automate the UCA measurement so that the technique can be made fast and scalable.

Clear demarcation of lateral bony features is a vital step in UCA measurement. In traditional X-Ray, the images are relatively noise-free and clear features can be identified easily. However, in case of ultrasound modality, scanned images have low contrast and are often plagued by speckle noise. This makes the differentiation of the bony features challenging. Also the fact, that the lateral bony features are numerous and can significantly vary in shape, size and location, adds to the overall complexity of the problem. The first step of automatic UCA measurement is clear segmentation of the lateral bony features and this chapter attempts to tackle this problem using deep learning.

U-Net is one of the most successful architectures for biomedical image processing. As a starting point, the ultrasound spine images were segmented using basic U-Net (Ronneberger, Fischer & Brox 2015). But, the segmentation output was not satisfactory (Avg. Dice Score: 0.8133) as the output images had a lot of missed or conjoint bony features and the architecture required a large number of parameters which directly increased the computation cost and memory size (Gadosey et al. 2020). As the first modification, the conventional convolution layers of U-Net were replaced with proposed light dense block and depthwise convolutional operations replaced the conventional convolution operations. The new architecture was called Light Dense (LD) model. In LD model, each depthwise convolution unit is densely connected with another depthwise convolution unit within the same light dense block (shown in Fig 6). This architecture incorporates feature propagation and boost feature reuse while effectively reducing the number of parameters. Though the LD model gave better segmentation accuracy (Avg. Dice Score: 0.8204) with less number of parameters, it was still unable to adequately distinguish the TBFs and LBFs and generated conjoint bony features in the output as the input images were noisy. The two issues, high noise in images and variability of bony features, were tackled separately as follows:

a) To improve the segmentation clarity and tackle the 'noisy' information, trainable selection gates were employed in the LD model as the next modification and the new architecture was called Light Convolution Dense Selection (LCS) model. By this modification, it was anticipated that the gating mechanism (Oktay et al. 2018) would smartly suppress irrelevant information, such as noise, from the feature maps and be able to identify more number of bony features from the useful information in the image. Past research shows that the gating mechanism can be effectively used to extract the selective features from the noisy ultrasound images as it suppresses noise and enables the network to make segmentation predictions based on class-specific features (Schlemper et al. 2019). In effect, the selection gate blocked extraneous information from information-rich feature maps and thereby improved the segmentation output for noisy ultrasound images (Avg. Dice Score: 0.8292).

b) To handle the problem of large variabilities of shape, size, and locations of the TBFs and LBFs, the conventional skip-pathways were replaced by multi-scale (Ibtehaz & Rahman 2020) skip-paths and the architecture was named as Light-Dense Inception (LDI) model. Feature fusion was enhanced by using this skip-path as the inconsistencies between encoder-decoder features were bridged. This path solves the problem of the semantic gap between the encoder and decoder side and enhances the feature fusion and improves the segmentation output (Ibtehaz & Rahman 2020). By this modification, the low-level features underwent additional processing and the multi-scale inception layers, adapted in this skip-path, enabled the network to extract more features from different scales. This gave a significant boost to the segmentation performance (Avg. Dice Score: 0.8419) compared to the initial LD model.

As a final step, the two sub-models were combined into the final architecture and is called Light-Convolution Dense Selection U-Net (LDS U-Net) model. As anticipated, the combination of selection gates and multi-scale skip pathway gave a much improved segmentation accuracy (Avg. Dice Score: 0.8694) compared to both LD model and basic U-Net model.

4.6 Key takeaways from LDS U-Net

In a nutshell, the proposed LDS U-Net has 3 main elements:

a. Light Dense blocks that reduce the number of parameters used in the architecture and thereby reduce the computation time,
b. Selection gates that smartly discards the extraneous regions and ensure explicit flow of only the relevant information within the network and,
c. Multi-scale paths that help with clearer identification of lateral bony features through improved feature propagation.

To understand performance of LDS U-Net vis-a-vie other contemporary models, the proposed model is compared with the basic U-Net, MultiResUNet, and Attention U-Net using the same 109 volume projection ultrasound spine image datasets. Firstly, LDS U-Net quantitatively outperforms the other models with 0.7415 Jaccard index, 0.8694 Dice coefficient, 0.8885 F1 score, and 0.9592 pixel accuracy. Secondly, it is able to better identify the thoracic and lumbar bony features most consistently when qualitatively compared to other models. The segmentation outcomes of LDS U-Net, through manual observation and measurement of distinguishable bony features, are also very promising. Thirdly, as it is constructed using light dense block, LDS U-Net is also found to be computationally efficient with least number of parameters and smallest program memory size. Finally, in handling noisy ultrasound images with high variability, the LDS U-Net is successful in correctly segmenting 76.59% of total images (shown in Table 4.3) with complete identification of all the 26 bony features, which is more consistent than any of the other models.

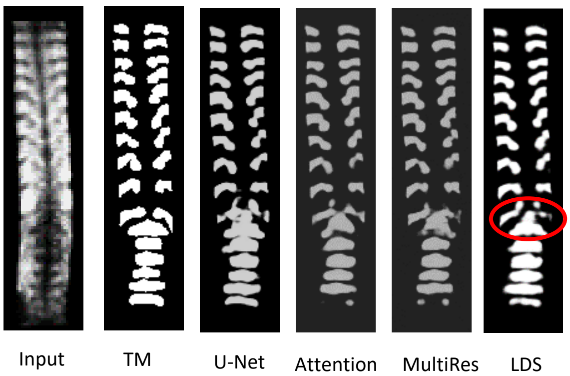


Fig 4.11. LDS U-Net is unable to distinguish T12 level and 1st LBF

4.7 Areas of improvement of LDS U-Net

The objective of segmentation is to clearly identify the bony features through an adequate segmentation technique. This would form the input for further automatic angle calculation. Upon scrutiny of this research, it is observed that there are few cases where the LDS U-Net did not perform well. For instance, in Fig 4.12, the demarcation between T12 level and 1st LBF is obscure.

Table 4.5: LDS U-Net Detection Rate

Index	LDS U-NET
% TBF Accuracy	79.7%
% LBF Accuracy	72.3%
Average % BF Accuracy	76.6%

Further, the detection rate of LDS U-Net is summarized in table 4.5. The bony features of about 23% of images are not detected clearly. It is postulated that there may be two main causes of this sub-optimal performance:

(a) Propagation of noise through the network: The information at the encoder side is more contextual and that at the decoder side is more localized. In LDS U-Net, the skip connections originate at the encoder side and, despite the selection gates, can result in speckle noises being transferred to the decoder.

(b) Filter sizes are fixed for a given level: During the convolution operation, LDS U-Net is not designed to choose the appropriate kernel size to handle the large variation in the locations, shapes, and sizes of the features. The variability of bony features is only handled through the multi-scale skip pathway. This resulted in loss of information of TBFs for about 20% of images and of LBFs for about 23% of images, as shown in the table 4.5.

As an improvement and for a clearer demarcation of bony features, a new approach is taken to better handle the noise and variabilities in bony features by making the architecture denser and adaptable. This forms the basis of the second approach to designing a novel architecture (SIU-Net) described in Chapter V.

CHAPTER V: ULTRASOUND SPINE IMAGE SEGMENTATION USING MULTI-SCALE FEATURE FUSION SKIP-INCEPTION U-NET (SIU-NET)

5.1 Introduction

As highlighted in Chapter 1, Figure 1.4, segmentation forms the basis of identifying the key feature points for further automatic angle measurement. After the evaluation of LDS U-Net's performance (described in Chapter IV), it is thought that the overall segmentation performance could be improved by focusing on the drawbacks through an alternate architecture design.

5.1.1 Re-visiting the capabilities of contemporary CNN architectures to handle variations

Usage of Convolutional Neural Network (CNN) for the analysis of medical images is a recent trend (Zhang et al. 2017) and has become the most used deep learning technique in this field (Krizhenvshky, Sutskever & Hinton), (Thong et al. 2015), (Liu et al. 2019), (Hong et al. 2018). Because CNN has an end-to-end characteristic, it does not require a manual design of features and has been proven excellent in image feature extraction (Zeiler & Fergus 2014). Kokabu et al. applied a basic CNN architecture to assess the performance of a 3D depth sensor imaging system in predicting the Cobb angle (Kokabu et al. 2021). However, due to the inherent process of depth scanning, the research faced a limitation in predicting the Cobb Angle with high accuracy. Also, no external validation dataset was considered, and the CNN architecture applied was computationally very expensive Also, in general, CNN's usefulness and accuracy are hampered by their need for large quantities of training data. In the case of medical images, getting hold of an adequate image can often be expensive, complicated and the prerequisite of accurate annotations adds to the complexity (Litjens et al. 2017).

In biomedical image segmentation, enormous success was achieved by using the U-Net architecture (Ronneberger, Fischer & Brox 2015). A U-Net is made up of two main sections: (a) Multilayer deep encoder network which helps to extract spatial features from an image, and (b) corresponding multilayer deep decoder network that up-samples the extracted feature maps to predict the final segmentation output. It utilizes the self-learning property of the convolution kernel to process the original image and delivers the classification result. Since its architecture is modular in construction, the U-Net can extract considerably complex and detailed image features just by increasing the depth or the number of layers in the architecture. Usually, the lower layers of U-Net are capable of extracting some common features of images, whereas the higher layers extract more targeted features (Liu, Li & Gong 2019). In a basic CNN model, the spatial information may get lost during the max-pooling and transposed convolution operation. To reduce this loss of information, U-Net employs skip pathways that connect the encoder to the corresponding decoder in the same layer.

U-Net can achieve very good results with a small number of training datasets, which is very useful in the case of medical image segmentation, where a large number of training data is mostly unavailable. For instance, a two-stage U-Net was used for the segmentation and detection of breast lesions with various shapes and locations from ultrasound images with artifacts with high accuracy (Amiri et al. 2020). Also, a modified version of U-Net, Oct-U-Net was used successfully to diagnose fetal spina bifida from a 3D ultrasound image and gave better segmentation accuracy than Fully Convolutional Network (Chen, Tian & Deng 2021). Research has also been carried out on the application of U-Net in scoliosis detection such as the basic U-Net was employed to automatically segment the bony features from 2D ultrasound images for scoliosis measurement. However, as the segmentation technique was applied on sparse 2D images, the predicted segmentation accuracy was low (Ungi et al. 2020).

An advanced version of U-Net, U-net with robustness to speckle and regular occlusion noise (RSN-U-net) was introduced to noise removal of spine ultrasound VPI images and segment the bony features. As the variability of shape and sizes of bony features was not addressed in that research, the segmentation accuracy was not high (Huang et al. 2020). Recently, dual-task ultrasound transverse vertebrae segmentation network (D-TVNet), another modified U-Net along with Atrous spatial pyramid pooling (ASPP) module was introduced to segment the ultrasound spine image (Lyu et al. 2021). The research aimed to clearly distinguish boundary edges of the bony features from noisy ultrasound scan images. Though it produces an overall promising segmentation output, the architecture was insufficient to segment the bony features properly when a larger area was occupied by noise.

However, for the convolution operation, U-Net is not designed to choose the appropriate kernel size to handle the large variation in the locations, shapes, and sizes of the features (Ibtehaz & Rahman 2020). Adding more layers to a conventional U-Net is not a good option as the network would become deeper and would produce redundant computation during training (Bi et al. 2017). Moreover, the skip-connections of U-Net impose a limiting fusion scheme, forcing combinations only at the same scale feature maps of the corresponding encoders and decoders (Zhou et al. 2018), (Zhou et al. 2019).

A sequence of fixed multi-scale Gabor filters was used in biomedical image processing to handle the variations in an input image. In CNN, the concept of Inception architecture was introduced (Szegedy et al. 2015) to perform the same operation. Also, since Inception blocks utilize convolutional layers of varying kernel sizes in parallel, they can be used to assess the region of interest from diverse scales (Szegedy et al. 2015), (Szegedy, Vanhoucke, et al.), (Szegedy et al. 2017), and to combine and carry the outputs deeper into the network.

5.1.2 Approach to re-designing a novel architecture to better handle variabilities and noise

During this research, it is found that for dealing with the complexity of an ultrasound image, same-scale feature fusion skip pathways alone will not suffice. As an antidote to U-Net's restrictive same scale feature fusion problem, UNet++ aimed to improve segmentation accuracy by including (a) dense blocks (Huang et al. 2017) and (b) convolution layers between the encoder and decoder (Zhou et al. 2018), (Zhou et al. 2019). Dense skip connections ensure that all prior feature maps are accumulated and transferred to the succeeding node along each skip pathway. This generates full resolution feature maps at multiple semantic levels and also the re-designed skip connections aggregates features of different semantic scales which produced highly flexible feature fusion schemes. However, the information at the encoder side is more contextual and that at the decoder side is more localized. In U-Net++, dense skip connections originate at the encoder side and can result in speckle noises getting transferred to the decoder during the concatenation operations. It may give a suboptimal performance while handling ultrasound images.

There are two major problems to be addressed in this research:

a) Choosing the appropriate kernel size for the convolution operation which can handle the large variation in locations, shapes, and sizes of the features.
b) Designing a suitable architecture that can extract semantically rich features and fuse multi-scale features for a better segmentation output.

In this research, a novel deep learning architecture, Skip-Inception U-Net or SIU-Net, is developed to suitably segment the TBFs and LBFs in the ultrasound spine image dataset. The standard U-Net is adopted as the main network architecture, and the simple convolutional layers are replaced with Inception blocks (Szegedy, Vanhoucke, et al. 2016). The encoders-decoders are bridged using newly designed decoder side skip-pathways (Zhou et al. 2019). The segmentation result from the proposed architecture is compared with the basic U-Net, UNet++, and MultiResUNet (Ibtehaz & Rahman 2020) models. The result shows that the new model outperforms the aforementioned models.

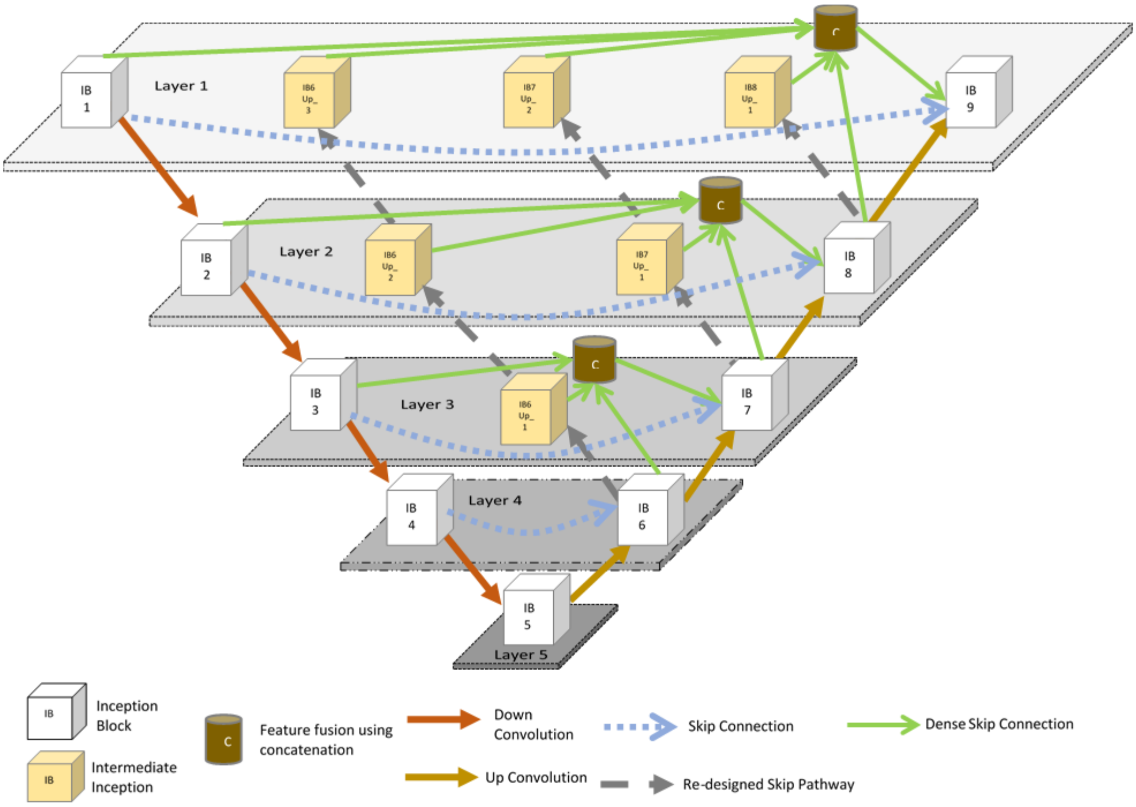


Fig 5.1. Overall architecture of proposed SIU-Net

5.2 Methodology of SIU-Net

The architecture of the proposed Skip-Inception U-Net or SIU-Net is shown in Fig 5.1. This network model employs the basic U-Net structure as the base framework and contains the improvised Inception blocks, the re-designed Dense-skip connection feature fusion using concatenation and the Down-sample path, and the Up-sample path.

5.2.1 Inception block

To solve the issue of choosing an appropriate kernel size to handle large variability in the spine image dataset, the concept of Inception block (IB) is adopted to develop a high-performance segmentation model (Szegedy et al. 2015). A modified IB (Fig. 5.2) of

Inception V2 architecture (Szegedy, Vanhoucke, et al. 2016) is introduced to replace the traditional convolutional layers of basic U-Net. There are two advantages of IB: (a) it increases the depth and width of the model without any increase of computational requirement and (b) it allows the flexibility of using multiple filter sizes within the same level (Zhang et al. 2020).

In the main U-Net architecture, a sequence of two 3×3 convolutional layers was used after each pooling layer and transposed convolutional layer. On the other hand, a basic Inception block involved the use of a 5×5 filter, which required high computation. According to (Szegedy, Vanhoucke, et al. 2016), factorizing one 5×5 convolution to two 3×3 convolution operations can improve the computational speed since one 5×5 convolution is 2.78 times more time consuming than one 3×3 convolution.

In SIU-Net, the selection of the filter sizes of the modified IB is optimized to 3×3 and

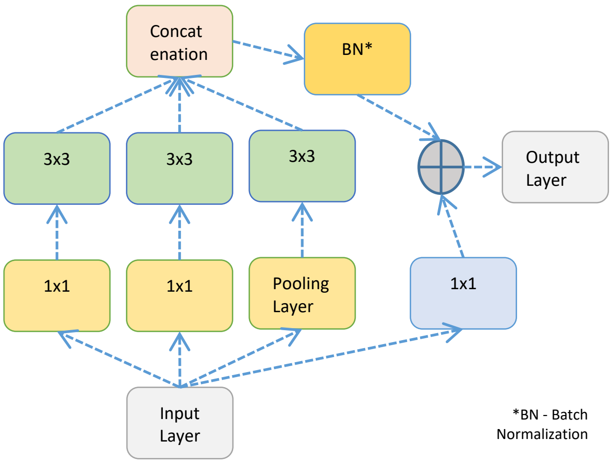


Fig 5.2. Inception Block in proposed architecture

1×1 to make it computationally efficient. The modification enables better aggregation of feature maps from different branches of kernels of different sizes and makes the network wider and capable of learning more features (Badrinarayanan, Kendall & Cipolla 2017). Also, the residual connection makes the learning easier since a residual IB learns a function with reference to the input feature maps, instead of learning an unreferenced function (Szegedy et al. 2017). Unlike the original Inception V2 architecture, each 1×1 convolutional layer is followed by a 3×3 convolution layer. The concatenated 3×3 layers are then followed by a batch normalization (BN) layer that avoids gradient vanishing while retaining the convolutional layers.

In the modified Inception block, the technique is adapted from (Ibtehaz & Rahman 2020) to control the number of filters of the convolution layers inside the block. A parameter P is assigned to control the number of filters as given by equation -

$$P = \lambda \times F \tag{5.1}$$

where, F is the number of filters in the corresponding layers similar to the basic U-Net, and λ is a scaler coefficient. Putrefying P to λ and F gives a suitable way to control the number of parameters as well as to keep them equivalent to basic U-Net. The number of filters is set to F = [32, 64, 128, 256, 512] along with the layers respectively as these numbers are comparable to basic U-Net. λ is chosen as 1.67 to ensure that the model structure is similar to the basic U-Net. Instead of retaining the filters number same, it is useful to progressively increase the them in the consecutive convolution layers inside a MultiRes block (Ibtehaz & Rahman 2020). The number of filters is set to $(\frac{P}{6}, \frac{P}{3}, \frac{P}{2})$ in the three corresponding convolutional layers for extracting multiscale features as this combination achieved the best results in previous experiments (Ibtehaz & Rahman 2020).

Layer	*m*	*Inception Blocks*	*n*
1	0	(0,0) (0,1) (0,2) (0,3) (0,4)	3
2	1	(1,0) (1,1) (1,2) (1,3)	2
3	2	(2,0) (2,1) (2,2)	1
4	3	(3,0) (3,1)	0
5	4	(4,0)	

Fig 5.3. Feature fusion scheme using the dense convolution operation

5.2.2 Dense-skip connection

A multi-scale feature fusion scheme impacts segmentation accuracy more than a single-scale (Zhou et al. 2019). In SIU-Net, the output of the previous IB of the decoder of the same dense block is merged with the corresponding up-sampled output of the lower dense block through dense skip connections (DSC) (Fig. 5.1). Through a dense convolution operation, each node in a decoder is presented with a final aggregated fused feature map containing: a) the feature from the previous decoder, b) intermediate block combined feature maps and c) the same-scale feature from the corresponding encoder.

Let,$(f^{m,n})$ denote the output of node $F^{m,n}$ where m is the down-sampling layer along the encoder side and n is the convolution layer of the dense block along with the skip connection. The whole feature fusion scheme is shown in Fig 5.3. The fused feature maps $f^{m,n}$ is calculated as:

$$f^{m,n} = \begin{cases} \dot{\mathrm{E}}\big((f^{m,n}), \mathrm{U}(f^{m+1,n-1})\big), & n = 0 \\ \dot{\mathrm{E}}(C\big[(f^{m,n}), U[f^{m+1,p}]_{p=0}^{n-1}, (f^{m+1,n-1})\big]) & n > 0 \end{cases} \tag{5.2}$$

where $\dot{E}(.)$ denotes a convolution operation followed by the activation function, U (.) denotes an up-sampling layer and C[ ] denotes the concatenation operation. Node at level n=0 receives only one input from the previous layer of the encoder, the nodes at level n=1 get features from the same level encoder, one encoder sub-network, and the previous layer decoder. Nodes at level n>1 get n+2 features, where the two features are from the same layer encoder and previous layer decoder respectively and the rest of the n features are from n encoder sub-networks of the same skip-connection. Encoder sub-networks are the up-sampled output from the lower level skip-connection.

Therefore, the top layer of the proposed network (Fig. 5.1) is:

$$IB9 = \dot{E}(C[(IB1, IB6up_3, IB7up_2, IB8up_1), IB8] \qquad (5.3)$$

where $\dot{E}(.)$ indicates a convolution operation followed by the activation function, C[ ] indicates the concatenation operation. Other layers follow similarly. $IB1, IB8, IB9$ indicates 1st , 8th and 9th inception blocks shown at Fig. 5.1. $IB6up_3$ indicates the 6th inception block after three up convolution operations. Similarly, $IB7up_2$ and $IB8up_1$ stands for 7th inception block after two up convolution operations and 8th inception block after one up convolution operation respectively.

