

Please cite the Published Version

Shaubari, Ezak Fadzrin Ahmad (2018) Automatic segmentation of the human thigh muscles in magnetic resonance imaging. Doctoral thesis (PhD), Manchester Metropolitan University.

Downloaded from: <https://e-space.mmu.ac.uk/621007/>

Usage rights:  Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

MANCHESTER METROPOLITAN UNIVERSITY

Automatic Segmentation of the Human