By this, the advantage offered by the skip pathway is retained. Since the shallower layers of the decoder side are used to extract the information, more localized information is extracted by SIU-Net. Also, these DSCs generate full resolution feature maps at multiple semantic levels and help to improve segmentation accuracy and gradient flow.

5.2.3 Ablation study

During the initial stages of research to find the best segmentation method for segmenting the TBFs and LBFs in the ultrasound spine images, several models were designed and continuously improved from the basic U-Net structure. Notably, three models are highlighted below.

In the first model, the improvised IB (Fig. 5.2) is employed within the basic U-Net architecture while the down-convolution, up-convolution, and skip pathways are unchanged. This network is named as IU model (Inception+ U-Net). The output is not satisfactory as the issue of the restrictive single-scale feature aggregation is not addressed.

In the second model, the idea of the Residual path (Ibtehaz & Rahman 2020) of MultiResUNet is deployed by incorporating some convolution layers with residual connections along the traditional skip path. This is done to add some additional non-linear transformations on feature propagation from the corresponding encoder to the decoder. This new model is similar to the normal MultiResUNet except that the conventional IB is replaced with the modified Inception block. This network is named as IRs model (Inception + Res path). However, the performance of this model is found unsatisfactory because the residual paths do not extract features from deep layers and are inadequate to handle noisy ultrasound images.

In the third model, along with the improvised IB, the encoder side skip-pathways (similar to UNet++) are introduced. It is named as ISP model (Inception + encoder skip path). Though this model gives better output than the previous two models, it is found that the absence of the skip pathways in the decoder side results in a lack of richness of information especially during edge detection of bony features amidst image noises.

5.3 Experimental setup

5.3.1 Dataset

The dataset used for this approach is same as the LDS U-Net approach and mentioned in section 4.3.1 in Chapter IV. As in LDS U-Net approach, each input ultrasound image, of size 2574 × 640 pixels, is resized to 256 × 64 pixels through image pre-processing and maintaining the aspect ratio of the original image. However, this time, data augmentation is done to increase the size of the training dataset by randomly flipping and rotating the available images.

5.3.2 Implementation details

Software: Spyder, Anaconda.

Libraries: Keras with Tensorflow backend (Abadi et al. 2016).

Machine: GPU laptop having NVIDIA GeForce RTX 2060.

The mode of operation of any semantic segmentation algorithm is to analyze each pixel and to predict whether they represent a point of interest, or are merely a part of the background. This means that the task can also be looked at as a pixel-wise binary classification problem where the objective of the segmentation algorithm would be to minimize the binary cross-entropy loss function.

For an image X, let the corresponding Truth Mask (TM) be Y, and the predicted segmentation output be Y'. For a pixel px, the TM value is y_{px} and the network predicted output is y'_{px}, The binary cross-entropy loss for that image is defined as:

$$Cross\ Entropy\ (X, Y, Y') = \sum_{px\epsilon X}[\![-(y_{px}]\!] \log(y'_{px}) + (1 - y_{px}) \log(1 - y'_{px})) \quad (5.4)$$

For a batch containing n images, the loss function J becomes,

$$J = \frac{1}{n}\sum_{i=1}^{n} Cross\ Entropy\ (X_i, Y_i, Y'_i) \quad (5.5)$$

The binary cross-entropy loss is minimized, and the model is trained using Adam optimizer (Kingma & Ba 2014). Adam optimizer adaptively computes different learning rates for different parameters from estimates of the first and second moments of the gradients. All the models are trained up to 150 epochs since 150 epochs are found to be the saturation point for model accuracy, and no further improvement is observed beyond this point. Finally, for validating the reliability of the model, 5-fold cross-validation of the dataset is done.

5.3.3 Evaluation metrics

For quantitative performance evaluation, three very popular evaluation indices are employed – Jaccard similarity (JS) (Chapter IV, equation 4.7), Dice Coefficient (DC) (Chapter IV, equation 4.8), and Euclidean distance (d) is given by,

$$Histogram\ Euclidean\ distance, d = \sqrt{\sum_i (hist(i1) - hist(i2))^2} \quad (5.6)$$

$hist(i2)$ is the histogram of the predicted image and $hist(i1)$ is the histogram of the corresponding truth mask.

5.4 Analysis of performance of SIU-Net

A thorough analysis of the proposed model, SIU-Net, is undertaken and compared with basic U-Net, UNet++, and MultiResUNet. Also, the performances of the intermediary models, i.e. IU model, IRs model, and ISP model, are analyzed. All the architectures have the same settings as illustrated in section 5.3.2.

Table 5.1: Quantitative evaluation

Method	Avg. Jaccard Index (Std. Dev.)	Avg. Dice Score (Std. Dev.)
U-Net	0.709 (0.034)	0.817 (0.037)
MultiResUNet	0.732 (0.034)	0.853 (0.029)
UNet++	0.748 (0.035)	0.861 (0.034)
Incep. + U-net	0.711 (0.035)	0.846 (0.033)
Incep. + Res path	0.734 (0.035)	0.855 (0.029)
Incep. + Encoder Skip Path	0.757 (0.036)	0.876 (0.026)
Proposed SIU-Net	**0.781 (0.033)**	**0.883 (0.025)**

5.4.1 SIU-Net outperforms U-Net, UNet++, and MultiResUNet in the segmentation of ultrasound spine image dataset

Table 5.1 details the segmentation performance (JS and DC) of the SIU-Net, UNet++, MultiResUNet, U-Net, and other three intermediary models i.e. IU, IRs, and ISP. It can be observed that each intermediary model (i.e. IU, IRs, and ISP) has a significant improvement on segmentation performance over their corresponding basic methods. This

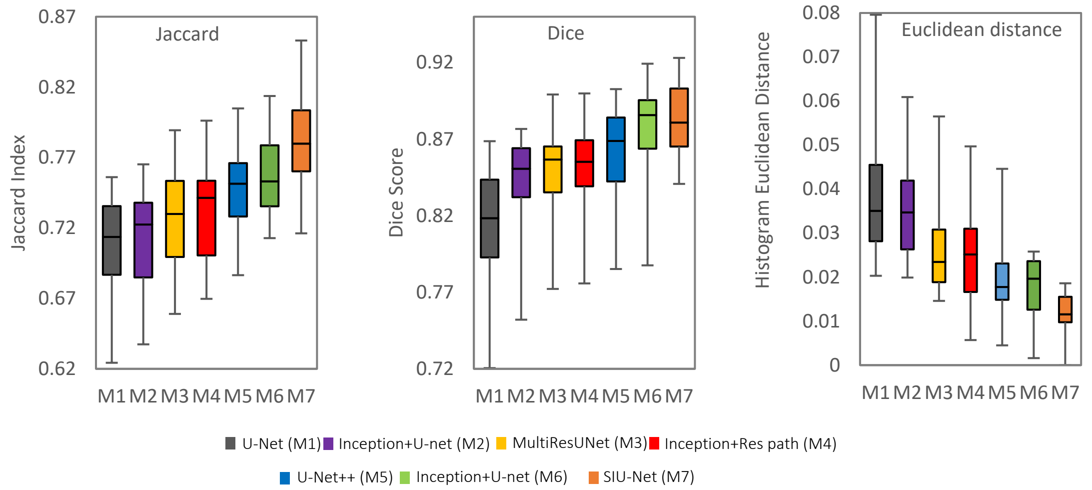


Fig 5.4. Performance comparison of models using (a) Jaccard Index, (b) Dice Index and (c) Histogram Euclidean distance

is because an IB, by its nature of construction, is more adept in extracting features from different locations, shapes, and sizes than conventional convolution layers. Fig 5.4 illustrates the performances of all the models in terms of the Jaccard index, Dice index, and Histogram Euclidean distance. The evaluation result shows that SIU-Net is deeper and more capable of learning features from datasets and achieves a better performance in terms of TBF and LBF segmentation.

5.4.2 SIU-Net performs best for bony feature edge detection

The main aim of this research is to appropriately locate LBFs and TBFs from the ultrasound spine images with a high degree of clarity for accurate scoliosis detection. This requires the segmentation of each bony feature with a proper edge boundary and location. Because of speckle noise, ultrasound images suffer from the lack of clear boundaries which complicates the segmentation process. Though U-Net and MultiResUNet can segment the bony features from the input images, UNet++ and SIU-Net perform better to clearly distinguish the edge boundaries (Fig. 5.5a). Further detailing suggests that UNet++ performs better than MultiResUNet, but it cannot outperform SIU-Net. The overall performance of a particular image is explained in Fig. 5.5a. The 8^{th}, 9^{th} and 10^{th}

TBFs pair's edge boundaries are not distinguishable using U-Net (JS: 0.7198, DC: 0.8201) and MultiResUNet (JS: 0.7241, DC: 0.8367), but it is prominent using UNet++ (JS: 0.7322, DC: 0.8484) and SIU-Net (JS: 0.7415, DC: 0.8503).

For more challenging images (Fig. 5.5b), the 10th right-side TBF and 1st LBF are

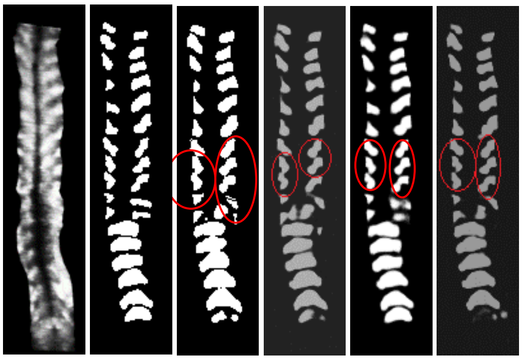


Fig 5.5a. SIU-Net is able to distinguish TBF pairs

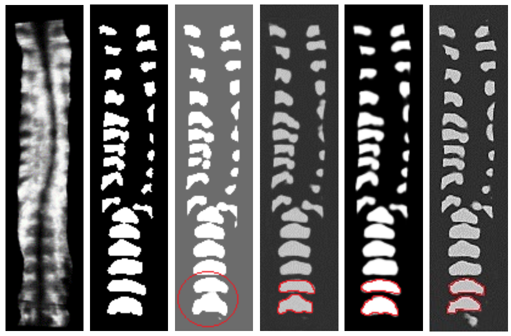


Fig 5.6a. SIU-Net is able to segment LBF

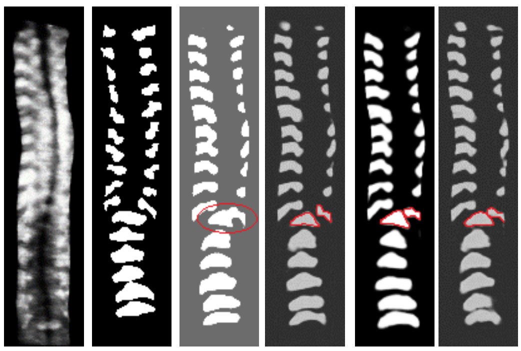


Fig 5.5b. SIU-Net is able to distinguish TBF and LBF

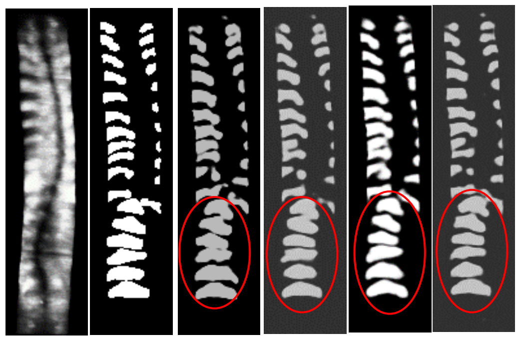


Fig 5.6b. SIU-Net is able to identify six individual LBFs

indistinguishable in U-Net (JS: 0.7313, DC: 0.8393) but clearly distinguishable in MultiResUNet (JS: 0.751, DC: 0.8578), UNet++ (JS: 0.7649, DC: 0.8601) and SIU-Net (JS: 0.7723, DC: 0.8679).

5.4.3 SIU-Net performs the best identification of LBF

There are six lumbar bony features visible in an ultrasound spine image. Though basic U-Net and MultiResUNet give good Jaccard and Dice value, in some images, the LBFs are not segmented. In Fig. 5.6a, the two bottom LBFs are indistinguishable in U-Net segmentation (JS: 0.7677, DC: 0.8542), but they are segmented in MultiResUNet (JS: 0.7801, DC: 0.8708), UNet++ (JS: 0.7953, DC: 0.8839) and SIU-Net (JS: 0.8001, DC: 0.8973). With closer observation, it can be seen that in SIU-Net, the edge boundary is better than the other two.

Again from Fig. 5.6b, it is clear that, in case of U-Net (JS: 0.7074, DC: 0.8158) and MultiResUNet (JS: 0.7179, DC: 0.82), many LBFs appear to be conjoined, whereas outputs of UNet++ (JS: 0.7265, DC: 0.834) and SIU-Net (JS: 0.7319, DC: 0.8418) are able to distinguish all the 6 LBFs properly.

5.4.4 SIU-Net is most capable of identifying TBF

Speckle noise poses considerable challenges, especially during identifying TBFs, which are more in number and smaller in size than LBFs. In Fig. 5.7a, as shown in the marked red area, the right TBFs are quite vague in the input image. While the other two methods (U-Net (JS: 0.6633, DC: 0.7976), UNet++ (JS: 0.6598, DC: 0.7936)) are able to

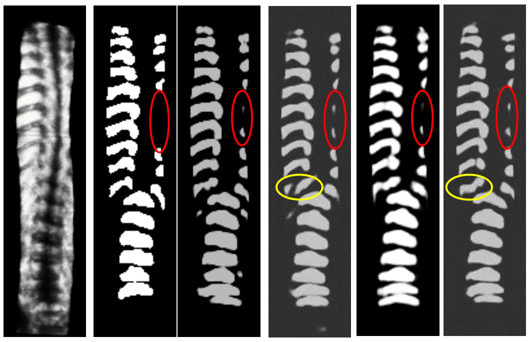


Fig 5.7a. SIU-Net is able to identify TBFs that are invisible in TM

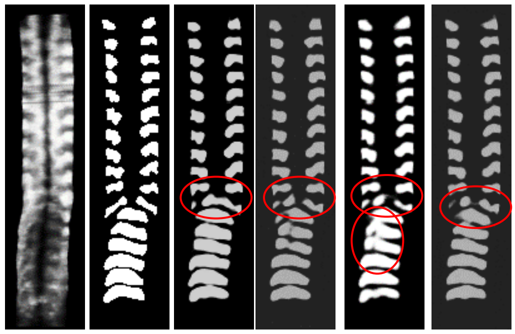


Fig 5.8a. SIU-Net outperforms MultiResUNet in middle bony feature segmentation

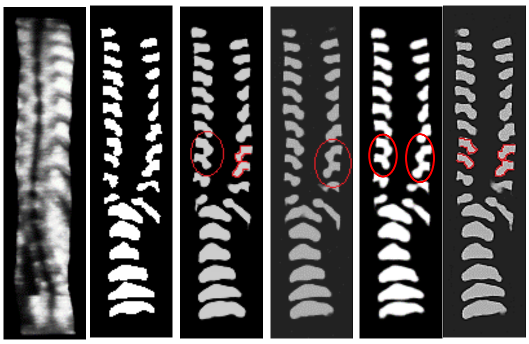


Fig 5.7b. SIU-Net is able to detect the edges of TBFs that are indistinguishable in other two methods

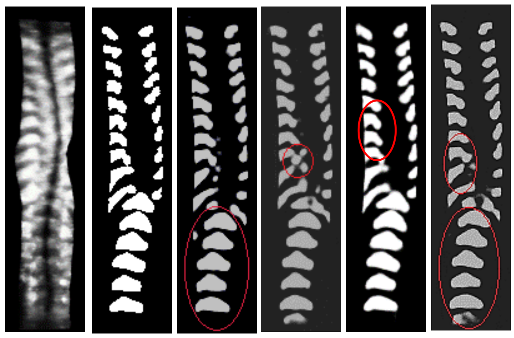


Fig 5.8.b. SIU-Net is able to distinguish both TBFs and LBFs clearly

identify one or two of those TBFs, MultiResUNet (JS: 0.6583, DC: 0.792) and SIU-Net (JS: 0.6494, DC: 0.7874) are better in differentiating all of them. Overall, SIU-Net performs better than MultiResUNet as the T12 level is fragmented in the case of the latter (shown in the yellow marked area). In fact, SIU-Net performs better than the manual segmentation process as those three TBFs are missing even in the truth mask.

In Fig. 5.7b, the 8^{th} and 9^{th} TBFs are conjoined in U-Net (left side) (JS: 0.7812, DC: 0.8623), MultiResUNet (right 7^{th} and 8^{th} TBFs) (JS: 0.808, DC: 0.8801) and UNet++ (both sides) (JS: 0.8113, DC: 0.8867). SIU-Net (JS: 0.831, DC: 0.9056) can define both the pairs clearly and also gives the highest Dice and Jaccard score. It can be concluded that SIU-Net performs better identification of TBF pairs than the other three methods.

5.4.5 SIU-Net is more consistent than MultiResUNet and UNet++

From the previous cases, it is ascertained that, while all of the three models (UNet++, MultiResUNet, and SIU-Net) outperform U-Net, the performance of SIU-Net is the best. Upon closer scrutiny of the last pair of TBFs in Fig. 5.8a, it is found that they are indiscernible in U-Net (JS: 0.7703, DC: 0.8588). But both are identified in MultiResUNet (JS: 0.7893, DC: 0.8607), UNet++ (JS: 0.7948, DC: 0.8798) and SIU-Net (JS: 0.8203, DC: 0.8974). However, the performances of the other methods with respect to the overall clarity of LBFs and TBFs are inconsistent.

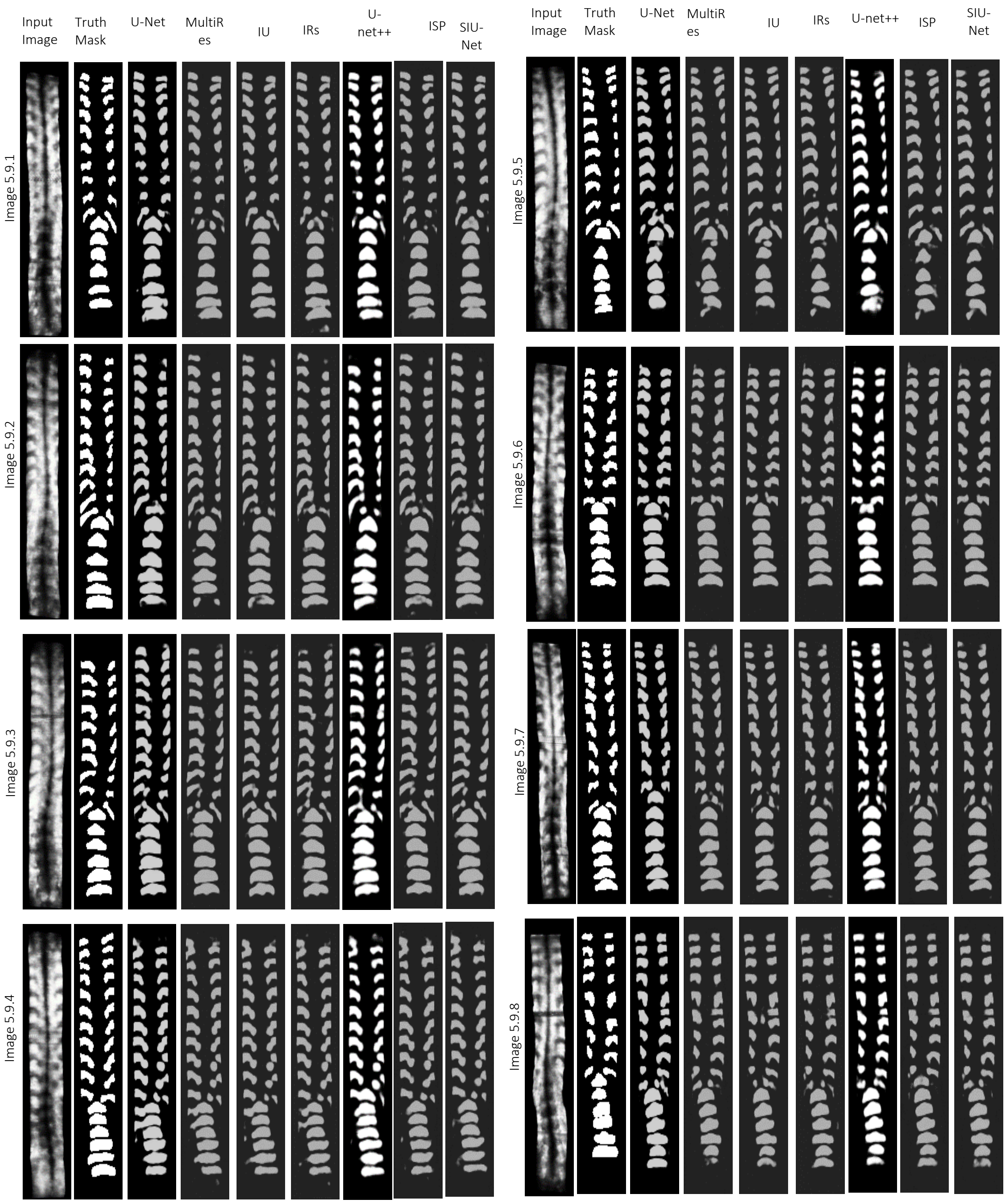


Fig 5.9. Qualitative comparison of all models using ultrasound spine image

Also, in Fig. 5.8b, U-Net (JS: 0.711, DC: 0.8375) is unable to identify 6 LBFs but able to segment all the TBFs clearly, whereas MultiResUNet (JS: 0.7509, DC: 0.8551) and UNet++ (JS: 0.7698, DC: 0.8698) identify 6 LBFs but fail to identify the 8th and 9th left TBFs. Unlike the other three methods, SIU-Net (JS: 0.7813, DC: 0.8916) identifies all the 6 LBFs clearly and segments all TBF pairs. However, the SIU-Net outperforms all the other models in the successful segmentation of both the TBFs and LBFs for its unique architecture. Hence, it can be concluded that SIU-Net gives the most consistent performance when compared to basic U-Net, UNet++, and MultiResUNet.

5.4.6 Further examples of qualitative and quantitative comparison of SIU-Net with other models

Fig. 5.9 shows a few more qualitative visual comparisons of bony feature segmentation results of the ultrasound spine image using SIU-Net and other methods: basic U-Net (Ronneberger, Fischer & Brox 2015), UNet++ (Zhou et al. 2019), MultiResUNet (Ibtehaz & Rahman 2020), IU, IRs model, and ISP model. Table 5.2 summarizes the Jaccard and Dice values of each image. In each of these cases, SIU-Net also qualitatively outperforms all the other segmentation methods.

Table 5.2: Quantitative evaluation of few selected individual images

Image No.	U-Net		MultiResUNet		Inception + U-Net		Inception + Res path		Unet++		Incep. + Encoder Skip Path		SIU-Net	
	JAC.	DICE	JAC.	DICE	JAC.	DICE	JAC.	DICE	JAC.	DICE	JAC.	DICE	JAC.	DICE
5.9.1	0.702	0.825	0.689	0.836	0.679	0.829	0.685	0.838	0.729	0.847	0.741	0.850	**0.758**	**0.853**
5.9.2	0.713	0.832	0.728	0.839	0.713	0.833	0.729	0.839	0.732	0.842	0.759	0.847	**0.763**	**0.851**
5.9.3	0.707	0.828	0.731	0.844	0.722	0.830	0.738	0.845	0.747	0.861	0.758	0.869	**0.767**	**0.873**
5.9.4	0.717	0.840	0.729	0.849	0.719	0.841	0.730	0.849	0.738	0.853	0.751	0.859	**0.759**	**0.863**
5.9.5	0.735	0.847	0.760	0.863	0.750	0.849	0.763	0.860	0.789	0.871	0.801	0.879	**0.820**	**0.883**
5.9.6	0.808	0.880	0.809	0.895	0.808	0.885	0.81	0.896	0.820	0.903	0.828	0.911	**0.841**	**0.919**
5.9.7	0.760	0.863	0.775	0.873	0.764	0.867	0.77	0.870	0.786	0.887	0.791	0.893	**0.803**	**0.901**
5.9.8	0.636	0.778	0.656	0.792	0.638	0.779	0.657	0.791	0.678	0.823	0.692	0.844	**0.703**	**0.853**

5.4.7 Comparison of accuracy of bony feature detection

The goal of a segmentation model is to accurately detect all the bony features which are present in their respective truth masks. A quantitative study is done to assess the accuracy in which the cases of the conjoint and broken features in an output segmentation image are considered to be failure cases. Also in this study, it is found that, in general, the chances of conjuncts are particularly high in the case of LBFs and the overall accuracy

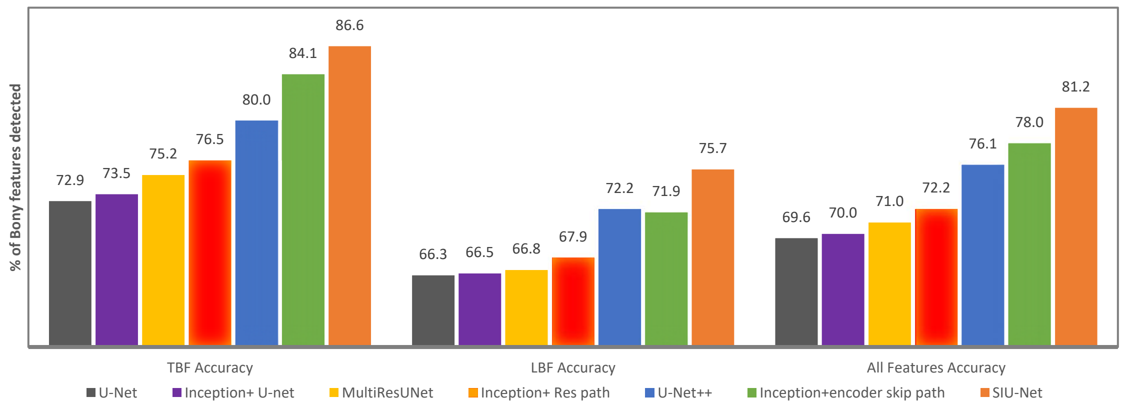


Fig 5.10. Proportion (%) of Images where segmentation outputs can detect all Bony features from respective Truth masks

of detecting LBFs is lower than that of the TBFs.

Fig. 5.10 depicts the comparison of all models in terms of percentage of images where all the bony features as well as individual thoracic and lumber bony features are detected clearly. From the figure, it is can be concluded that SIU-Net is the most accurate model, amongst the other models, to detect the bony features.

5.5 Discussion on SIU-Net

The quality of input images significantly impacts the quality of segmentation output. Among the imaging modalities used for scoliosis assessment, X-ray, MRI, and CT produce superior quality images while the non-radiating imaging modalities such as ultrasound are susceptible to noise and low contrast. U-Net is one of the most popular segmentation networks in deep learning. In previous work, U-Net performed very well in spine segmentation from X-ray images (Ronneberger, Fischer & Brox 2015). In a fully automated measurement of sagittal spinopelvic balance, U-net performed well in detecting anatomical landmark probability maps on the sacral endplate from sagittal X-ray images (Korez, Putzier & Vrtovec 2020) because the X-ray image is more apparent than the ultrasound image and, therefore, access to segment. There are inherent limitations of ultrasound imaging compared to X-ray, due to which ultrasound cannot always compete with X-ray imaging. Acoustic shadowing hides all bone surfaces that are deeper than the posterior surface of vertebrae and ribs. Therefore, basic U-Net is unable to segment all relevant features and provide missing bony features in segmentation output consistently. Table 5.3 compares the Dice score of multiple segmentation methods, both automatic and manual, using X-Ray (Horng et al. 2019), CT, and MRI (Li et al. 2021) with ultrasound. In the case of CT, the authors indicated that their method was not designed to tackle low-dose CT images and images with implants and only the medium to high-quality CT images was selected (dataset 3) for further segmentation (Khandelwal, Collins & Siddiqi 2021). The proposed method, SIU-Net, explicitly designed to manage the inherent drawbacks of ultrasound images, produces results comparable (by 93-95%) to other methods which used higher quality input images as input. Additionally, from the respective publications, it can be inferred that when compared to respective baseline U-Net outputs, SIU-Net gives the highest improvement (0.883 vs 0.817 – 8%) when compared to Residual U-Net (0.951 Vs 0.9410 - 1%) and MANet (0.925 Vs 0.9008 – 3%).

Table 5.3: Comparison of segmentation performance of various architectures in spine images using various imaging modalities

Author	Objective	Imaging Modality	Method	Method	Description of dataset used	Avg. Dice Score
Horng et al. 2019	Segmentation of vertebrae for scoliosis assessment	X-ray	Residual U-Net	Automatic	595 vertebra images; Five-fold cross validation (each fold images augmented to 1000, 10% validation images)	0.951 ± 0.03
Khandelwal et al. 2021	Segmentation of entire spine and individual vertebra to aid surgical planning	Computed Tomography (CT)	Region-based segmentation of spine	Manual	Dataset 1 (Lumbar vertebrae; 50 vertebrae extracted from 10 images)	0.924 ± 0.013
				Manual	Dataset 2 (Thoracular-Lumbar Vertebrae; 120 thoracic and 50 lumbar vertebrae across 10 subjects)	0.949 ± 0.022
				Manual	Dataset 4 (Cervical, thoracic and lumbar region; 43 medium-to-high dose CT images)	0.849 ± 0.753
			Shape-prior based flows for individual vetrebra segmentation	Automatic	Dataset 3: (Lumbar vertebrae; lumbar region of 30 patients, no. of slice range 55-200) (25 for training, 5 for evaluation), six-fold cross-validation	0.838 ± 0.031
Li et al. 2021	Segmentation of spine T2-weighted images to assess lumbar spinal stenosis	Magnetic Resonance Imaging (MRI)	Dual branch multi-scale attention network (MANet)	Automatic	1080 images of 120 patients (70% training, 20% validation, 10% testing)	0.925
Proposed Method	Segmentation of thoracic and lumbar bony features to assess Scoliosis	Ultrasound	Skip-Inception U-Net (SIU-Net)	Automatic	109 images from 109 patients (79 training, 30 testing), Five-fold cross validation	**0.883 ± 0.024**

Ultrasound imaging modalities that are non-radiating, safe, portable, and capable of real-time operation are explored for scoliosis assessment (Zheng, Lee, et al. 2016). Table 5.4 illustrates some popular ultrasound imaging techniques and the correlation of their respective measured scoliosis curvature angles to the radiographic Cobb angle method (Huang, Deng, et al. 2019), (Huang, Zeng & Li 2017), (Chen et al. 2021), (Brink et al. 2018). Huang et al. introduced a double-sweep 2.5-dimensional extended field-of-view

(EFOV) method that showed a good correlation of 0.993 (Huang, Zeng & Li 2017). However, in their research, the authors deduced that the surfaces of the transverse process may not be consistently visualized when scanning patients with scoliosis. The reconstruction of the panorama images was more time-consuming than the original image-cutting method. Chen et al. worked on 3D ultrasound with a fast reconstruction algorithm and found that it is impossible to apply the conventional slicing technique to visualize the hidden bony features of each layer from reconstructed 3D images (Chen et al. 2021). Zheng and his team developed a dedicated system called Scolioscan which employs 3D ultrasound scanning using the volume projection imaging (VPI) technique (Jiang et al. 2019). A coronal view of the spine similar to a posterior-anterior radiograph can be obtained from the reconstructed volume data using a simple re-slicing or volume rendering technique. One 3D ultrasound image generated is then split into nine 2D VPI images of different depths (Zhou et al. 2017). The re-slicing process helps to make the hidden bony features inside the spine volume more visible. This makes images from Scolioscan better suited for automatic scoliosis curvature assessment. Further, using VPI images, Brink et al. demonstrated high correlations of scoliosis curvature angle with the traditional Cobb angle method (Brink et al. 2018), for both thoracic and lumbar regions as shown in table 5.4.

Table 5.4: Comparison of scoliosis curvature angles measured using various types of ultrasound images with traditional Cobb angle

Author	Dimensions	Type of ultrasound images	Correlation (R^2) with traditional radiographic cobb angle
Huang et al. 2017	2.5D	2.5D extended field-of-view (EFOV)	-
Huang et al. 2019	2.5D	Double-sweep 2.5-dimensional extended field-of-view (EFOV)	0.993
Chen et al. 2021	3D US with fast reconstruction algorithm	Fast Dot-Projection algorithm- Voxel-based Nearest Neighbour (FDP-VNN)	Rater 1: 0.95 Rater 2: 0.90
		Fast Dot-Projection algorithm- Multiple Plane Interpolation (FDP-MPI4)	Rater 1:0.97 Rater 2: 0.97
Brink et al. 2018	3D	Volume Projection Imaging- Spinous Process (VPI-SP)	Automatic: 0.991 (Thoracic), 0.983 (lumbar) Manual: 0.987 (Thoracic), 0.970 (lumbar)
	3D	Volume Projection Imaging- Transverse Process (VPI-TP)	Manual: 0.992 (Thoracic), 0.985 (lumbar)

Table 5.5 summarizes the research work on all the three indices that are used with

ultrasound imaging for scoliosis measurement, namely Cobb Angle (Zheng, Young, et al. 2016), (Young et al. 2015); Spinous Process Angle (SPA) (Zhou et al. 2017), (Zheng, Lee, et al. 2016), (Brink et al. 2018), (Zeng et al. 2019), (Zeng, Ge, Gao, Zhou, Zhou, He, Lou, et al. 2021), (de Reuver et al. 2021), (Banerjee et al. 2020); and Transverse Process Angle (TPA) (Lee et al. 2021), (Ungi et al. 2020), (Brink et al. 2018). Due to the absence of vertebral bodies in ultrasound images, the angle between vertebra endplates of a scoliotic curve i.e. Cobb angle, cannot be sufficiently measured (Ungi et al. 2020) and the average correlation with radiographic Cobb Angle was 0.58 (blinded) and 0.85 (Aid Of previous Radiographs or AOR) (Zheng, Young, et al. 2016). Using

Table 5.5: Comparison of performance of various ultrasound scoliosis measurement indices with radiographic Cobb angle

Author	Ultrasound System	No. of patients	Method	Region of Interest (ROI)	Outcome vs Radiographic Cobb Angle	
					MAD	Correlation (R2)
ULTRASOUND COBB ANGLE						
Zheng et al. 2016	SonixTABLET	65	Manual (Blinded & AOR)	Coronal Curvature	Rater 1: 4.9°± 3.8° Rater 2: 4.6°± 3.8° (Blinded) Rater 1: 2.8°± 2.2° Rater 2: 2.7°± 1.9° (AOR)	Rater 1: 0.58 Rater 2: 0.58 (Blinded) Rater 1: 0.84 Rater 2: 0.87 (AOR)
Young et al. 2015	SonixTABLET US system	20	Manual (centre of lamina (COL) method)	Coronal Curvature	Rater 1: 2.6°± 2.0° Rater 2: 4.1°± 2.6° Rater 3: 3.8°± 3.3° Rater 4: 3.7°± 3.5°	-
SPINOUS PROCESS ANGLE (SPA)						
Zeng et al. 2019	SonixTABLET	50	Semi-Automatic (Using gradient vector flow (GVF) snake model)	Spinous process	Rater 1: 5.8° Rater 2: 6.6°	Rater 1: 0.75 Rater 2: 0.73
Zeng et al. 2021	SonixONE	92	Semi-Automatic (Using Stacked Hourglass Network)	vertebral spinous process (SP) & laminae	Rater 1: 5.7°± 4.5° Rater 2: 6.1°± 4.8°	Rater 1: 0.80 Rater 2: 0.75
S. Reuver et al. 2021	Scolioscan	70	Manual	Spinous Process	Thoracic: 6.5°± 3.9° Lumbar: 7.3°± 4.7°	Thoracic: 0.968 Lumbar: 0.923
Banerjee et al. 2020	Scolioscan	109	Automatic (U-Net segmentation)	Spinous Process	-	-
Brink et al. 2018	Scolioscan	33	Manual	Spinous column Profile	4.5° ± 3.1°	Thoracic: 0.987 Lumbar: 0.970
Brink et al. 2018	Scolioscan	33	Automatic	Spinous column Profile	4.9°± 3.2°	Thoracic: 0.991 Lumbar: 0.983

Zhou et al. 2017	Scoliosc an	99	Automatic (using Phase congruency method)	Spinous column Profile	-	0.83
Zheng et al. 2016	Scoliosc an	49	Manual (VPI-SP method)	Spinous column Profile	-	Thoracic: 0.784 Lumbar: 0.727
TRANSVERSE PROCESS ANGLE (TPA)						
Lee et al. 2021	Scoliosc an	164	Manual	Thoracic and Lumbar	Thoracic: 3.0° (0°-9.9°) Lumbar: 2.8° (0°-11.9°)	Thoracic: 0.893 Lumbar: 0.884
Ungi et al. 2020	MicrUs EXT-1H	8	Automatic (U-Net segmentation)	Thoracic and Lumbar	2.2°	-
Brink et al. 2018	Scoliosc an	33	Manual	Thoracic and Lumbar	4.7°± 3.6°	Thoracic: 0.992 Lumbar: 0.985

posterior anatomical landmarks, alternate ways to measure scoliosis in ultrasound are being worked on. Two such measures are SPA and TPA. SPA is the angle formed between the lines drawn through the most tilted part of the spinous column profile of coronal ultrasound images while the TPA is measured from the lateral bony features of the spine. Both SPA and TPA have been proven comparable to traditional radiographic Cobb angle (Ungi et al. 2020), (Brink et al. 2018). However, the SPA measurement process has limitations in that the raters are disturbed by scattered and invalid points when manually tuning the inflection points of the spinous process curve, and thus the measurements of SPA were affected significantly in the lumbar area (Zheng, Lee, et al. 2016), (Brink et al. 2018). In addition, SPA measurement uses different landmarks than the radiographic Cobb angle, introducing more variation sources for each measurement (Lee et al. 2021). The variation between SPA and traditional Cobb angle is more pronounced for greater curvature of the spine (Wu, Liu & Wong 2020). The transverse process is an alternative way to assess scoliosis and the measuring index is called the transverse process angle (TPA). TPA shows an excellent correlation with radiographic Cobb angle (Brink et al. 2018). In (Ungi et al. 2020), the transverse process and ribs were segmented on 2-D transverse ultrasound images and reconstructed through 3D reconstruction on the coronal plane. Measured from a 3D reconstructed image, their research gave a mean average difference (MAD) of 2° between X-ray-based angle and ultrasound TP angle and this difference is within the clinically significant error range. Ultrasound curvature angle (UCA) is a type of TPA, which is measured on coronal curvature on the AIS spine and uses thoracic and lumbar bony features as regions of interest (Lee et al. 2021). Hence, given the advantages of TPA/UCA over SPA and ultrasound Cobb angle, this research is focused on automation of the UCA measurement process.

Table 5.6: Summary of segmentation architectures applied to ultrasound images of spine and other body parts

Author	Field of work	Challenges	Seg. Arch. used	Special Features	Summary of dataset used	Dice Score	AUC	Jaccard/mIOU
				OTHER BODY PARTS				
Amiri et al. 2020	Breast Lesion	Complexity of lesion shape and location	Two stage U-Net	one U-Net for ROI detection, one for segmentation	Training - 1398, Testing - 1892, Total - 3290 Cross Validation: Five Fold	80.5	-	-
Chen et al. 2021	Foetal spina bifida	Large number of noise spot	Oct-U-Net	Octave feature to reduce redundant information	3,300 pregnant women's foetus images	-	-	91.7
				SPINE				
Ungi et al. 2020	Sagittal spine image	Automatic segmentation with a CNN and volume reconstruction	Basic U-Net	Basic U-Net features	No of Images: Training - 1398, Testing - 1892, Total - 3290 Cross Validation: One Cross	-	97	-
Huang et al. 2020	Thoracic and lumbar bony features	Noise	RSN-U-net	Total variance loss to improve the robustness against the speckle and regular occlusion noise	No of Images: 109; Training - 80, Testing - 29	78.38	98	-
Lyu et al. 2021	Thoracic and lumbar bony features	Noise, Variability of ROIs	D-TV Net	Two branches to estimate semantic region and contour segmentation, ASPP module to concatenate different scales of features	No of Images: 109;Training - 80, Testing - 29, Cross Validation: 3 fold	86.68	-	-
Banerjee et al. 2021	Thoracic and lumbar bony features	Noise, Variability of ROIs	LDS U-Net (Chapter IV)	Light Dense Block to increase computation efficiency, Multiscale Skip-pathway to enhance feature fusion & Selection Gates to identify target bony features.	No of Images: 109;Training - 79, Testing - 30, Cross Validation: 3 fold	86.94	-	-
Proposed Method	Thoracic and lumbar bony	Noise, Variability of ROIs	SIU-Net	Improvised inception block to handle large variability in spine images & re-designed Dense-skip connection	No of Images: 109;Training - 79, Testing - 30, Cross Validation: 5 fold	**88.3**	**99**	**78.0**

feature s	for multi-scale feature fusion	

For automation of these measurement processes, research has been carried out to efficiently segment the relevant region of interest (ROI) from an ultrasound image as a step before the actual angle measurement. Ultrasound images have high speckle and scan noise which makes demarcation of necessary information challenging. For automatic SPA measurement, the segmentation was done for the mid spine line (single ROI) and its segmentation was done using a basic U-net (Banerjee et al. 2020). On the other hand, for automatic UCA measurement, the proper identification and segmentation of thoracic and lumbar bony features are of prime importance. Single ROI segmentation techniques such as basic U-Net will not handle the variability in locations, sizes, and shapes of multiple ROI (TBFs and LBFs) in a noisy ultrasound image. Table 5.6 summarizes the performance of various segmentation architectures applied to ultrasound images of a breast lesion, fetal spina bifida, and spine.

In (Huang et al. 2020), an improvised version of U-Net i.e. U-Net with robustness to speckle and regular occlusion noise (RSN-U-Net) was introduced to segment lateral bony features from noisy ultrasound spine images. A new technique, i.e. total variance loss, was employed to improve the robustness against the speckle and regular occlusion noise. The research showed that RSN-U-Net only marginally outperformed basic U-Net with a segmentation result of 0.78 (Dice score) against the score of 0.76 by basic U-Net; indicating the variability of location, shape, and sizes of bony features was not specifically and fully addressed through the RSN-U-Net architecture.

Lyu et al. employed a dual-task ultrasound transverse vertebrae segmentation network (D-TVNet) to segment lateral bony features from noisy ultrasound spine images (Lyu et al. 2021). The Atrous Spatial Pyramid Pooling (ASPP) (Lian et al. 2021) module was adopted to extract effective features from the spine images. An ASPP module is composed of four parallel atrous convolution layers with different dilated rates. Feature fusion was done by merging all the features extracted by each atrous convolution layer from different receptive fields. The four parallel atrous convolution layers in ASPP can achieve the effect equivalent to applying multiple filters with their different receptive fields on a given input image. In that work, the number of components in ASPP was increased by one by fusing another extra feature stream into ASPP concatenation. Three limitations could be observed: a) ASPP with limited sampling ranges will not be able to comprehensively extract the features of the target entities with variable sizes, b) some entities are so far from the ranges enclosed by the convolution kernels of ASPP that the features they have cannot be sampled, and c) to reduce the computational burden, all the points, but for the operative sampling points, in the convolution kernels are filled with zeros. This may result in a situation wherein, for the final result, the convolution kernel samples the information surrounding a particular pixel and may disregard the delicate local features corresponding to the positions with zeroes (Lian et al. 2021). The sizes of the entities from which the ASPP module collect information, only vary within a limited range which, in the real scenario, will not suffice as the variability in location, shape, and sizes of entities would hinder the sampling of adequate information to generate the complete and precise features required for automatic UCA measurement.

In Chapter IV, a novel hybridized lightweight convolutional neural network architecture is presented, called Light-Convolution Dense Selection U-Net or LDS U-Net. This architecture had two main features – (a) Attention gates that improved the segmentation clarity by tackling the 'noisy' information and (b) Multi-scale skip-pathways replaced the conventional skip-pathways so that the problem of large variabilities in shape, size, and locations of the TBFs and LBFs could be handled. However, the attention gate (Oktay et al. 2018) has a drawback, when used in basic U-Net architecture, it requires a significantly more number of parameters but produces only marginal improvement in segmentation output.

A novel hybridized CNN architecture, multi-scale feature fusion skip-inception U-Net or SIU-Net is introduced in this research for ultrasound spine segmentation. 109 2D ultrasound spine images and their expert suggested truth masks are used as the input image dataset. Basic U-Net was taken as the starting architecture (Ronneberger, Fischer & Brox 2015). Basic U-Net gave segmentation output with missed and conjoint bony features (Avg. Jaccard: 0.709 and avg. Dice Score: 0.817). After analyzing the output images, it was deduced that several bony features got missed out due to the lack of flexibility of U-Net to choose multiple sizes of filters. As the first modification to increase the segmentation performance, the option of selecting multiple sizes of filters was incorporated with the basic U-Net architecture and the network was called the Inception+U-Net (IU) model. Instead of unnecessarily adding more layers in the basic U-Net, an improvised version of the inception block was used to replace the conventional convolution operation. As the inception model was equipped with convolutional layers of varying kernel sizes, it was anticipated that this concept would solve the variability issues. However, after evaluation of the segmentation output, it was seen that the IU model did not perform satisfactorily, especially in and around the noisy areas (avg. Jaccard: 0.711 and avg. Dice Score: 0.846).

Further analysis showed a high degree of inconsistency between the features passed from the encoder network and features transmitted through the decoder network. The combination of two inconsistent sets of features caused incongruity during the learning of the network and affected the segmentation result (Ibtehaz & Rahman 2020). The conventional skip path of the IU model was then replaced by the Residual path (Ibtehaz & Rahman 2020) to resolve the two incompatible sets of features in encoder and decoder sides and the model was named as Inception+Res path (IRs) model. This model performed better (avg. Jaccard: 0.734 and avg. Dice Score: 0.855) than the IU model as some further processing was incorporated in skip connections to make the feature maps from both sides more consistent. With this result, it was inferred that the IRs model could perform better in scenarios with less noise and clear visibility. However, as the input dataset is noisy ultrasound spine images, single-scale feature fusion was insufficient for this segmentation work. A dense network was needed to perform multi-scale feature fusion to extract deeper layer features and fuse with the features from shallower layers. Then, the residual skip paths were replaced with encoder side dense skip connection (Zhou et al. 2019) and it generated promising segmentation output. The new model was named as Inception + encoder skip path (ISP) model. As ultrasound image noise is the biggest challenge for this research, the ISP model, in many cases, failed to provide a clear segmentation edge boundary (avg. Jaccard: 0.757 and avg. Dice Score: 0.876). Therefore, for more semantically rich feature extraction and fusion, the ISP model skip connection was replaced with a decoder side dense skip connection. This model, SIU-Net, finally solved the problem of variability of bony features and speckle noise with better identification and segmentation of bony features with proper bone edge boundary detection (avg. Jaccard: 0.78 and avg. Dice Score: 0.883).

The proposed network shows a promising segmentation output compared to baseline network U-Net and advanced networks, such as MultiResUNet and Unet++. The performance of SIU-Net is evaluated quantitatively using three popular indices, i.e. Jaccard index, Dice coefficient, and histogram Euclidean distance. Each time, SIU-Net outperforms all the other popular segmentation techniques giving the maximum Jaccard index (0.78) and Dice coefficient (0.883) and minimum histogram Euclidean distance (0.011) in comparison with baseline network U-Net, MultiResUNet, and Unet++. This is because an inception block, by its nature of construction, is more adept in extracting features from different locations, shapes, and sizes than conventional convolution layers. Also, closer examination of the individual segmentation result, for some special cases, points to the fact that the SIU-Net does a superior job of identifying and segmenting the TBF and LBF pairs than the other networks (Fig 5.5, 5.6, 5.7 & 5.8). Finally, though MultiResUNet and Unet++ provide promising segmentation output in a few individual cases, the result demonstrates that the proposed network is more consistent in giving the adequate segmentation output for the overall range of input images.

5.6 Key takeaways of SIU-Net

In the manual UCA process, the evaluators must locate the proper points for line placements on the ultrasound images. Identifying the feature below and above the T12 level depends on the expertise and judgment of the evaluators (Lee et al. 2021). The overall process flow of automatic UCA measurement is shown in Chapter I, figure 1.4. A key element in the automatic UCA measurement is identifying and segmentation of the lateral bony features (thoracic and lumbar) as they play a pivotal role in angle calculation. This can be measured using an index called the bony feature detection rate and a comparison of detection rates between other contemporary segmentation techniques is presented in Table 5.7.

Table 5.7: Comparison of detection rates for various segmentation architectures in ultrasound scoliosis measurement

Author	Method	Avg. Detection Rate (%)
Huang et al. 2020	RSN-U-Net	69.34
Lyu et al. 2021	D-TV Net	75.29
Banerjee et al. 2021	LDS U-Net (Chapter IV)	76.59
Proposed method	**SIU-Net**	**81.20**

It can be concluded that SIU-Net gives the best detection rate not only against general segmentation architectures such as U-Net, U-Net++, and MultiResUNet (Fig. 5.10) but also outperforms other architectures dedicated specifically to ultrasound spine images (Table 5.7). The future step of this research is to calculate the UCA angle automatically from segmented images and validate it against existing techniques and is detailed in Chapter VI.

CHAPTER VI: AUTOMATIC CALCULATION OF ULTRASOUND CURVE ANGLE

6.1 Introduction

In clinical AIS diagnosis with 3D ultrasound, experts need to observe hundreds of images of the whole spine region in a sequence. This process is tedious and time-consuming. For faster diagnosis and better visualization of the spine structure, volume projection imaging (VPI) was proposed to project the voxels of 3D ultrasound volume data onto a sequence of 2D spine coronal planes (Cheung, Zhou, Law, Mak, et al. 2015). However, owing to the low quality (speckle noise and low saturation) of the ultrasound images and the acoustic shadow caused by the high acoustic impedance of the bones (Pandey et al. 2020), the experts conducting the examinations have to be highly qualified in sonographic assessments. Nevertheless, the subjective factors behind personal experience are inevitable in manual ultrasonic scoliosis diagnosis. Therefore, the current clinical workflow can benefit significantly from an automatic method for spine deformity measurement (Cheung, Zhou, Law, Mak, et al. 2015), (Brink et al. 2018).

Automatic spine angle measurement would require pre-processing, namely image segmentation of the ultrasound spine image, so that the regions of interest (thoracic and lumbar regions) can be demarcated and the angle can be subsequently measured. The steps for automatic UCA measurement are as follows:

1. **Collection of data:** A total of 109 2D spine ultrasound images of 109 patients with varying degrees of scoliosis are collected from Scolioscan (Hong Kong Polytechnic University) (discussed in Chapter III).
2. **Segmentation of bony features:** All 109 spine ultrasound images are segmented using six contemporary and two proposed segmentation techniques for demarcation of the thoracic and lumbar bony features. Depending upon the segmentation accuracy and detection rate, the best segmentation technique is selected among the eight for automatic UCA calculation (discussed in Chapter IV and V).
3. **RGB conversion of segmented binary images (proposed Algorithm 1):** In binary segmented images, isolation of ribs and thoracic bony features (TBFs) are difficult since they are connected. The binary segmented images are, therefore, converted to RGB images using specific colour codes for ribs and

TBFs and for LBFs (discussed in this chapter).

4. **Automatic UCA calculation by centroid pairing and inscribed rectangle slope (CPI-SLO) method:** In this step, the key features needed for automatic UCA calculation, i.e. TBFs and LBFs, are identified and separated from RGB images. Then, the main thoracic and thoraco-lumbar angles are calculated separately as follows (discussed in this chapter):
   i. Calculation of main thoracic angles (proposed Algorithm 2): The main thoracic angles are calculated using the proposed slope calculation using the centroid pairing method.
   ii. Calculation of thoraco-lumbar angles (proposed Algorithm 3): The thoraco-lumbar angles are calculated using the proposed slope calculation involving the largest inscribed rectangle method.

Fig 6.1 presents the complete flowchart for automatic UCA calculation using the centroid pairing and inscribed rectangle slope (CPI-SLO) method.

6.2 Selection of the most suitable segmentation technique for subsequent angle calculation

In an ideal case, the segmented output should be able to demarcate all the TBFs and LBFs. However, since the images are ultrasound noisy images, some segmentation outputs have multiple missing or conjoined bony features and are of little use for the calculation of UCA and cannot be taken into account for angle calculation. Also, the variability between thoracic and lumbar features add to the complexity, as both are necessary to estimate the main thoracic and thoraco-lumbar angles for a complete UCA estimation.

6.2.1 Quantitative evaluation of various segmentation techniques

One baseline segmentation technique, U-Net (Ronneberger, Fischer & Brox 2015), and seven state-of-the-art segmentation techniques, such as Attention U-Net (Oktay et al. 2018), UNet++ (Zhou et al. 2019), MultiResUNet (Ibtehaz & Rahman 2020), RSN U-Net (Huang et al. 2020), D-TV Net (Lyu et al. 2021), LDS U-Net (Chapter IV) and SIU-Net (Chapter V), have been employed to segment the noisy ultrasound spine VPI images. These images were compared quantitatively using the Dice score and Jaccard Index, as shown in Table 6.1.

The SIU-Net model outperformed the baseline model U-Net by a huge margin. When

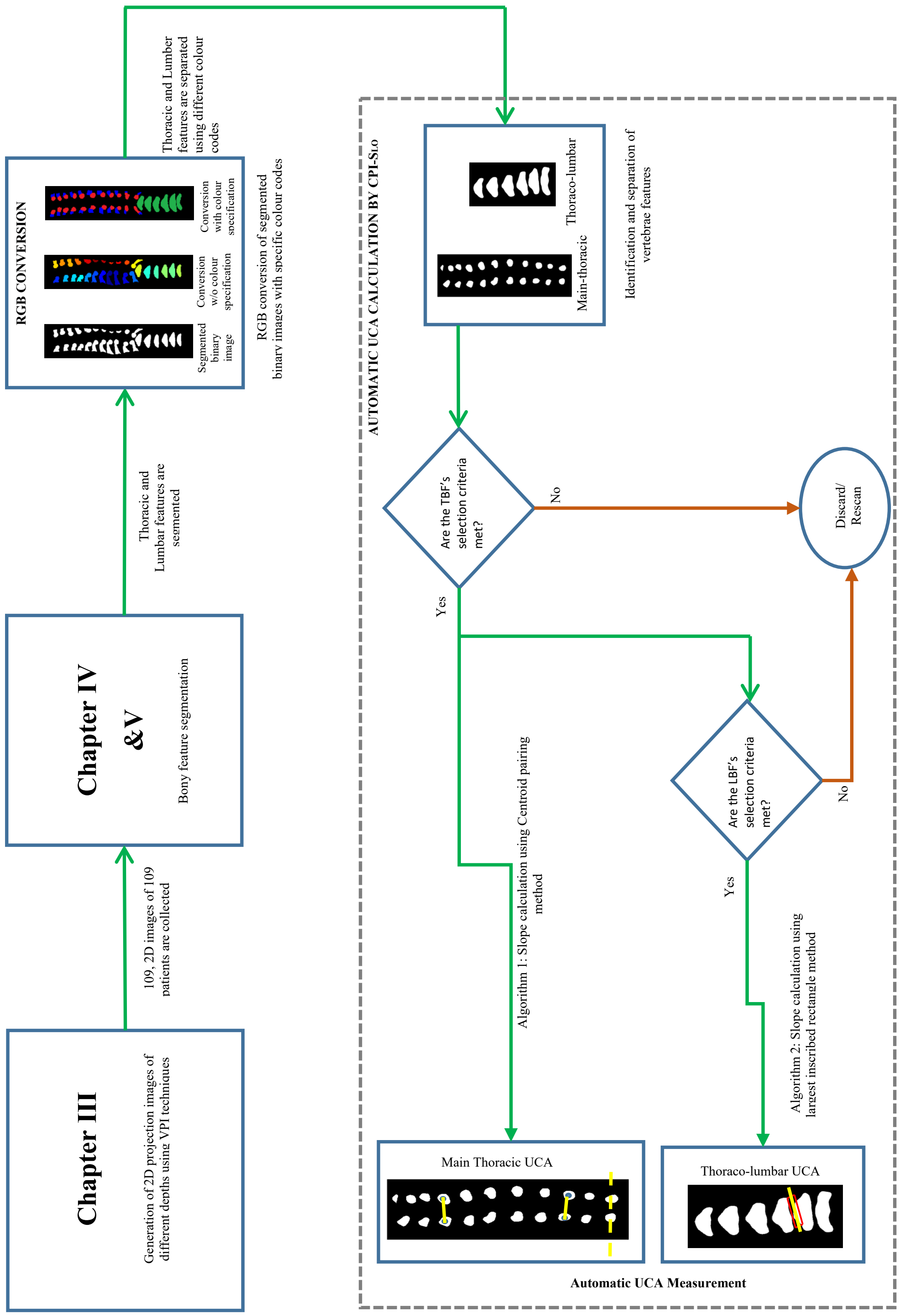


Fig 6.1. Process flow of Centroid Pairing and Inscribed rectangle Slope (CPI-SLO) method

compared with the two popular medical image segmentation models, MultiResUNet and Unet++, SIU Net showed a 4% and 2.5% improvement, respectively, on the averaged Dice score.

Further, D-TV Net and LDS U-Net were developed specifically to handle ultrasound noisy spine images; SIU-Net outperformed them approximately by 1.8% and 1.5% on the averaged Dice score, respectively. Hence, the SIU-Net model is superior in the segmentation of noisy ultrasound images. Therefore, it can be considered for automatic UCA calculation.

Table 6.1: Comparison of quantitative segmentation results for various segmentation architectures

METHOD	DICE	JACCARD
U-Net (Ronneberger, Fischer & Brox 2015)	0.8133	0.7015
Attention U-Net (Oktay et al. 2018)	0.8297	0.7189
Unet++ (Zhou et al. 2019)	0.861	0.748
MultiResUNet (Ibtehaz & Rahman 2020)	0.8458	0.7264
RSN U-net (Huang et al. 2020)	0.7838	-
D-TV Net (Lyu et al. 2021)	0.8668	-
LDS U-Net (Chapter IV)	0.8694	0.7415
SIU-Net (Chapter V)	0.883	0.781

The key element in the automatic UCA measurement is identifying the lateral bony features (thoracic and lumbar) as they play a pivotal role in angle calculation. This can be measured using an index called the bony feature detection rate. A comparison of detection rates between baseline and other contemporary segmentation techniques is presented in Table 6.2. It can be concluded that SIU-Net gives the best detection rate (81.2%) not only against general segmentation architectures such as U-Net, Attention U-Net, U-Net++ and MultiResUNet but also outperforms other architectures dedicated specifically to ultrasound spine images.

Table 6.2: Comparison of detection rates for various segmentation architectures

METHOD	AVG. DETECTION RATE (%)
U-Net	69.6
Attention U-Net	70.2
Unet++	76.1
MultiResUNet	71.0
RSN-U-net	69.34
D-TV Net	75.29
LDS U-Net (Chapter IV)	76.59
SIU-Net (Chapter V)	81.20

6.2.2 Evaluation of various segmentation techniques for usability

It must be noted that between the two segmented sections of TBFs and LBFs if LBFs are not segmented satisfactorily, it is still possible to calculate the main thoracic angle. On the other hand, if TBFs are not segmented satisfactorily, both the main thoracic and thoraco-lumbar angles are impossible to measure. However, for understanding the performance of CPI-SLO, both the cases are tackled separately as

Set A: Images where only the main thoracic angles can be calculated and

Set B: Images where both the main thoracic and thoraco-lumbar angles, i.e. complete UCA, can be calculated.

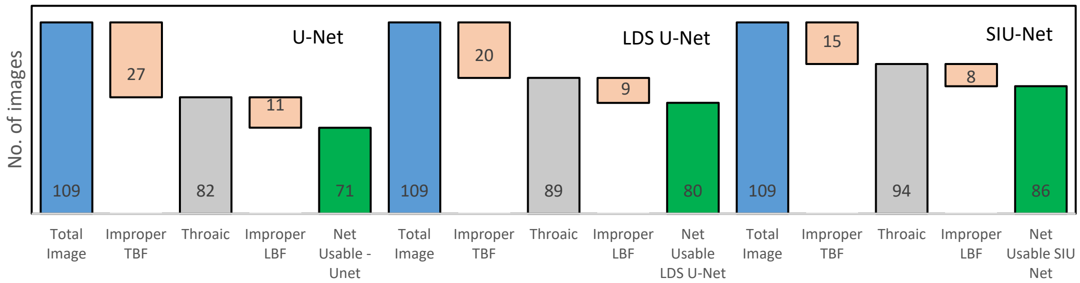


Fig 6.2. Comparison of usability of images for automatic scoliosis assessment

Fig 6.2 outlines the comparison between the usability of output images from the three segmentation architectures for the two sets of images mentioned above. Out of the total 109 images, SIU-Net produced the most usable images (86 images) for complete estimation of UCA compared to LDS U-Net (80 usable images) and U-Net (71 usable images). Also, SIU-Net generates the least loss during TBF segmentation (15 images) when compared to SIU-Net (20 images) and U-Net (27 images) for estimation of just the main thoracic angles.

The outcome of the bony feature segmentation using SIU-Net showed that, unlike some baseline segmentation architecture results, the output segmented images using SIU-Net were not affected by the ultrasound scan noise. Moreover, SIU-Net demarcated the boundary areas of TBFs and LBFs more clearly than other contemporary segmentation techniques.

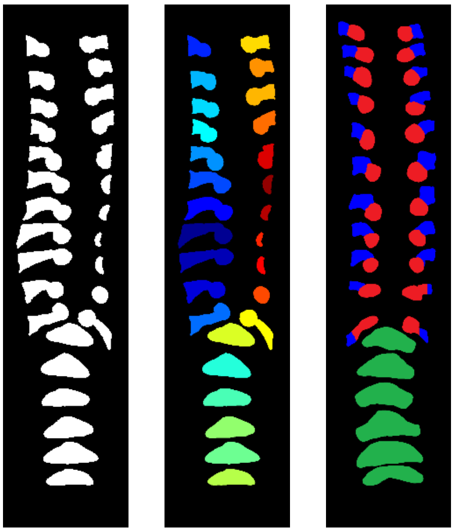
6.3 Demarcation of Regions of Interest

The regions of interest for automatic UCA calculation are TBFs and LBFs. LBFs can be demarcated easily as they appear as distinct entities in the binary segmented images. However, the rib and thoracic bony features (TBF) appear as a single entity without any noticeable boundary separation, which makes their demarcation particularly challenging. To distinguish the TBFs from the ribs in binary segmented images, it is proposed to convert the binary images to RGB images using specific colour codes. However, there are two main challenges to this approach:

i. Simple binary to RGB conversion of the segmented image cannot determine the ribs, TBFs and the LBFs in three colours, rather it is open-ended and gives a multi-coloured output.

ii. A proper boundary condition between the connected ribs and TBFs must be defined to demarcate them using two distinct colours.

Therefore, a novel algorithm for binary to RGB conversion with specific colour codes is proposed before automatic UCA calculation.



(a) (b) (c)

Fig 6.3. RGB conversion of binary-segmented ultrasound spine images (a) segmented binary image (b) RGB conversion without specific colour code (without algorithm 1) & (c) RGB conversion with specific colour code (with proposed algorithm 1)

6.3.1 Methodology: binary to RGB conversion of segmented images

For better identification of thoracic and lumbar bony features, the segmented binary outputs are converted to RGB images with three different colour codes: red for thoracic bony features, blue for ribs, and green for lumbar bony features.

The steps are as follows:

Let A_C denote the total number of contours in binary spine images, where C represents the total number of bony features detected in a single segmented spine image. The procedure to convert the binary segmented images to RGB with specific colors in specific areas is described in the sequence of operations below:

1. RGB Conversion without specific colour coding:

The segmented binary images are converted to RGB images with no specific colour definition, i.e. TBF, LBF, and ribs are not in different colours (Fig 6.3b).

2. Demarcating the LBFs:

i. High dense and low orientation areas are selected $[A_L]$ and marked as class 3.

ii. All class 3 areas are converted to green colour.

3. Demarcating the TBFs and ribs:

i. The contours $A_{T+R} = [A_C\text{-}A_L]$ are separated as $[A_{(T+R)Left}\ ]\ and\ [A_{(T+R)Right}]$, respectively.

ii. The horizontal profile, i.e. the height of the contours of each column are calculated as $H^m_{left}\ and\ H^n_{right}$, where m and n are the total numbers of contours in the left and right columns.

iii. The narrowest part of each contour, i.e. valleys are calculated as V^m_{left} and V^n_{right}.

iv. The left column contours are separated into two classes using the valley. The right part of the valleys $[A_{T\ Left}]$ with height $H^m_{left\ T} = [H^m_{left} - V^m_{left}]$ are marked as class 2 and converted to the color red. The left part of the valleys $[A_{R\ Left}] = [A_{(T+R)Left} - A_{T\ Left}]$ are marked as class1 and converted to the color blue.

v. The right column contours are separated into two classes using the valley. The left part of the valleys $[A_{T\ Right}]$ with height $H^n_{right\ T} = [H^n_{right} - V^n_{right}]$ are marked as class 2 and converted to the color red. The right part of the valleys $[A_{R\ Right}] = [A_{(T+R)Right} - A_{T\ Right}$ are marked as class1 and converted to color blue (Fig 6.3c).

Algorithm 1: Proposed algorithm for binary to RGB conversion

Initialization: A_C denotes the total number of contours of segmented bony features in binary spine images. T_d and T_A are the threshold values for density and orientation of contours and selected as 0.95 and 20 respectively. m and n are the total number of segmented bony features in left and right columns respectively.

Output: The ribs, thoracic, and lumbar bony features will be separately marked in color blue, red, and green respectively.

Convert the segmented binary images to RGB images with no specific color definition, i.e. TBF, LBF, and ribs are not in any specific colors.

Calculate density and orientation for each segmented contours.

for i = 1, . . . , A_c

if density> T_d && orientation < T_A

Mark the contours $[A_L]$ as class 3 and colour them as green.

end if

Separate $A_{T+R} = [A_C\text{-}A_L]$ contours in left and right columns as $[A_{(T+R)Left}\ ]\ and\ [A_{(T+R)Right}]$ respectively.

for $A_{(T+R)Left}$= 1 ,…., m

Calculate heights of contours $[A_{(T+R)Left}]$ as H^m_{left}.

Calculate the narrowest parts i.e. valleys of contours $\{A_{(T+R)Left}\}$ as V^m_{left}.

Mark contours $[A_{T\ Left}]$ with heights $H^m_{left\ T} = [H^m_{left} - V^m_{left}]$ as class 2 and colour them as red.

Mark the rest of the contours $[A_{R\ Left}] = [A_{(T+R)Left} - A_{T\ Left}]$ as class 1 and colour them as blue.

end for

for $A_{(T+R)Right}$= 1 ,…., n

Calculate heights of contours $[A_{(T+R)Right}]$ as H^{n}_{right}.

Calculate the narrowest parts i.e. valleys of contours $\{A_{(T+R)Right}\}$ as V^{n}_{right}.

Mark contours $[A_{T\ Right}]$ with heights $H^{n}_{right\ T} = [H^{n}_{right} - V^{n}_{right}]$ as class 2 and colour them as red.

Mark the rest of the contours $[A_{R\ Right}]$ = $[A_{(T+R)Right} - A_{T\ Right}]$ as class 1 and colour them as blue.

end for

end for

6.4 Automatic calculation of Scoliosis angle (UCA)

Manual UCA measurement has been established to have a good agreement with the traditional radiographic Cobb angle method (Lee et al. 2021). Therefore, the sequence of the manual UCA measurement process is encompassed in this work for automatic angle measurement. Notably, in UCA measurement, unlike SPA, the main thoracic and thoraco-lumbar angles are calculated separately. The steps for manual UCA angle calculation are as follows:

1. The T12 pair ribs, generally the last pair of ribs from the bottom of the spine ultrasound image, are first identified to classify vertebrae levels.
2. Among all the thoracic bony feature pairs above the T12 level, the most tilted pairs are identified manually, and lines are drawn through the centres of the most tilted pairs.
3. For the region below the T12 level, the widest lump (lumbar feature) or the most tilted lump is selected manually, and a line is drawn through the centre of that lump.
4. The most tilted lines in the main thoracic and thoraco-lumbar areas are selected for calculating the main thoracic and thoraco-lumbar angles.

However, due to the intrinsic nature of noisy ultrasound imaging modality, inadequacy and ambiguity of bony feature information are an inevitable part of a segmented image. In the manual UCA technique, the experts make a judgment call when such cases are encountered. But for automatic measurement of UCA, the presence of all the bony features is essential for an accurate calculation. The need for the selection of the right image, which has adequate lateral feature information, makes the process of automatic calculation more complicated. The proposed algorithms of UCA measurement are constructed such that they can smartly adapt themselves depending on the quality of bony feature information in the segmented outputs. The algorithms use centroid pairing and the largest inscribed rectangle to find the slope of the thoracic and lumbar regions, respectively. This is called the CPI-SLO method.

6.4.1 Identification of vertebrae features

The RGB conversion of the segmented binary images enables the identification of the bony features with different colour provisions, as shown in Fig 6.4b. The image is then bifurcated using the bilateral features, along the medial dark line formed by spinous process shadow, into two separate images: one for the thoracic transverse process (Fig 6.4c) and another for lumbar lumps (Fig 6.4d). Also, the rib regions (blue) are not required for the UCA calculation and are not included in any of the outputs.

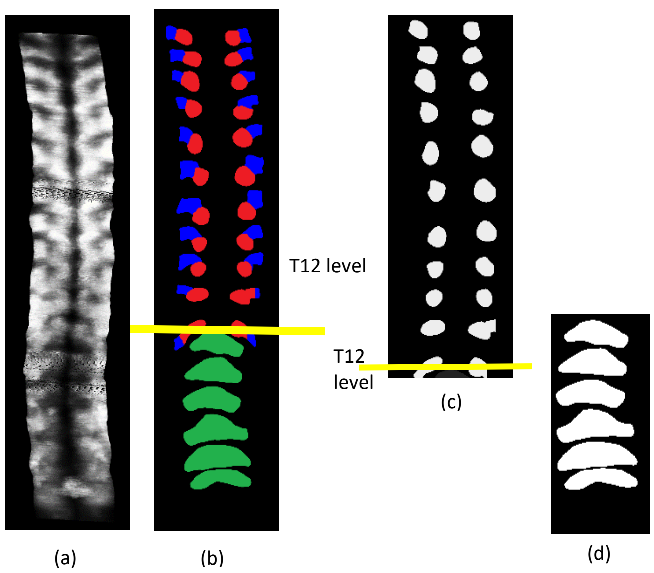


Fig 6.4. Identification of vertebrae features (a) Input image, (b) segmented bony feature (c) identified TBFs (d) identified LBFs

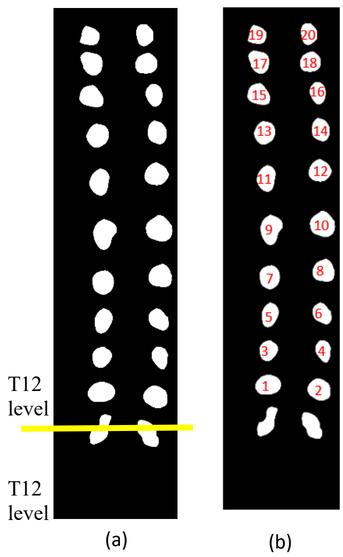


Fig 6.5. Measurement of main thoracic angle using Centroid pairing method

6.4.2 Measurement of main thoracic angle

Let $C^{T}_{Contour}$ denote the contours of the thoracic transverse processes detected on the segmented spine image, where T represents the total number of thoracic bony features detected in a single segmented spine image. The procedure to calculate the thoracic angle from the contours is described in the sequence of operations below:

1. **Centroid calculation and sorting**: The centroid of each contour is calculated in a row-by-row manner and then sorted according to their positions in the left and right columns. $\{[Cen^{M}_{left}]$, M = 1, . . . , $P_m\}$ determines the centroids of the left column contours, M denotes the number of contours present in the left column, and P_m is the total number of contours present in the left column. $\{[Cen^{N}_{right}]$, N = 1, . . . , $P_n\}$ determines the centroids of the right column contours, N denotes the number of

contours present in the right column, and P_n is the total number of contours present in the right column, as shown in Fig 6.5.

2. **Ascertaining the T12 region**: The traverse operation goes from bottom to top, as the lower-most two centroids (left and right) are the centroids of the T12 region. Let the two centroids of the T12 region be $[Cen_{left}^{1}]$ and $[Cen_{right}^{1}]$; they are considered to be the reference regions for angle calculation.

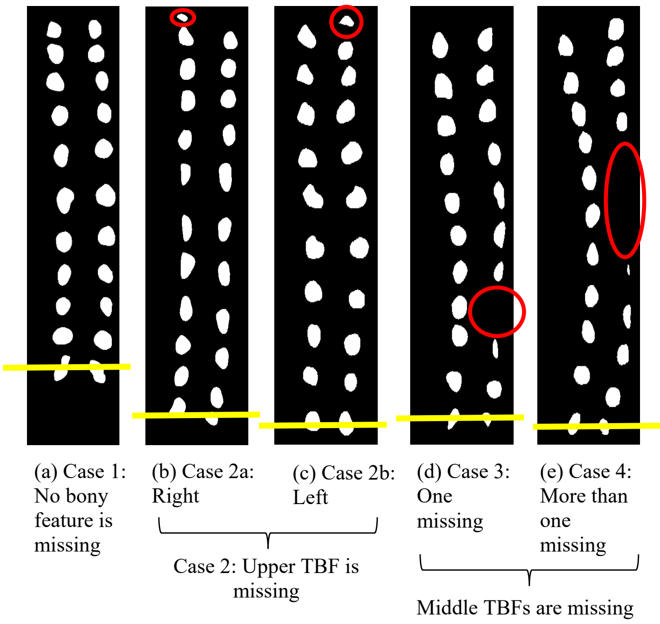


Fig 6.6. Different cases of measurement of main thoracic angle using Centroid pairing method

3. **Identification of missing TBFs**: Sometimes, the segmentation model fails to identify one or more thoracic bony features in either side or both sides of the spinous column profile; such special cases need to be considered for thoracic angle measurement.

i) **Case I - when no bony feature is missing in either side of the spinous column profile, i.e. $P_m = P_n$** (Fig 6.6a)

Slopes $\{[M_{slope}^{P_x}]$, x = 2, . . . , $P_m\}$ are calculated for each corresponding centroid pair from $\{[Cen_{left}^{M}]$ and $[Cen_{right}^{N}]$, where $M = N\}$ using the following equation:

$$m = \frac{y_2 - y_1}{x_2 - x_1} \quad (6.1)$$

Where m is the calculated slope between two points, and (x_1, y_1) *and* (x_2, y_2) are the coordinates of two points.

ii) **Case II - when one bony feature is missing in either side of the spinous column profile, i.e. $P_m \neq P_n$ and $|P_m - P_n| = 1$**

There are two possibilities of missing bony features: there may be one bony feature missing on the top of any column, or any bony feature missing from somewhere in the middle of any column. If the top bony feature in any of the columns is missing, then the corresponding bony feature of the other column is discarded since it cannot make a pair for angle measurement.

a) **If the bony feature in the right column is missing, i.e. $\boldsymbol{P_m - P_n = 1}$** (Fig 6.6b)

To detect the missing top bony feature, the distance (d_{12}) between the $Cen_{left}^{P_m}$ and $Cen_{right}^{P_n}$ and (d_{32}) between $Cen_{left}^{P_{m-1}}$ and $Cen_{right}^{P_n}$ is calculated using the following equation:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{6.2}$$

where d is the calculated distance between two points and (x_1, y_1) and (x_2, y_2) are the coordinates of two points.

If $d_{12} > d_{32}$, then the uppermost bony feature in the right column is missing. Then the slopes {[$M_{slope}^{P_z}$], z = 2, . . . , P_{m-1}} are calculated for each corresponding centroid pair from {[Cen_{left}^{M}] and [Cen_{right}^{N}], where $M = N = P_{m-1}$}, using equation (6.1).

If $d_{32} > d_{12}$, then the uppermost bony feature is present on either side and the bony feature is missing somewhere in the middle of the right column (Fig 6.6d). In that case, a missing pair will be present in between the thoracic transverse process and will affect the angle calculation. So, the particular image cannot be used for angle calculation.

b) **If the bony feature in the left column is missing, i.e. $\boldsymbol{P_n - P_m = 1}$** (Fig 6.6c)

To detect the missing top bony feature, the distance (d_{21}) between the $Cen_{right}^{P_n}$ and $Cen_{left}^{P_m}$ and (d_{23}) between $Cen_{right}^{P_{n-1}}$ and $Cen_{left}^{P_m}$ is calculated using equation (6.1).

If $d_{21} > d_{23}$, then the uppermost bony feature in the left column is missing. Then the slopes {[$M_{slope}^{P_x}$], x = 2, . . . , P_{n-1}} are calculated for each corresponding centroid pair from {[Cen_{left}^{M}] and [Cen_{right}^{N}], where $M = N = P_{n-1}$} using equation (6.1).

If d_{23}> d_{21}, then the uppermost bony feature is present on either side and the bony feature is missing from somewhere in the middle of the left column. In that case, a missing pair will be there in between the thoracic transverse process, and it will affect the angle calculation. So, the particular image cannot be used for angle calculation.

iii) **Case III - when more than one bony feature is missing in either side of the spinous column profile, i.e. $P_m \neq P_n$ and $|P_m - P_n| > 1$** (Fig 6.6e)

No slope is calculated, as extra bony features are present or missing in either side of the spinous column profile.

4. **Main thoracic angle calculation**: {[$M_{slope}^{P_x}$], x = 2, . . . , h, h is the total no. of calculated angles} are divided in two sets {[M_{slope}^{b}], b = 2, . . . , $\frac{P_y}{2}$} and {[M_{slope}^{t}], x = $\frac{P_y}{2} + 1$, . . . , P_y}. The highest two slopes $M_{slope}^{b_{high}}$ *and* $M_{slope}^{t_{high}}$ are selected from [M_{slope}^{b}] and [M_{slope}^{t}] and the final thoracic angle $M^{Thoracic}$ is calculated from them using the following equation:

$$M^{Thoracic} = \tan^{-1}\left(\frac{M_{slope}^{b_{high}} - M_{slope}^{t_{high}}}{1 + M_{slope}^{b_{high}} \cdot M_{slope}^{t_{high}}}\right) \quad (6.3)$$

Where $M^{Thoracic}$ is the calculated angle between the two most tilted thoracic transverse processes and $M_{slope}^{b_{high}}$ and $M_{slope}^{t_{high}}$ are the two most tilted slopes.

Algorithm 2: Proposed algorithm for main thoracic angle calculation.

Initialization: $C_{Contour}^{T}$ denotes the detected thoracic transverse process from segmented spine images, the two bottom thoracic bony features are denoted as T12 level. P_m *and* P_n are the total number of contours present in the left and right column of the spinous column profile respectively.

Output: The thoracic angle $M^{Thoracic}$.

for *M* = 1, . . . , P_m

for *N* = 1, . . . , P_n

Calculate centroid [Cen_{left}^{M}] and [Cen_{right}^{N}] for each contour of the left and right column.

end for

end for

if $P_m = P_n$

for P_x= 2, …. , P_m

Obtain Slopes [$M_{slope}^{P_x}$] for each corresponding centroid pairs of {[Cen_{left}^{M}] and [Cen_{right}^{N}], where *M*= *N*}

end for

elseif $P_m \neq P_n$

if $P_m - P_n = 1$

Calculate the distance (d_{12}) between the $Cen_{left}^{P_m}$ and $Cen_{right}^{P_n}$ and (d_{32}) between $Cen_{left}^{P_{m-1}}$ and $Cen_{right}^{P_n}$

if $d_{12} > d_{32}$

Discard the upper-most bony feature in the left column

for P_x = 2, . . . , P_{m-1}

Obtain slopes $[M_{slope}^{P_x}]$ for each corresponding centroid pair from {$[Cen_{left}^{M}]$ and $[Cen_{right}^{N}]$, where M= N = P_{m-1}}

end for

elseif $d_{32} > d_{12}$

Do not calculate any angle

end if

elseif $P_n - P_m = 1$

Calculate the distance (d_{21}) between the $Cen_{right}^{P_n}$ and $Cen_{left}^{P_m}$ and (d_{23}) between $Cen_{right}^{P_{n-1}}$ and $Cen_{left}^{P_m}$

if $d_{21} > d_{23}$

Discard the upper-most bony feature in the right column

for P_x = 2, . . . , P_{n-1}

Obtain slopes $[M_{slope}^{P_x}]$ for each corresponding centroid pair from {$[Cen_{left}^{M}]$ and $[Cen_{right}^{N}]$, where $M = N = P_{n-1}$}

end for

elseif $d_{23} > d_{21}$

Do not calculate any angle

end if

elseif $|P_m - P_n| > 1$

Do not calculate any angle

end if

end if

Divide{ $[M_{slope,}^{P_x}]$ x = 2, . . . , h, h is the total no of calculated angles} in two sets {$[M_{slope}^{b}]$, b = 2, . . . , $\frac{P_y}{2}$} and {$[M_{slope}^{t}]$, x = $\frac{P_y}{2} + 1, \ldots, P_y$}.

Select the highest two slopes $M_{slope}^{b_{high}}$ *and* $M_{slope}^{t_{high}}$ from $[M_{slope}^{b}]$ and $[M_{slope}^{t}]$

Calculate thoracic angle $M^{Thoracic}$

6.4.3 Measurement of thoraco-lumbar angle

Let $C^{L}_{Contour}$ denote the contours of the lumbar lumps detected on the segmented spine image, where L represents the total number of lumbar bony features detected below the T12 region in a single segmented spine image. Six, sometimes five (depending on scan length), lumbar regions should be identified below the T12 level. Sometimes, the conjoined lumbar features give less than five feature counts (Fig 6.7); then the algorithm

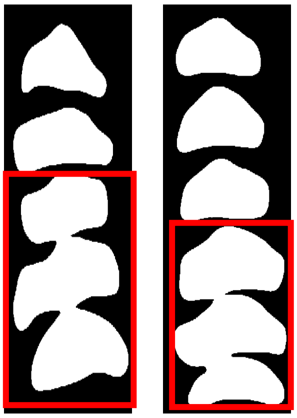


Fig 6.7. Conjoint LBF

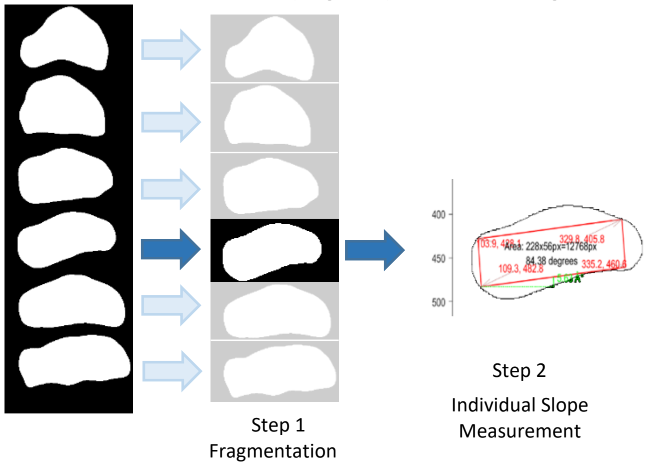


Fig 6.8. Measurement of thoraco-lumbar angle using largest inscribed rectangle method

fails to differentiate the conjoint features. Thus, such images cannot be taken into consideration for angle calculation. The procedure to calculate the thoraco-lumbar angle from the contours is described in the following sequence of operations (Fig 6.8):

1. **Separation of individual lumbar regions:** All the L contours of $\{[C^{L}_{Contour}]$, L= 0, ….., v; $v \geq 5$, v is the total lumbar bony feature detected$\}$are separated into different images to calculate the individual slope of each contour. Let each contour be denoted by $I^1, \ldots, I^L$.
2. **Calculation of largest inscribed rectangle inside lumbar regions and corresponding slopes:**
   i. Each contour $\{I^1, \ldots, I^L\}$ is rotated to detect the arbitrarily oriented rectangle.
   ii. Each right white pixel that is next to a left black pixel is searched for the largest rectangle. If the current rectangle is larger than the previous largest rectangle, then the current one is considered to be the largest rectangle inside the contour.
   iii. If all left border pixels are checked, the image is rotated by the rotation step = RS. The range of rotation is $0 < RS \leq 90°$. The default rotation step is set as 5°. If RS > (Current Angle - Previous Angle), then the image is rotated once.

iv. Iteration is needed to refine the rotation steps to find the largest possible rectangle. Iteration is set at 1. If First Angle = Last Angle, then no rotation happens. After one iteration, the rotation step is divided by 2 and the largest rectangles are checked. The iteration continues till the rotation step reaches 0.01°.

v. After all rotations up to nearly 90° are done, the largest rectangle and its slope are considered the results.

3. **Calculation of most titled slope and thoraco-lumbar angle:**

i. The slopes for all the contours $\{I^1, \dots, I^L\}$ are calculated as { $[M^L_{slope}]$, L= L= 0, ….., v, v is the total lumbar bony feature detected}.

ii. The highest slope $M^{L_{high}}_{slope}$ is selected among $[M^L_{slope}]$.

iii. The thoraco-lumbar angle $M^{Thoracic}$ is calculated using $M^{b_{high}}_{slope}$ and $M^{L_{high}}_{slope}$ using equation (6.3).

Algorithm 3: Proposed algorithm for thoraco-lumbar angle calculation.

Initialization: $\{[C^L_{Contour}]$, L= 0, ….., v, v is the total lumbar bony feature detected} are the identified lumbar bony features,

Output: Thoraco-lumbar angle $M^{Thoraco-lumbar}$

for L= 1, …., v

Calculate the slope $[M^L_{slope}]$ of the largest inscribed rectangle inside the contour.

end for

Select the highest slope $M^{L_{high}}_{slope}$

Obtain thoraco-lumbar angle from $M^{b_{high}}_{slope}$ and $M^{L_{high}}_{slope}$.

6.5 Experiments

6.5.1 Dataset

A total of 109 images collected from 109 patients (82 females and 27 males) with an average age of 15.6 ± 2.7 years and different degrees of spine deformities were used in this study retrospectively. The experimental procedures involving human subjects were approved by the Institutional Review Board. The subjects gave informed consent to their inclusion in this study as required, and the work adheres to the Declaration of Helsinki.

Scolioscan (Model SCN801, Telefield Medical Imaging Ltd), developed in Hong Kong, was used to generate 3D volume projection image (VPI) using the conventional 3D ultrasound imaging technique. The instrument has a rigid frame with two movable and four fixed supporting boards to aid the patient in maintaining a stable standing posture during diagnosis. It also has two LCD screens and a touch screen that is used by the operator for uploading patient information, setting parameters during the scan, generating VPI images, conducting measurements, and producing reports. Scolioscan also has a proprietary software for scanning, analysing VPI images and measuring the angle of the spine. A volume projection image (VPI) is a volumetric image with an averaged intensity of all voxels within a selected depth along with the front and back directions. Ethical approval of the project was obtained from the Hong Kong Polytechnic University Ethics Committee (Approval number: HSEARS20180906005). The study received human subject ethical approvals from The Hong Kong Polytechnic University.

6.5.2 Implementation Details

Software: MATLAB (version: R2021b)

Machine: GPU laptop equipped with NVIDIA GeForce RTX 2060.

6.5.3 Analysis of the performance of CPI-SLO with various segmentation methods

Firstly, the three segmentation techniques, i.e. (a) SIU-Net, (b) LDS U-Net and (c) U-Net, as the baseline segmentation technique, were used to provide outputs with respect to the identification of TBFs and LBFs. The output segmented images were then used with CPI-SLO and the performances were assessed.

Secondly, the performance of CPI-SLO was statistically compared with manual UCA measurement to establish the reliability and validity of the proposed method with respect to the current practice. For evaluating the reliability, conventional statistical measures of range, mean and standard deviations were compared. For assessing the validity, three measures were used: (a) a linear regression and correlation analysis with zero intercepts ($y = kx; k = regression\ coefficient$) was implemented to describe the relationship between the main thoracic and thoraco-lumbar manual and automatic UCA; a correlation coefficient of 0.25–0.50 indicates a poor correlation, 0.50–0.75 indicates moderate to good correlation, and 0.75–1.00 indicates very well to excellent correlation (Dawson & Trapp 2004); (b) the Pearson correlation coefficient (r) was calculated to check the reliability of automatic UCA; the level of significance is accepted at $p < 0.05$; (c) Bland and Altman's method of differences (Bland & Altman 1986) was also applied to inspect the agreement between manual and automatic UCA for both main thoracic and thoraco-lumbar angles.

6.6 Results

6.6.1 Performance of CPI-SLO compared with conventional manual UCA measurement

Multiple statistical analyses are done to determine the reliability and validity of the proposed automatic CPI-SLO method as compared with the prevalent manual ultrasound curvature angle (UCA) method.

Table 6.3: Mean range of UCA angle

METHOD	SPREAD (RANGE)		MEAN (μ) AND STD DEV	
	THORACIC	THORACO-LUMBAR	THORACIC	THORACO-LUMBAR
U-Net	5.01° to 58.20°	5.24° to 49.21°	25.63° ± 12.57°	22.4° ± 8.10°
LDS U-Net	5.24° to 50.61°	7.31° to 56.47°	25.29° ± 12.79°	22.50° ± 8.92°
SIU-Net	7.31° to 56.47°	6.75° to 49.91°	25.10° ± 10.73°	21.46° ± 8.82°
Manual	7.90° to 59.90°	10.40° to 47.20°	25.90° ± 11.20°	23.91° ± 8.32°

6.6.1.1 Reliability

For the 109 patients, the range of the UCA measured using the manual method is 7.9–59.9° (μ 25.9 ± 11.2) for the main thoracic region and 10.4–47.2° (μ 23.9 ± 8.3) for the thoraco-lumbar region. Table 6.3 shows that the UCA angle range measured by the SIU-Net automatic method is 7.31–56.47° (μ 25.10± 10.73) for the main thoracic region and 6.75–49.91°(μ 21.46 ± 8.82) for the thoraco-lumbar region, which are closer to the manual measurement range compared to those calculated using the LDS U-Net and U-Net method.

Table 6.4 shows that, using the SIU-Net method, 92% (86 of 94 angles) of the main thoracic automatic UCA and 91% (78 of 86 angles) of thoraco-lumbar automatic UCA are within ±5° difference from the corresponding manually measured UCA angles, which is better than the results produced using the other two segmentation methods. Hence, the SIU-Net + CPI-SLO method has lesser spread, closer mean and standard deviation and is most consistent with the manual UCA method compared to other segmentation methods. Thus, it shows the highest reliability for future adaptations.

6.6.1.2 Validity

Main thoracic and thoraco-lumbar automatic UCA using SIU-Net shows very good linear correlations with manual UCA for main thoracic ($R^2 = 0.885$) and thoraco-lumbar curves ($R^2 = 0.801$) (Fig 6.9). The scaling factors between manual and automatic UCA obtained from the linear regression analysis with zero intercepts ($y = kx$, k= regression coefficient) are 1.01 and 0.933 for the main thoracic and thoraco-lumbar angles, respectively.

Table 6.4: Comparison of accuracy between automatic UCAs and manual UCA

METHOD	% OF IMAGES WITH DIFFERENCES WITHIN ±5° FROM RESP. MANUAL UCA	
	THORACIC	THORACO-LUMBAR
U-Net	75	78
LDS U-Net	87	84
SIU-Net	92	91

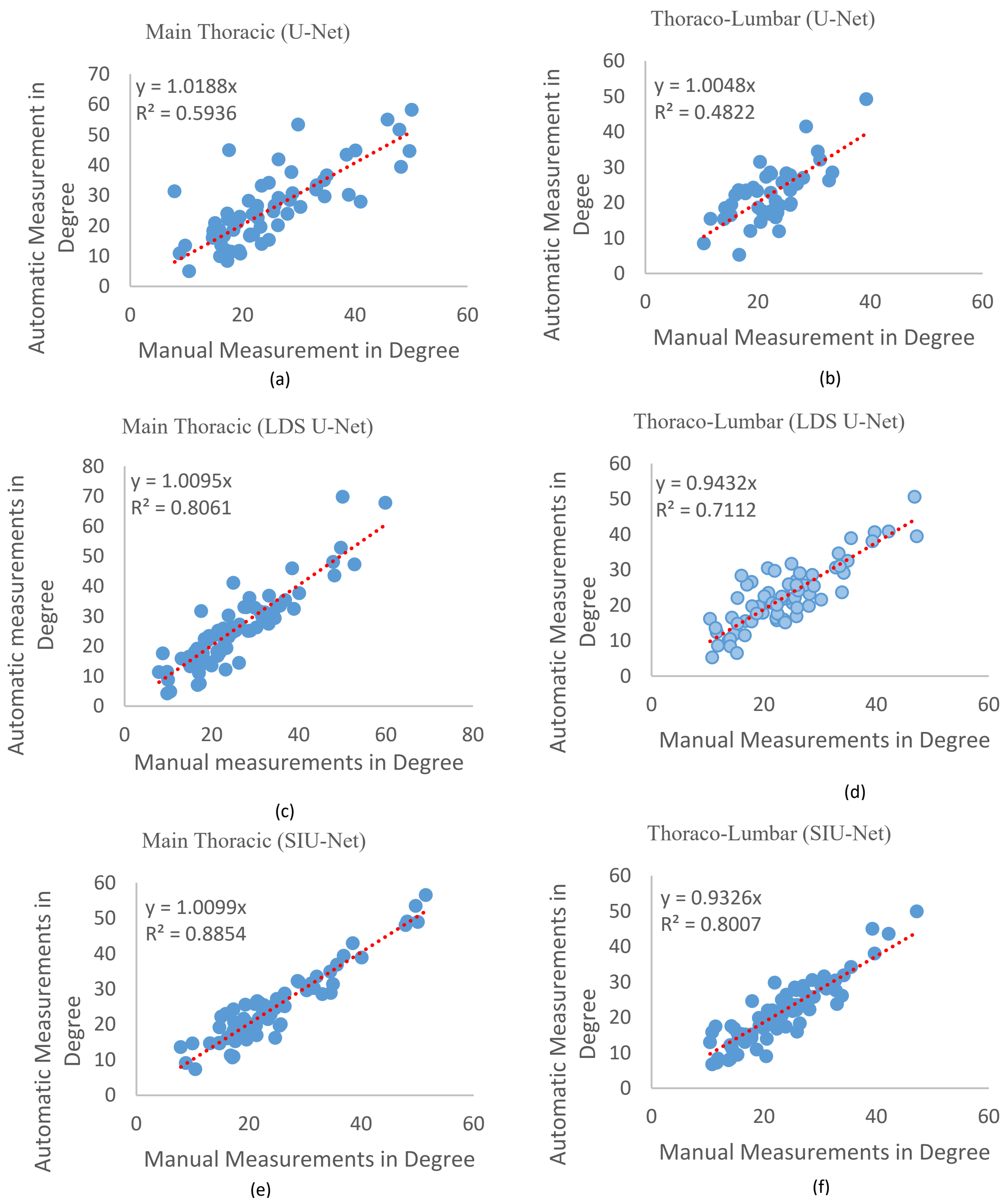


Fig 6.9. Correlations (R^2) and regression equations between (a) Manual-Unet, (b) Manual-LDS U-Net & (c) Manual-SIU-Net and Thoraco-lumbar angles (e) Manual-Unet, (f) Manual-LDS U-Net & (g) Manual-SIU-Net

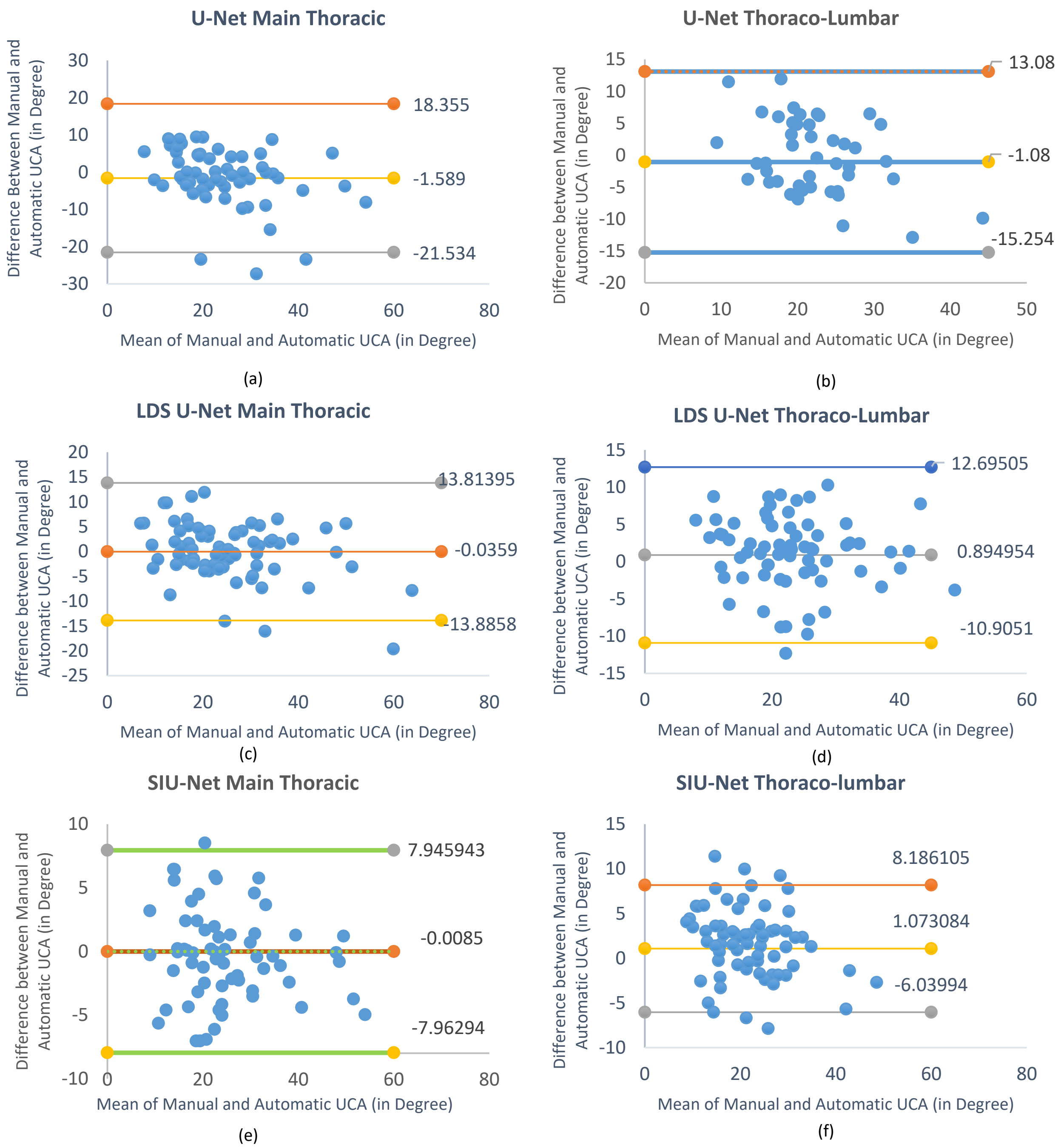


Fig 6.10. Bland–Altman plots which demonstrates the differences between the (a) Manual-Unet, (b) Manual-LDS U-Net & (c) Manual-SIU Net for main thoracic angles and (e) Manual-Unet, (f) Manual-LDS U-Net & (g) Manual- SIU Net for Thoraco-lumbar angles. The central line represents the bias and the dotted lines represent the 95% limits of agreement.

On the other hand, the linear relationship between automatic UCA measured using LDS U-Net (for main thoracic: $R^2 = 0.806$, scaling factor 1.01; for thoraco-lumbar: $R^2 = 0.711$, scaling factor 0.943) and that using U-Net (for main thoracic: $R^2 = 0.594$, scaling factor 1.02; for thoraco-lumbar: $R^2 = 0.482$, scaling factor 1.005) are not as high as that using SIU-Net.

Furthermore, the reliability of the automatic UCA measured using all three segmentation methods is assessed by the Pearson correlation coefficient (r). The results in Table 6.5 reveal a significant linear correlation between the SIU-Net automatic and manual UCA ($r = 0.941$ for main thoracic and $r = 0.899$ for thoraco-lumbar, $p < 0.001$).

Bland and Altman's plot is employed to graphically evaluate the agreement between the two measurements. It is a scatter plot where the difference between the manual and automatic UCA measurements is plotted against their mean values. It provides two main pieces of information, namely the average of all the differences (bias) and the 95% limits of agreement (+1.96 of standard deviation and -1.96 of standard deviation). Fig 6.10 shows the Bland and Altman's plot along with the values of bias, 95% upper and lower limits of agreement between all the automatic and manual measurements of UCA angles. In the case of SIU-Net and CPI-SLO method, the bias, 95% upper and lower limits of agreements are (-0.0085, 7.946 and -7.963) for main thoracic curves and (1.073, 8.186 and -6.039) for thoraco-lumbar angles, respectively, which is better than the other two methods, showing the best agreement with the manual UCA measurement.

Table 6.5: Correlation between automatic UCAs and manual UCA

Method	Pearson correlation coefficient (r)	
	Thoracic	Thoraco-Lumbar
U-Net automatic vs. manual	0.773	0.697
LDS automatic vs. manual	0.898	0.845
SIU automatic vs. manual	0.941	0.899

These results show that the automatic UCA obtained by SIU-Net + CPI-SLO method has a very good correlation with manual UCA. Hence, the proposed method is suitable for implementation in the automatic scoliosis diagnosis process.

6.7 Discussion & analysis

Adolescence idiopathic scoliosis (AIS) requires periodic monitoring of patients for continuous observation of small changes in spine curvature. The application of ultrasound imaging in the assessment of scoliosis could reduce radiation exposure risk for young adults. Previously, spinous process angles (SPA) were measured from the Scolioscan ultrasound images using the VPI-SP method, where the spinal column profile was taken as an anatomical reference (Zheng, Lee, et al. 2016). However, this leads to an underestimation of the traditionally used Cobb angles. Measurement of the ultrasound curvature angle (UCA), on the other hand, is an alternative and reliable technique of scoliosis assessment, where more lateral features are used for angle measurement (Lee et al. 2021). The manual measurement of UCA is currently in practice, which requires a lot of time and human judgment. The objective of this work is to automate the UCA measurement and subsequently reduce the human intervention to fasten the process.

Unlike the SPA measurement technique, the manual measurement of UCA (Lee et al. 2021) requires the visualization of the lateral features of the spine to evaluate coronal curvature. Therefore, an alternative VPI method, which uses different depth profiles extracted from the sagittal ultrasound images, is used for input coronal images. Nine 2D VPI images of different depth profiles are extracted from one 3D image for all patients (Zheng, Lee, et al. 2016). Then, the layer that best reveals the tilted lateral features of the spine is selected by two raters. In the manual UCA process, the evaluators need to locate the proper points for line placements on the ultrasound images. The identification of the feature below and above the T12 level is dependent on the expertise and judgment of the evaluators. The first key element in the automatic UCA measurement is the identification and segmentation of the lateral bony features (thoracic and lumbar), as they play a pivotal role in angle calculation. In the proposed automatic UCA measurement technique, 109 ultrasound spine images are segmented using a novel multi-scale feature fusion skip-inception U-Net method and are proven to give the best performance against seven other contemporary segmentation techniques.

In the manual UCA process, for the line drawing of the vertebra structure, the uppermost and lowermost tilted region of the curve must be selected. Among all the thoracic bony feature pairs above T12 level, lines are drawn through the centre of the most tilted pairs, and the tilted regions are selected manually. Similarly, for the region below the T12 level, the widest lump (lumbar feature) or the most tilted lump is selected manually, and one line is drawn through the centre of that lump. In the proposed automatic process, all the slopes between the bony feature pairs above the T12 level are calculated, and the two most tilted pairs are selected among them for main thoracic angle calculation. In the case of the bony features below the T12 level, the slopes of each feature are calculated and the feature with the highest slope is selected for thoraco-lumbar angle calculation. Therefore, no human involvement is required in automatic angle calculation, thus making the process more reliable.

In this research, the main thoracic and thoraco-lumbar angles are measured automatically from the segmented images by basic U-Net (Ronneberger, Fischer & Brox 2015), LDS U-Net (Chapter IV), and SIU-Net (Chapter V). The feasibility of the automatic UCA, i.e. SIU-Net + CPI-SLO, is demonstrated by the significant correlation (scalability = 1.02, R = 0.941, $p < 0.001$ for main thoracic; scalability = 1.01, r = 0.899, $p < 0.001$ for thoraco-lumbar) between the results obtained by the manual and automatic UCA methods for 109 patients with varying degrees of scoliosis. The results ascertain that the proposed automatic method can facilitate reliable scoliosis assessment without the need for human intervention and variations in manual measurements and is significantly faster than the manual method since it is done by a processor. Moreover, the manual UCA measurement has a very high correlation with the gold standard X-ray Cobb angle (Lee et al. 2021). Hence, it can be concluded that this automatic SIU-Net + CPI-SLO method shows good agreement with the traditional method. Therefore, the proposed SIU-Net + CPI-SLO method can enable automatic diagnosis using 3D ultrasound imaging for scoliosis assessment.

The proposed automatic UCA measurement is limited by a few factors. It must be noted that, given the inherent limitations of ultrasound imaging modality, all the bony features of some ultrasound spine images are not identified properly. In those cases, the features are either conjoined or missing. For thoracic bony features, if one top TBF is missing in either the left or right side of the spinous column profile, the extra TBF, which does not make a pair, is not included and the slopes of other pairs are taken into consideration. But if more than one TBF from the top is missing on either side, pairing will be not viable; therefore, the images cannot be included for further angle calculation. Similarly, the image would be discounted from measurement if one or more TBFs are missing in the

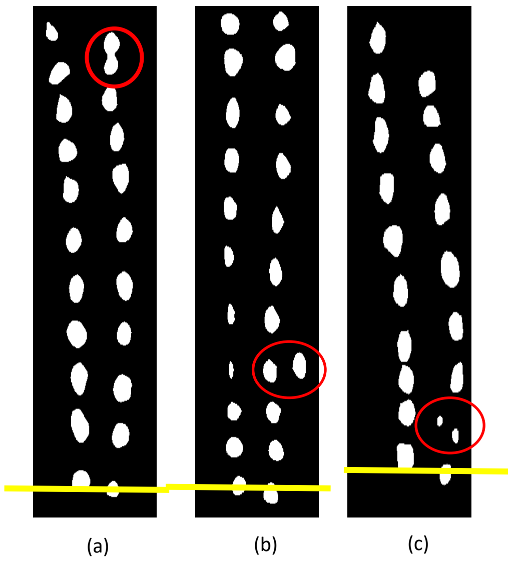


Fig 6.11. Special cases for main thoracic UCA measurement (a) upper right two bony features are conjoint, extra bony feature is present in right side (b) $P_n - P_m = 1$ & (c) $P_m = P_n$

middle and conjoined features are present, as forming TBF pairs will be erroneous. Also, it is observed that there are three special cases where Algorithm 1 (for main thoracic angle calculation) gives incorrect outputs, as shown in Fig 6.11: (a) Top right TBFs are conjoint; (b) in the right column, an extra TBF is present; hence, the algorithm incorrectly assigns them under the $P_n - P_m = 1$ situation; (c) one right bony feature is broken, so the algorithm wrongly assumes it to be under the $P_n = P_m$ situation.

In the case of lumbar bony features, a conjoined feature may result in incorrect calculation. Therefore, if an image with less than five features is present, the image will be considered to have either missing or conjoined bony features and, therefore, will not be included for thoraco-lumbar angle measurement. Further, during our validation, it was noted that there are cases where Algorithm 2 (for thoraco-lumbar angle calculation) would not give a correct output if the lumber is fragmented, increasing the LBF count, as shown in Fig 6.12. Rescanning will be required for the images with missing or conjoined bony features for proper feature identification. Further improvement in the proposed method could be done by giving real-time feedback to the scanner so that higher-quality images can be used to calculate the UCA automatically.

6.8 Conclusion

In this research, a novel automatic technique is presented to measure ultrasound curvature angle from spine VPI images by utilizing a selection of best images by experts, ultrasound spine image segmentation, and a newly developed angle measurement technique. The performance evaluation shows that this automatic method can produce

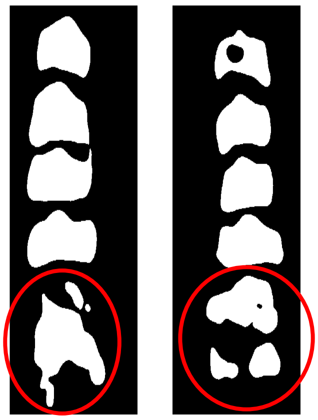


Fig 6.12. Special cases for Thoraco-lumbar UCA measurement: (a) & (b) lower most LBF is broken and is counted as 7

results comparable to manual measurements of UCA. The outcome of this automatic spine curvature measurement method is demonstrated by the significant correlation (for main thoracic angle: $R^2 = 0.885$, scalability = 1.01, r = 0.941, p < 0.001; for thoraco-lumbar angle: $R^2 = 0.801$, scalability = 0.933, r = 0.899, p < 0.001) between the results obtained by manual and automatic curvature assessment methods among patients with varying degrees of scoliosis. The advantage of the automatic method is that it overcomes subjective manual annotation, reduces the human error in measurements, and improves the scalability of the system. Also, frequent monitoring of curve progression for AIS patients is safely possible using this technique, as no harmful radiation is emitted in this imaging process. Hence, the automatic UCA measurement technique using 3D ultrasound imaging modality is robust, fast, economic, and suitable for mass diagnosis of scoliosis measurement in the adolescent population.

CHAPTER VII: CONCLUSIONS & FUTURE WORK

7.1 Research summary

This research focuses on developing a suitable technique for the automatic measurement of spine curvature angle using ultrasound images. In the current practice, the accuracy of scoliosis diagnosis is largely dependent on the awareness and experience of the observer and the quality of images. It is difficult for the experts to make accurate measurements because of the large anatomical variations among patients from different age groups and could result in many inter- or intra-observer errors. Manual assessment technique is also quite limited in terms of scalability due to the need for skilled experts. Further, there are two drawbacks specifically associated with ultrasound images:

(a) Unlike the clear images obtained from radiative modalities, an ultrasound image suffers from scan and speckle noise. As speckle noise behaves as information, it makes the clinical data hard to analyse.

(b) Ultrasound images are low contrast images because the speed of sound varies depending on the tissues, thereby making it tough to differentiate between fat- and water-based tissues. This makes it difficult to demarcate all the necessary feature information from the image.

This research had to overcome unique challenges in developing the technique for automatic diagnosis of scoliosis, such as (a) dealing with spine images which have very high variability in terms of shape, size, and location of bony features and (b) using images that are taken using ultrasound which are of low contrast and plagued with scan and speckle noises.

The overall structure of the research is presented in Chapter I, Fig. 1.4. The major steps in the automatic measurement of spine curvature angle using ultrasound images can be summarized as follows:

1. Generation and collection of data: This research focuses on using ultrasound images for scoliosis assessment, as this imaging modality is safe, economic, and scalable. Chapter III describes the data collection process in detail.
2. Segmentation: The input spine image data is simplified and/or changed into something more meaningful and easier to analyse for scoliosis diagnosis. Chapter IV and V detail the segmentation methodology.
3. Automatic angle calculation: The segmented output images, which identify the relevant ROIs, are used to estimate the spine curvature angle automatically. This

methodology is discussed in Chapter VI.

7.2 Research contributions

The key contributions of this research are as follows:

1. Two novel segmentation techniques are developed for ultrasound spine imaging:

   The proper identification and segmentation of thoracic and lumbar bony features are of prime importance. Single ROI segmentation techniques such as basic U-Net are incapable of handling the high variability in locations, sizes, and shapes of multiple ROIs (TBFs and LBFs) in a noisy ultrasound image. To this end, new segmentation architectures are developed using two different approaches:

   a) Light-convolution dense selection (LDS) U-Net: The design employs a smart selection gate to filter noise, multi-scale skip pathways for effective feature propagation, and light dense block to lower the overall computational requirement.

      The performance of LDS U-Net is compared with the basic U-Net, MultiResUNet, and Attention U-Net. Firstly, LDS U-Net quantitatively outperforms the other models with 0.7415 Jaccard index, 0.8694 Dice coefficient, 0.8885 F1 score, and 0.9592-pixel accuracy. Secondly, it can identify the thoracic and lumbar bony features most consistently. Thirdly, in handling noisy ultrasound images with high variability, the LDS U-Net is successful in correctly segmenting 76.6% of total images with complete identification of all the 26 bony features, which is more consistent than other models. Finally, its lightweight design translates to a much lower parameter requirement and a significantly smaller memory footprint. The details of LDS U-Net are presented in Chapter IV.

   b) Multi-scale feature fusion skip-inception U-Net (SIU-Net): As an alternative approach to effective segmentation, the segmentation architecture is re-designed by focusing more on tackling variability and noise without primarily focusing on the low computational requirement. This resulted in a new segmentation architecture design called SIU-Net. It is designed using improvised inception blocks for variable feature extraction and novel dense skip pathways for improved feature fusion and enhanced segmentation output.

      SIU-Net shows a promising segmentation output compared to baseline network U-Net and advanced networks, such as MultiResUNet and Unet++. The performance of SIU-Net is evaluated quantitatively using three popular indices, i.e. Jaccard index, Dice coefficient, and histogram Euclidean distance. SIU-Net outperforms all the other popular segmentation techniques, with the maximum Jaccard index (0.78) and Dice coefficient (0.883) and minimum histogram Euclidean distance (0.011). In handling noisy ultrasound images with high variability, the SIU-Net is successful in correctly segmenting 81.2% of total images with complete identification of all the 26 bony features, which is more consistent than any of the other models. SIU-Net is detailed in Chapter V.

Table 7.1: Summary of performance of proposed segmentation methods

Index	LDS U-NET	SIU-NET
Dice Score	0.8694	0.883
% TBF Accuracy	79.7%	86.0%
% LBF Accuracy	72.3%	75.7%
Average % BF Accuracy	76.6%	81.2%

Comparison between LDS and SIU-Net (Table 7.1) shows that SIU-Net can identify the TBFs more effectively, at 86%, while showing improvement in the detection of LBFs as well. Hence, for angle calculation, SIU-Net is found to be the most suitable architecture for the generation of segmented images and identifying the relevant ROIs for angle calculation.

2. Automatic UCA calculation:

A novel automatic technique, the centroid pairing and inscribed rectangle slope (CPI-SLO) method can measure ultrasound curvature angle automatically from spine VPI images. For this technique, three algorithms have been proposed for UCA calculation:

a) Algorithm 1: Proposed algorithm for binary to RGB conversion of segmented spine ultrasound images.
b) Algorithm 2: Proposed algorithm for main thoracic angle calculation.
c) Algorithm 3: Proposed algorithm for thoraco-lumbar angle calculation.

The performance evaluation of the automatic UCA measurement technique shows that this automatic method can produce results comparable to manual measurements of UCA. The advantage of the automatic method is that it overcomes subjective manual annotation, reduces the human error in measurements, and improves the scalability of the system, as reported in Chapter VI.

The feasibility of this automatic spine curvature measurement method is demonstrated by the significant correlation (for main thoracic angle: $R^2 = 0.885$, scalability = 1.01, r = 0.941, p < 0.001; for thoraco-lumbar angle: $R^2 = 0.801$, scalability = 0.933, r= 0.899, p < 0.001) between the results obtained by manual and automatic curvature assessment methods among patients with scoliosis of differing degrees. The results ascertain that this proposed automatic method can provide reliable scoliosis assessment without the need for human intervention and variations in manual measurements and is significantly faster than the manual method since it is performed by a processor. Moreover, manual measurement shows a very high correlation with the gold standard X-ray Cobb angle. Hence, this automatic method shows good agreement with the traditional method.

The societal impact of this research are as follows:

1. Safe: This research develops a radiation-free imaging system that eliminates the risk of cancer among the adolescent population.
2. Economic: Ultrasound imaging modality is more affordable compared to other imaging modalities.
3. Scalable: Automatic and fast detection is possible using this technique as it uses Artificial Intelligence for scoliosis assessment.
4. Reliable: The technique is highly reliable and innovative and is comparable to the current established practice for scoliosis assessment.

7.3 Recommendations for future studies

Some of the possible recommendations regarding future work related to the current study are as follows:

- The automatic ultrasound curve angle (UCA) measurement system can be used as a mobile application, and if the scanned spine image is fed to the application, it can provide the curvature angle. In this way, the periodic monitoring of spine curvature will be a very easy and fast process.
- The image dataset has a low sample size. Though the segmentation and angle calculation results are promising, the generalization of performance needs to be tested in larger and more diverse patient populations in the future.
- For almost 20% of the images, segmentation could not be done properly. In such cases, the proposed method either fails to differentiate the bony features or misses out on

some features completely. For this, the authors postulate that rescanning will be required for the images with missing or conjoined bony features for proper feature identification. Going forward, automatic segmentation can be combined with the scanning process so that the scanner can gets real-time feedback to conduct a targeted re-scan of the unclear areas, if necessary.

- Image quality could be further improved by developing a flexible or small probe with good penetration to overcome the limitations of ultrasound imaging. Using different ultrasound machines, imaging protocols, and different sonographers will help improve the richness of input information for future studies.

REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G. & Isard, M. 2016, 'Tensorflow: A system for large-scale machine learning', *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, pp. 265-83.

Abbasian Ardakani, A., Bitarafan-Rajabi, A., Mohammadzadeh, A., Mohammadi, A., Riazi, R., Abolghasemi, J., Homayoun Jafari, A. & Bagher Shiran, M. 2019, 'A hybrid multilayer filtering approach for thyroid nodule segmentation on ultrasound images', *Journal of Ultrasound in Medicine*, vol. 38, no. 3, pp. 629-40.

Amiri, M., Brooks, R., Behboodi, B. & Rivaz, H. 2020, 'Two-stage ultrasound image segmentation using U-Net and test time augmentation', *International journal of computer assisted radiology and surgery*, vol. 15, no. 6, pp. 981-8.

Badrinarayanan, V., Kendall, A. & Cipolla, R. 2017, 'Segnet: A deep convolutional encoder-decoder architecture for image segmentation', *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481-95.

Banerjee, S., Ling, S.H., Lyu, J., Su, S. & Zheng, Y.-P. 2020, 'Automatic Segmentation of 3D Ultrasound Spine Curvature Using Convolutional Neural Network', *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, pp. 2039-42.

Bengio, Y. 2012, 'Deep learning of representations for unsupervised and transfer learning', *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, pp. 17-36.

Bi, L., Kim, J., Kumar, A., Fulham, M. & Feng, D. 2017, 'Stacked fully convolutional networks with multi-channel learning: application to medical image segmentation', *The Visual Computer*, vol. 33, no. 6-8, pp. 1061-71.

Bland, J.M. & Altman, D. 1986, 'Statistical methods for assessing agreement between two methods of clinical measurement', *The lancet*, vol. 327, no. 8476, pp. 307-10.

Brink, R.C., Wijdicks, S.P., Tromp, I.N., Schlösser, T.P., Kruyt, M.C., Beek, F.J. & Castelein, R.M. 2018, 'A reliability and validity study for different coronal angles using ultrasound imaging in adolescent idiopathic scoliosis', *The Spine Journal*, vol. 18, no. 6, pp. 979-85.

Campbell Jr, R.M., Smith, M.D., Mayes, T.C., Mangos, J.A., Willey-Courand, D.B., Kose, N., Pinero, R.F., Alder, M.E., Duong, H.L. & Surber, J.L. 2003, 'The characteristics of thoracic insufficiency syndrome associated with fused ribs and congenital scoliosis', *JBJS*, vol. 85, no. 3, pp. 399-408.

Cassar-Pullicino, V. & Eisenstein, S. 2002, 'Imaging in scoliosis: what, why and how?', *Clinical radiology*, vol. 57, no. 7, pp. 543-62.

Cassella, M.C. & Hall, J.E. 1991, 'Current treatment approaches in the nonoperative and operative management of adolescent idiopathic scoliosis', *Physical therapy*, vol. 71, no. 12, pp. 897-909.

Cerrolaza, J.J., Sinclair, M., Li, Y., Gomez, A., Ferrante, E., Matthew, J., Gupta, C., Knight, C.L. & Rueckert, D. 2018, 'Deep learning with ultrasound physics for fetal skull segmentation', *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 564-7.

Chen, H., Zheng, R., Qian, L., Liu, F., Song, S. & Zeng, H. 2021, 'Improvement of 3D Ultrasound Spine Imaging technique using Fast Reconstruction Algorithm', *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*.

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K. & Yuille, A.L. 2017, 'Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs', *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834-48.

Chen, L.-C., Papandreou, G., Schroff, F. & Adam, H. 2017, 'Rethinking atrous convolution for semantic image segmentation', *arXiv preprint arXiv:1706.05587*.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F. & Adam, H. 2018, 'Encoder-decoder with atrous separable convolution for semantic image segmentation', *Proceedings of the European conference on computer vision (ECCV)*, pp. 801-18.

Chen, L., Bentley, P., Mori, K., Misawa, K., Fujiwara, M. & Rueckert, D. 2018, 'DRINet for medical image segmentation', *IEEE transactions on medical imaging*, vol. 37, no. 11, pp. 2453-62.

Chen, L., Tian, Y. & Deng, Y. 2021, 'Neural Network Algorithm-Based Three-Dimensional Ultrasound Evaluation in the Diagnosis of Fetal Spina Bifida', *Scientific Programming*, vol. 2021.

Chen, W., Lou, E.H. & Le, L.H. 2011, 'Using ultrasound imaging to identify landmarks in vertebra models to assess spinal deformity', *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, pp. 8495-8.

Cheng, J.C., Castelein, R.M., Chu, W.C., Danielsson, A.J., Dobbs, M.B., Grivas, T.B., Gurnett, C.A., Luk, K.D., Moreau, A. & Newton, P.O. 2015, 'Adolescent idiopathic scoliosis', *Nature reviews disease primers*, vol. 1, no. 1, pp. 1-21.

Cheung, C.-W.J., Zhou, G.-Q., Law, S.-Y., Lai, K.-L., Jiang, W.-W. & Zheng, Y.-P. 2015, 'Freehand three-dimensional ultrasound system for assessment of scoliosis', *Journal of Orthopaedic Translation*, vol. 3, no. 3, pp. 123-33.

Cheung, C.-W.J., Zhou, G.-Q., Law, S.-Y., Mak, T.-M., Lai, K.-L. & Zheng, Y.-P. 2015, 'Ultrasound volume projection imaging for assessment of scoliosis', *IEEE transactions on medical imaging*, vol. 34, no. 8, pp. 1760-8.

Chollet, F. 2015, 'keras'.

Chollet, F. 2017, 'Xception: Deep learning with depthwise separable convolutions', *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251-8.

Choudhry, M.N., Ahmad, Z. & Verma, R. 2016, 'Adolescent idiopathic scoliosis', *The open orthopaedics journal*, vol. 10, p. 143.

Cobb, J. 1948, 'Outline for the study of scoliosis', *Instr Course Lect AAOS*, vol. 5, pp. 261-75.

Dawson, B. & Trapp, R. 2004, 'Research questions about relationships among variables', *Basic and Clinical Biostatistics. 4th edn. Lange Medical Books/McGraw-Hill*, pp. 190-220.

de Reuver, S., Brink, R.C., Lee, T.T., Zheng, Y.-P., Beek, F.J. & Castelein, R.M. 2021, 'Cross-validation of ultrasound imaging in adolescent idiopathic scoliosis', *European Spine Journal*, vol. 30, no. 3, pp. 628-33.

Deng, Q. & Huang, Q. 2019, 'Detecting Vertebra Landmarks From Ultrasound Image Using Single Shot MultiBox Detector', *2019 6th International Conference on Systems and Informatics (ICSAI)*, IEEE, pp. 757-61.

Denil, M., Shakibi, B., Dinh, L., Ranzato, M.A. & De Freitas, N. 2013, 'Predicting parameters in deep learning', *Advances in neural information processing systems*, vol. 26, pp. 2148-56.

Deschênes, S., Charron, G., Beaudoin, G., Labelle, H., Dubois, J., Miron, M.-C. & Parent, S. 2010, 'Diagnostic imaging of spinal deformities: reducing patients radiation dose with a new slot-scanning X-ray imager', *Spine*, vol. 35, no. 9, pp. 989-94.

Dice, L.R. 1945, 'Measures of the amount of ecologic association between species', *Ecology*, vol. 26, no. 3, pp. 297-302.

Doody, M.M., Lonstein, J.E., Stovall, M., Hacker, D.G., Luckyanov, N., Land, C.E. & Collaborators, U.S.C.S. 2000, 'Breast cancer mortality after diagnostic radiography: findings from the US Scoliosis Cohort Study', *Spine*, vol. 25, no. 16, pp. 2052-63.

Dubousset, J., Charpak, G., Dorion, I., Skalli, W., Lavaste, F., Deguise, J., Kalifa, G. & Ferey, S. 2005, 'A new 2D and 3D imaging approach to musculoskeletal physiology and pathology with low-dose radiation and the standing position: the EOS system', *Bulletin de l'Academie nationale de medecine*, vol. 189, no. 2, pp. 287-97; discussion 97.

Dunn, J., Henrikson, N.B., Morrison, C.C., Blasi, P.R., Nguyen, M. & Lin, J.S. 2018, 'Screening for adolescent idiopathic scoliosis: evidence report and systematic review for the US preventive services task force', *Jama*, vol. 319, no. 2, pp. 173-87.

Fong, D.Y., Cheung, K.M., Wong, Y.-W., Wan, Y.-Y., Lee, C.-F., Lam, T.-P., Cheng, J.C., Ng, B.K. & Luk, K.D. 2015, 'A population-based cohort study of 394,401 children followed for 10 years exhibits sustained effectiveness of scoliosis screening', *The Spine Journal*, vol. 15, no. 5, pp. 825-33.

Gadosey, P.K., Li, Y., Adjei Agyekum, E., Zhang, T., Liu, Z., Yamak, P.T. & Essaf, F. 2020, 'Sd-unet: Stripping down u-net for segmentation of biomedical images on platforms with low computational budgets', *Diagnostics*, vol. 10, no. 2, p. 110.

Girdler, S., Cho, B., Mikhail, C.M., Cheung, Z.B., Maza, N. & Cho, S.K.-W. 2020, 'Emerging techniques in diagnostic imaging for idiopathic scoliosis in children and adolescents: a review of the literature', *World neurosurgery*, vol. 136, pp. 128-35.

Girshick, R., Donahue, J., Darrell, T. & Malik, J. 2014, 'Rich feature hierarchies for accurate object detection and semantic segmentation', *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 580-7.

Gstoettner, M., Sekyra, K., Walochnik, N., Winter, P., Wachter, R. & Bach, C.M. 2007, 'Inter-and intraobserver reliability assessment of the Cobb angle: manual versus digital measurement tools', *European Spine Journal*, vol. 16, no. 10, pp. 1587-92.

Guan, S., Khan, A.A., Sikdar, S. & Chitnis, P.V. 2019, 'Fully Dense UNet for 2-D Sparse Photoacoustic Tomography Artifact Removal', *IEEE journal of biomedical and health informatics*, vol. 24, no. 2, pp. 568-76.

Guo, J., Li, Y., Lin, W., Chen, Y. & Li, J. 2018, 'Network decoupling: From regular to depthwise separable convolutions', *arXiv preprint arXiv:1808.05517*.

Han, M., Bao, Y., Sun, Z., Wen, S., Xia, L., Zhao, J., Du, J. & Yan, Z. 2019, 'Automatic segmentation of human placenta images with U-Net', *IEEE Access*, vol. 7, pp. 180083-92.

He, K., Zhang, X., Ren, S. & Sun, J. 2016, 'Deep residual learning for image recognition', *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-8.

Herzenberg, J.E., Waanders, N.A., Closkey, R.F., Schultz, A.B. & Hensinger, R.N. 1990, 'Cobb angle versus spinous process angle in adolescent idiopathic scoliosis. The relationship of the anterior and posterior deformities', *Spine*, vol. 15, no. 9, pp. 874-9.

Hoffman, D.A., Lonstein, J.E., Morin, M.M., Visscher, W., Harris III, B.S. & Boice Jr, J.D. 1989, 'Breast cancer in women with scoliosis exposed to multiple diagnostic x rays', *JNCI: Journal of the National Cancer Institute*, vol. 81, no. 17, pp. 1307-12.

Hong, R., Cheng, W.-H., Yamasaki, T., Wang, M. & Ngo, C.-W. 2018, *Advances in Multimedia Information Processing–PCM 2018: 19th Pacific-Rim Conference on Multimedia, Hefei, China, September 21-22, 2018, Proceedings, Part III*, vol. 11166, Springer.

Horne, J.P., Flannery, R. & Usman, S. 2014, 'Adolescent idiopathic scoliosis: diagnosis and management', *American family physician*, vol. 89, no. 3, pp. 193-8.

Horng, M.-H., Kuok, C.-P., Fu, M.-J., Lin, C.-J. & Sun, Y.-N. 2019, 'Cobb Angle Measurement of Spine from X-Ray Images Using Convolutional Neural Network', *Computational and Mathematical Methods in Medicine*, vol. 2019.

Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. & Adam, H. 2017, 'Mobilenets: Efficient convolutional neural networks for mobile vision applications', *arXiv preprint arXiv:1704.04861*.

Hu, R., Dollár, P., He, K. & Darrell, T. 2017, 'Girshick RLearning to Segment Every Thing', *Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy*, pp. 22-7.

Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K.Q. 2017, 'Densely connected convolutional networks', *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700-8.

Huang, H., Xu, H., Wang, X. & Silamu, W. 2015, 'Maximum F1-score discriminative training criterion for automatic mispronunciation detection', *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 4, pp. 787-97.

Huang, Q.-H., Zheng, Y.-P., Li, R. & Lu, M.-H. 2005, '3-D measurement of body tissues based on ultrasound images with 3-D spatial information', *Ultrasound in medicine & biology*, vol. 31, no. 12, pp. 1607-15.

Huang, Q., Deng, Q., Li, L., Yang, J. & Li, X. 2019, 'Scoliotic imaging with a novel double-sweep 2.5-dimensional extended field-of-view ultrasound', *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 66, no. 8, pp. 1304-15.

Huang, Q., Zeng, Z. & Li, X. 2017, '2.5-D extended field-of-view ultrasound', *IEEE transactions on medical imaging*, vol. 37, no. 4, pp. 851-9.

Huang, Y., Han, L., Dou, H., Luo, H., Yuan, Z., Liu, Q., Zhang, J. & Yin, G. 2019, 'Two-stage CNNs for computerized BI-RADS categorization in breast ultrasound images', *Biomedical engineering online*, vol. 18, no. 1, p. 8.

Huang, Z., Wang, L.-W., Leung, F.H., Banerjee, S., Yang, D., Lee, T., Lyu, J., Ling, S.H. & Zheng, Y.-P. 2020, 'Bone Feature Segmentation in Ultrasound Spine Image with Robustness to Speckle and Regular Occlusion Noise', *arXiv preprint arXiv:2010.03740*.

Huang, Z., Zhao, R., Leung, F.H., Banerjee, S., Lee, T.T.-Y., Yang, D., Lun, D.P., Lam, K.-M., Zheng, Y.-P. & Ling, S.H. 2022, 'Joint Spine Segmentation and Noise Removal from Ultrasound Volume Projection Images with Selective Feature Sharing', *IEEE Transactions on Medical Imaging*.

Hwang, B.Y., Mampre, D., Ahmed, A.K., Suk, I., Anderson, W.S., Manbachi, A. & Theodore, N. 2021, 'Ultrasound in Traumatic Spinal Cord Injury: A Wide-Open Field', *Neurosurgery*.

Ibtehaz, N. & Rahman, M.S. 2020, 'MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation', *Neural Networks*, vol. 121, pp. 74-87.

Imran, A.-A.-Z., Huang, C., Tang, H., Fan, W., Cheung, K., To, M., Qian, Z. & Terzopoulos, D. 2020, 'Analysis of scoliosis from spinal x-ray images', *arXiv preprint arXiv:2004.06887*.

Inoue, M., Minami, S., Nakata, Y., Otsuka, Y., Takaso, M., Kitahara, H., Tokunaga, M., Isobe, K. & Moriya, H. 2005, 'Preoperative MRI analysis of patients with idiopathic scoliosis: a prospective study', *Spine*, vol. 30, no. 1, pp. 108-14.

Jaccard, P. 1912, 'The distribution of the flora in the alpine zone. 1', *New phytologist*, vol. 11, no. 2, pp. 37-50.

Jada, A., Mackel, C.E., Hwang, S.W., Samdani, A.F., Stephen, J.H., Bennett, J.T. & Baaj, A.A. 2017, 'Evaluation and management of adolescent idiopathic scoliosis: a review', *Neurosurgical focus*, vol. 43, no. 4, p. E2.

Janumala, T. & Ramesh, K. 2020, 'Development of an Algorithm for Vertebrae Identification Using Speeded up Robost Features (SURF) Technique in Scoliosis X-Ray Images', *International Conference on Image Processing and Capsule Networks*, Springer, pp. 54-62.

Jarrett, K., Kavukcuoglu, K. & LeCun, Y. 2009, 'What is the best multi-stage architecture for object recognition?', *2009 IEEE 12th international conference on computer vision*, IEEE, pp. 2146-53.

Jiang, W.W., Zhou, G.Q., Lai, K.-L., Hu, S.Y., Gao, Q.Y., Wang, X.Y. & Zheng, Y.P. 2019, 'A fast 3-D ultrasound projection imaging method for scoliosis assessment'.

Jin, K.H., McCann, M.T., Froustey, E. & Unser, M. 2017, 'Deep convolutional neural network for inverse problems in imaging', *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509-22.

Kaiser, L., Gomez, A.N. & Chollet, F. 2017, 'Depthwise separable convolutions for neural machine translation', *arXiv preprint arXiv:1706.03059*.

Kalagara, H., Nair, H., Kolli, S., Thota, G. & Uppal, V. 2021, 'Ultrasound Imaging of the Spine for Central Neuraxial Blockade: a Technical Description and Evidence Update', *Current Anesthesiology Reports*, pp. 1-14.

Keenan, B.E., Izatt, M.T., Askin, G.N., Labrom, R.D., Pearcy, M.J. & Adam, C.J. 2014, 'Supine to standing Cobb angle change in idiopathic scoliosis: the effect of endplate pre-selection', *Scoliosis*, vol. 9, no. 1, pp. 1-10.

Khandelwal, P., Collins, D.L. & Siddiqi, K. 2021, 'Spine and Individual Vertebrae Segmentation in Computed Tomography Images Using Geometric Flows and Shape Priors', *Frontiers in Computer Science*, p. 66.

Kim, H., Kim, H.S., Moon, E.S., Yoon, C.-S., Chung, T.-S., Song, H.-T., Suh, J.-S., Lee, Y.H. & Kim, S. 2010, 'Scoliosis imaging: what radiologists should know', *Radiographics*, vol. 30, no. 7, pp. 1823-42.

Kingma, D.P. & Ba, J. 2014, 'Adam: A method for stochastic optimization', *arXiv preprint arXiv:1412.6980*.

Kokabu, T., Kanai, S., Kawakami, N., Uno, K., Kotani, T., Suzuki, T., Tachi, H., Abe, Y., Iwasaki, N. & Sudo, H. 2021, 'An algorithm for using deep learning convolutional neural networks with three dimensional depth sensor imaging in scoliosis detection', *The Spine Journal*, vol. 21, no. 6, pp. 980-7.

Konieczny, M.R., Senyurt, H. & Krauspe, R. 2013, 'Epidemiology of adolescent idiopathic scoliosis', *Journal of children's orthopaedics*, vol. 7, no. 1, pp. 3-9.

Korez, R., Putzier, M. & Vrtovec, T. 2020, 'A deep learning tool for fully automated measurements of sagittal spinopelvic balance from X-ray images: performance evaluation', *European Spine Journal*, vol. 29, no. 9, pp. 2295-305.

Krizhenvshky, A., Sutskever, I. & Hinton, G., 'Imagenet classification with deep convolutional networks', *Proceedings of the Conference Neural Information Processing Systems (NIPS)*, pp. 1097-105.

Krizhevsky, A., Sutskever, I. & Hinton, G.E. 2012, 'Imagenet classification with deep convolutional neural networks', *Advances in neural information processing systems*, pp. 1097-105.

Kumar, K. 1996, 'Spinal deformity and axial traction', *Spine (Phila Pa 1976)*, vol. 21, no. 5, pp. 653-5.

Kundu, R., Chakrabarti, A. & Lenka, P.K. 2012, 'Cobb angle measurement of scoliosis with reduced variability', *arXiv preprint arXiv:1211.5355*.

Lai, K.K.-L., Lee, T.T.-Y., Lee, M.K.-S., Hui, J.C.-H. & Zheng, Y.-P. 2021, 'Validation of Scolioscan Air-Portable Radiation-Free Three-Dimensional Ultrasound Imaging Assessment System for Scoliosis', *Sensors*, vol. 21, no. 8, p. 2858.

Langensiepen, S., Semler, O., Sobottke, R., Fricke, O., Franklin, J., Schönau, E. & Eysel, P. 2013, 'Measuring procedures to determine the Cobb angle in idiopathic scoliosis: a systematic review', *European Spine Journal*, vol. 22, no. 11, pp. 2360-71.

Le Bras, A., Laporte, S., Mitton, D., De Guise, J. & Skalli, W. 2003, 'Three-dimensional (3D) detailed reconstruction of human vertebrae from low-dose digital stereoradiography', *European Journal of Orthopaedic Surgery & Traumatology*, vol. 13, no. 2, pp. 57-62.

LeCun, Y., Kavukcuoglu, K. & Farabet, C. 2010, 'Convolutional networks and applications in vision', *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, IEEE, pp. 253-6.

Lee, T.T.-Y., Lai, K.K.-L., Cheng, J.C.-Y., Castelein, R.M., Lam, T.-P. & Zheng, Y.-P. 2021, '3D ultrasound imaging provides reliable angle measurement with validity comparable to X-ray in patients with adolescent idiopathic scoliosis', *Journal of Orthopaedic Translation*, vol. 29, pp. 51-9.

Levy, A.R., Goldberg, M.S., Mayo, N.E., Hanley, J.A. & Poitras, B. 1996, 'Reducing the lifetime risk of cancer from spinal radiographs among people with adolescent idiopathic scoliosis', *Spine*, vol. 21, no. 13, pp. 1540-7.

Li, H., Luo, H., Huan, W., Shi, Z., Yan, C., Wang, L., Mu, Y. & Liu, Y. 2021, 'Automatic lumbar spinal MRI image segmentation with a multi-scale attention network', *Neural Computing and Applications*, pp. 1-14.

Li, J., Wu, Y., Zhao, J., Guan, L., Ye, C. & Yang, T. 2017, 'Pedestrian detection with dilated convolution, region proposal network and boosted decision trees', *2017 International Joint Conference on Neural Networks (IJCNN)*, IEEE, pp. 4052-7.

Li, M., Cheng, J., Ying, M., Ng, B., Zheng, Y., Lam, T., Wong, W. & Wong, M.S. 2010, 'Application of 3-D ultrasound in assisting the fitting procedure of spinal orthosis to patients with adolescent idiopathic scoliosis', *Research into Spinal Deformities 7*, IOS Press, pp. 34-7.

Li, M., Cheng, J., Ying, M., Ng, B., Zheng, Y., Lam, T., Wong, W. & Wong, M.S. 2012, 'Could clinical ultrasound improve the fitting of spinal orthosis for the patients with AIS?', *European Spine Journal*, vol. 21, no. 10, pp. 1926-35.

Li, S., Yang, J., Zhu, L., Li, Y., Peng, H., Lin, Y., Li, X., Huang, Z. & Wang, H. 2017, 'Left ventricular mechanics assessed by 2-dimensional speckle tracking echocardiography in children and adolescents with idiopathic scoliosis', *Clinical spine surgery*, vol. 30, no. 4, pp. E381-E9.

Lian, X., Pang, Y., Han, J. & Pan, J. 2021, 'Cascaded hierarchical atrous spatial pyramid pooling module for semantic segmentation', *Pattern Recognition*, vol. 110, p. 107622.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B. & Sánchez, C.I. 2017, 'A survey on deep learning in medical image analysis', *Medical image analysis*, vol. 42, pp. 60-88.

Liu, D., Yang, Y., Yu, X., Yang, J., Xuan, X., Yang, J. & Huang, Z. 2020, 'Effects of specific exercise therapy on adolescent patients with idiopathic scoliosis: A prospective controlled cohort study', *Spine*, vol. 45, no. 15, p. 1039.

Liu, J., Yuan, C., Sun, X., Sun, L., Dong, H. & Peng, Y. 2021, 'The measurement of Cobb angle based on spine X-ray images using multi-scale convolutional neural network', *Physical and Engineering Sciences in Medicine*, vol. 44, no. 3, pp. 809-21.

Liu, S. & Huang, D. 2018, 'Receptive field block net for accurate and fast object detection', *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 385-400.

Liu, S., Wang, Y., Yang, X., Lei, B., Liu, L., Li, S.X., Ni, D. & Wang, T. 2019, 'Deep learning in medical ultrasound analysis: a review', *Engineering*, vol. 5, no. 2, pp. 261-75.

Liu, T., Xie, S., Yu, J., Niu, L. & Sun, W. 2017, 'Classification of thyroid nodules in ultrasound images using deep model based transfer learning and hybrid features', *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 919-23.

Liu, W., Li, W. & Gong, W. 2019, 'Ensemble of fine-tuned convolutional neural networks for urine sediment microscopic image classification', *IET Computer Vision*, vol. 14, no. 1, pp. 18-25.

Long, J., Shelhamer, E. & Darrell, T. 2015, 'Fully convolutional networks for semantic segmentation', *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431-40.

Lonstein, J. & Carlson, J. 1984, 'The prediction of curve progression in untreated idiopathic scoliosis', *J Bone Jt Surg*, vol. 3, no. 2, pp. 1061-71.

Loupas, T., McDicken, W. & Allan, P.L. 1989, 'An adaptive weighted median filter for speckle suppression in medical ultrasonic images', *IEEE transactions on Circuits and Systems*, vol. 36, no. 1, pp. 129-35.

Lyu, J., Bi, X., Banerjee, S., Huang, Z., Leung, F.H., Lee, T.T.-Y., Yang, D.-D., Zheng, Y.-P. & Ling, S.H. 2021, 'Dual-task ultrasound spine transverse vertebrae segmentation network with contour regularization', *Computerized Medical Imaging and Graphics*, vol. 89, p. 101896.

McKenna, C., Wade, R., Faria, R., Yang, H., Stirk, L., Gummerson, N., Sculpher, M. & Woolacott, N. 2012, 'EOS 2D/3D X-ray imaging system: a systematic review and economic evaluation', *Health technology assessment (Winchester, England)*, vol. 16, no. 14, p. 1.

Meiburger, K.M., Acharya, U.R. & Molinari, F. 2018, 'Automated localization and segmentation techniques for B-mode ultrasound images: A review', *Computers in biology and medicine*, vol. 92, pp. 210-35.

Melhem, E., Assi, A., El Rachkidi, R. & Ghanem, I. 2016, 'EOS® biplanar X-ray imaging: concept, developments, benefits, and limitations', *Journal of children's orthopaedics*, vol. 10, no. 1, pp. 1-14.

Mitulescu, A., Skalli, W., Mitton, D. & De Guise, J. 2002, 'Three-dimensional surface rendering reconstruction of scoliotic vertebrae using a non stereo-corresponding points technique', *European spine journal*, vol. 11, no. 4, pp. 344-52.

Morrison, D.G., Chan, A., Hill, D., Parent, E.C. & Lou, E.H. 2015, 'Correlation between Cobb angle, spinous process angle (SPA) and apical vertebrae rotation (AVR) on posteroanterior radiographs in adolescent idiopathic scoliosis (AIS)', *European Spine Journal*, vol. 24, no. 2, pp. 306-12.

Morrissy, R., Goldsmith, G., Hall, E., Kehl, D. & Cowie, G. 1990, 'Measurement of the Cobb angle on radiographs of patients who have', *J Bone Joint Surg Am*, vol. 72, no. 3, pp. 320-7.

Morvan, G., Mathieu, P., Vuillemin, V., Guerini, H., Bossard, P., Zeitoun, F. & Wybier, M. 2011, 'Standardized way for imaging of the sagittal spinal balance', *European Spine Journal*, vol. 20, no. 5, pp. 602-8.

Musson, R.E., Warren, D.J., Bickle, I., Connolly, D.J. & Griffiths, P.D. 2010, 'Imaging in childhood scoliosis: a pictorial review', *Postgraduate medical journal*, vol. 86, no. 1017, pp. 419-27.

Nash Jr, C., Gregg, E., Brown, R. & Pillai, K. 1979, 'Risks of exposure to X-rays in patients undergoing long-term treatment for scoliosis', *The Journal of bone and joint surgery. American volume*, vol. 61, no. 3, pp. 371-4.

Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y. & Kainz, B. 2018, 'Attention u-net: Learning where to look for the pancreas', *arXiv preprint arXiv:1804.03999*.

Pan, Y., Chen, Q., Chen, T., Wang, H., Zhu, X., Fang, Z. & Lu, Y. 2019, 'Evaluation of a computer-aided method for measuring the Cobb angle on chest X-rays', *European Spine Journal*, vol. 28, no. 12, pp. 3035-43.

Pandey, P.U., Quader, N., Guy, P., Garbi, R. & Hodgson, A.J. 2020, 'Ultrasound Bone Segmentation: A Scoping Review of Techniques and Validation Practices', *Ultrasound in Medicine & Biology*, vol. 46, no. 4, pp. 921-35.

Pasha, S., Rajapaske, C.R., Reddy, R., Diebo, B., Knott, P., Jones, B.C., Kumar, D., Zhu, W., Lou, E. & Shapira, N. 2021, 'Quantitative imaging of the spine in adolescent idiopathic scoliosis: shifting the paradigm from diagnostic to comprehensive prognostic evaluation', *European Journal of Orthopaedic Surgery & Traumatology*, pp. 1-13.

Phan, H.T.H., Kumar, A., Kim, J. & Feng, D. 2016, 'Transfer learning of a convolutional neural network for HEp-2 cell image classification', *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, IEEE, pp. 1208-11.

Pleiss, G., Chen, D., Huang, G., Li, T., van der Maaten, L. & Weinberger, K.Q. 2017, 'Memory-efficient implementation of densenets', *arXiv preprint arXiv:1707.06990*.

Presciutti, S.M., Karukanda, T. & Lee, M. 2014, 'Management decisions for adolescent idiopathic scoliosis significantly affect patient radiation exposure', *The Spine Journal*, vol. 14, no. 9, pp. 1984-90.

Pruijs, J., Hageman, M., Keessen, W., Van Der Meer, R. & Van Wieringen, J. 1994, 'Variation in Cobb angle measurements in scoliosis', *Skeletal radiology*, vol. 23, no. 7, pp. 517-20.

Q Zheng, S.L.F., G E Tasian, Y Fan October 2018, 'Computer aided diagnosis of congenital abnormalities of the kidney and urinary tract in children based on ultrasound imaging data by integrating texture image features and deep transfer learning image features', *Journal of Pediatric Urology* vol. DOI: 10.1016/j.jpurol.2018.10.020.

Rajasekaran, S., Kamath, V., Kiran, R. & Shetty, A.P. 2010, 'Intraspinal anomalies in scoliosis: an MRI analysis of 177 consecutive scoliosis patients', *Indian journal of orthopaedics*, vol. 44, no. 1, pp. 57-63.
Ran, H., Zhi-hong, W. & Jiang-na, H. 2011, 'Scoliosis on pulmonary function', *Zhongguo yi xue ke xue Yuan xue bao. Acta Academiae Medicinae Sinicae*, vol. 33, no. 1, pp. 102-6.
Ravishankar, H., Venkataramani, R., Thiruvenkadam, S., Sudhakar, P. & Vaidya, V. 2017, 'Learning and incorporating shape models for semantic segmentation', *International conference on medical image computing and computer-assisted intervention*, Springer, pp. 203-11.
Reamy, B.V. & Slakey, J. 2001, 'Adolescent idiopathic scoliosis: review and current concepts', *American family physician*, vol. 64, no. 1, p. 111.
Ronneberger, O., Fischer, P. & Brox, T. 2015, 'U-net: Convolutional networks for biomedical image segmentation', *International Conference on Medical image computing and computer-assisted intervention*, Springer, pp. 234-41.
Roth, H.R., Shen, C., Oda, H., Oda, M., Hayashi, Y., Misawa, K. & Mori, K. 2018, 'Deep learning and its application to medical image segmentation', *Medical Imaging Technology*, vol. 36, no. 2, pp. 63-71.
Safari, A., Parsaei, H., Zamani, A. & Pourabbas, B. 2019, 'A semi-automatic algorithm for estimating cobb angle', *Journal of biomedical physics & engineering*, vol. 9, no. 3, p. 317.
Samuvel, B., Thomas, V. & Mini, M. 2012, 'A mask based segmentation algorithm for automatic measurement of Cobb angle from scoliosis x-ray image', *2012 International Conference on Advances in Computing and Communications*, IEEE, pp. 110-3.
Sardjono, T.A., Wilkinson, M.H., Veldhuizen, A.G., van Ooijen, P.M., Purnama, K.E. & Verkerke, G.J. 2013, 'Automatic Cobb angle determination from radiographic images', *Spine*, vol. 38, no. 20, pp. E1256-E62.
Schlemper, J., Oktay, O., Schaap, M., Heinrich, M., Kainz, B., Glocker, B. & Rueckert, D. 2019, 'Attention gated networks: Learning to leverage salient regions in medical images', *Medical image analysis*, vol. 53, pp. 197-207.
Schmitz-Feuerhake, I. & Pflugbeil, S. 2011, '‘Lifestyle’and cancer rates in former East and West Germany: the possible contribution of diagnostic radiation exposures', *Radiation protection dosimetry*, vol. 147, no. 1-2, pp. 310-3.
Schulte, T.L., Hierholzer, E., Boerke, A., Lerner, T., Liljenqvist, U., Bullmann, V. & Hackenberg, L. 2008, 'Raster stereography versus radiography in the long-term follow-up of idiopathic scoliosis', *Clinical Spine Surgery*, vol. 21, no. 1, pp. 23-8.
Shi, B., Mao, S., Wang, Z., Lam, T.P., Yu, F.W.P., Ng, B.K.W., Chu, W.C.-W., Zhu, Z., Qiu, Y. & Cheng, J.C.Y. 2015, 'How does the supine MRI correlate with standing radiographs of different curve severity in adolescent idiopathic scoliosis?', *Spine*, vol. 40, no. 15, pp. 1206-12.
Sifre, L. & Mallat, S. 2014, 'Rigid-motion scattering for image classification', *Ph. D. thesis*.
Simony, A., Hansen, E.J., Christensen, S.B., Carreon, L.Y. & Andersen, M.O. 2016, 'Incidence of cancer in adolescent idiopathic scoliosis patients treated 25 years previously', *European Spine Journal*, vol. 25, no. 10, pp. 3366-70.
Simonyan, K. & Zisserman, A. 2014, 'Very deep convolutional networks for large-scale image recognition', *arXiv preprint arXiv:1409.1556*.

Sun, W. & Wang, R. 2018, 'Fully convolutional networks for semantic segmentation of very high resolution remotely sensed images combined with DSM', *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 3, pp. 474-8.

Szabo, T.L. 2004, *Diagnostic ultrasound imaging: inside out*, Academic press.

Szegedy, C., Ioffe, S., Vanhoucke, V. & Alemi, A. 2016, 'Inception-v4, inception-resnet and the impact of residual connections on learning', *arXiv preprint arXiv:1602.07261*.

Szegedy, C., Ioffe, S., Vanhoucke, V. & Alemi, A.A. 2017, 'Inception-v4, inception-resnet and the impact of residual connections on learning', *Thirty-first AAAI conference on artificial intelligence*.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. & Rabinovich, A. 2015, 'Going deeper with convolutions', *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1-9.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. 2016, 'Rethinking the inception architecture for computer vision', *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818-26.

Tawfik, N.A., Ahmed, A.T., El-Shafei, T.E. & Habba, M.R. 2020, 'Diagnostic value of spinal ultrasound compared to MRI for diagnosis of spinal anomalies in pediatrics', *Egyptian Journal of Radiology and Nuclear Medicine*, vol. 51, no. 1, pp. 1-11.

Thomson, B.R., Nijkamp, J., Ivashchenko, O., van der Heijden, F., Smit, J.N., Kok, N.F., Kuhlmann, K.F., Ruers, T.J. & Fusaglia, M. 2019, 'Hepatic vessel segmentation using a reduced filter 3D U-Net in ultrasound imaging', *arXiv preprint arXiv:1907.12109*.

Thong, W.E., Labelle, H., Shen, J., Parent, S. & Kadoury, S. 2015, 'Stacked auto-encoders for classification of 3d spine models in adolescent idiopathic scoliosis', *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Springer, pp. 13-25.

Tu, Y., Wang, N., Tong, F. & Chen, H. 2019, 'Automatic measurement algorithm of scoliosis Cobb angle based on deep learning', *Journal of Physics: Conference Series*, vol. 1187, IOP Publishing, p. 042100.

Ungi, T., Greer, H., Sunderland, K.R., Wu, V., Baum, Z.M., Schlenger, C., Oetgen, M., Cleary, K., Aylward, S.R. & Fichtinger, G. 2020, 'Automatic spine ultrasound segmentation for scoliosis visualization and measurement', *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 11, pp. 3234-41.

Ungi, T., King, F., Kempston, M., Keri, Z., Lasso, A., Mousavi, P., Rudan, J., Borschneck, D.P. & Fichtinger, G. 2014, 'Spinal curvature measurement by tracked ultrasound snapshots', *Ultrasound in medicine & biology*, vol. 40, no. 2, pp. 447-54.

Vedula, S., Senouf, O., Bronstein, A.M., Michailovich, O.V. & Zibulevsky, M. 2017, 'Towards CT-quality ultrasound imaging using deep learning', *arXiv preprint arXiv:1710.06304*.

Vo, Q.N., Le, L.H. & Lou, E. 2019, 'A semi-automatic 3D ultrasound reconstruction method to assess the true severity of adolescent idiopathic scoliosis', *Medical & biological engineering & computing*, vol. 57, no. 10, pp. 2115-28.

Wahab, N., Khan, A. & Lee, Y.S. 2017, 'Two-phase deep convolutional neural network for reducing class skewness in histopathological images based breast cancer detection', *Computers in biology and medicine*, vol. 85, pp. 86-97.

Wang, Q., Li, M., Lou, E.H. & Wong, M.S. 2015, 'Reliability and validity study of clinical ultrasound imaging on lateral curvature of adolescent idiopathic scoliosis', *PloS one*, vol. 10, no. 8, p. e0135264.
Wright, N. 2000, 'Imaging in scoliosis', *Archives of disease in childhood*, vol. 82, no. 1, pp. 38-40.
Wu, H.-D., Liu, W. & Wong, M.-S. 2020, 'Reliability and validity of lateral curvature assessments using clinical ultrasound for the patients with scoliosis: a systematic review', *European Spine Journal*, vol. 29, no. 4, pp. 717-25.
Yan, B., Lu, X., Qiu, Q., Nie, G. & Huang, Y. 2020, 'Association between incorrect posture and adolescent idiopathic scoliosis among Chinese adolescents: Findings from a large-scale population-based study', *Frontiers in Pediatrics*, vol. 8.
Yeung, K.H., Man, G.C.W., Lam, T.P., Ng, B.K.W., Cheng, J.C.Y. & Chu, W.C.W. 2020, 'Accuracy on the preoperative assessment of patients with adolescent idiopathic scoliosis using biplanar low-dose stereoradiography: a comparison with computed tomography', *BMC Musculoskeletal Disorders*, vol. 21, no. 1, pp. 1-10.
Young, M., Hill, D.L., Zheng, R. & Lou, E. 2015, 'Reliability and accuracy of ultrasound measurements with and without the aid of previous radiographs in adolescent idiopathic scoliosis (AIS)', *European Spine Journal*, vol. 24, no. 7, pp. 1427-33.
Yu, L., Guo, Y., Wang, Y., Yu, J. & Chen, P. 2017, 'Segmentation of fetal left ventricle in echocardiographic sequences based on dynamic convolutional neural networks', *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 8, pp. 1886-95.
Zeiler, M.D. & Fergus, R. 2014, 'Visualizing and understanding convolutional networks', *European conference on computer vision*, Springer, pp. 818-33.
Zeng, H.-Y., Ge, S.-H., Gao, Y.-C., Zhou, D.-S., Zhou, K., He, X.-M., Lou, E. & Zheng, R. 2021, 'Automatic segmentation of vertebral features on ultrasound spine images using Stacked Hourglass Network', *arXiv preprint arXiv:2105.03847*.
Zeng, H.-Y., Ge, S.-H., Gao, Y.-C., Zhou, D.-S., Zhou, K., He, X.-M. & Zheng, R. 2021, 'Automatic segmentation of vertebral features on ultrasound spine images using Stacked Hourglass Network', *arXiv preprint arXiv:2105.03847*.
Zeng, H., Zheng, R., Le, L.H. & Ta, D. 2019, 'Measuring spinous process angle on ultrasound spine images using the GVF segmentation method', *2019 IEEE International Ultrasonics Symposium (IUS)*, IEEE, pp. 1477-80.
Zhang, J., Cui, X., Chen, S., Dai, Y., Huang, Y. & Zhang, S. 2021, 'Ultrasound-guided nusinersen administration for spinal muscular atrophy patients with severe scoliosis: an observational study', *Orphanet Journal of Rare Diseases*, vol. 16, no. 1, pp. 1-8.
Zhang, J., Li, H., Lv, L. & Zhang, Y. 2017, 'Computer-aided Cobb measurement based on automatic detection of vertebral slopes using deep neural network', *International journal of biomedical imaging*, vol. 2017.
Zhang, J., Lou, E., Shi, X., Wang, Y., Hill, D.L., Raso, J.V., Le, L.H. & Lv, L. 2010, 'A computer-aided Cobb angle measurement method and its reliability', *Clinical Spine Surgery*, vol. 23, no. 6, pp. 383-7.
Zhang, Z., Wu, C., Coleman, S. & Kerr, D. 2020, 'DENSE-INception U-net for medical image segmentation', *Computer Methods and Programs in Biomedicine*, vol. 192, p. 105395.
Zheng, R., Young, M., Hill, D., Le, L.H., Hedden, D., Moreau, M., Mahood, J., Southon, S. & Lou, E. 2016, 'Improvement on the accuracy and reliability of ultrasound coronal curvature measurement on adolescent idiopathic scoliosis with the aid of previous radiographs', *Spine*, vol. 41, no. 5, pp. 404-11.

Zheng, Y.-P., Lee, T.T.-Y., Lai, K.K.-L., Yip, B.H.-K., Zhou, G.-Q., Jiang, W.-W., Cheung, J.C.-W., Wong, M.-S., Ng, B.K.-W. & Cheng, J.C.-Y. 2016, 'A reliability and validity study for Scolioscan: a radiation-free scoliosis assessment system using 3D ultrasound imaging', *Scoliosis and spinal disorders*, vol. 11, no. 1, p. 13.

Zhou, G.-Q., Jiang, W.-W., Lai, K.-L., Lam, T.-P., Cheng, J.C.-Y. & Zheng, Y.-P. 2016, 'Semi-automatic Measurement of Scoliotic Angle Using a Freehand 3-D Ultrasound System Scolioscan', *XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016*, Springer, pp. 341-6.

Zhou, G.-Q., Jiang, W.-W., Lai, K.-L. & Zheng, Y.-P. 2017, 'Automatic measurement of spine curvature on 3-D ultrasound volume projection image with phase features', *IEEE transactions on medical imaging*, vol. 36, no. 6, pp. 1250-62.

Zhou, G.-Q., Li, D.-S., Zhou, P., Jiang, W.-W. & Zheng, Y.-P. 2020, 'Automating spine curvature measurement in volumetric ultrasound via adaptive phase features', *Ultrasound in medicine & biology*, vol. 46, no. 3, pp. 828-41.

Zhou, G.-Q. & Zheng, Y.-P. 2015, 'Assessment of scoliosis using 3-D ultrasound volume projection imaging with automatic spine curvature detection', *2015 IEEE International Ultrasonics Symposium (IUS)*, IEEE, pp. 1-4.

Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N. & Liang, J. 2018, 'Unet++: A nested u-net architecture for medical image segmentation', *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, pp. 3-11.

Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N. & Liang, J. 2019, 'Unet++: Redesigning skip connections to exploit multiscale features in image segmentation', *IEEE transactions on medical imaging*, vol. 39, no. 6, pp. 1856-67.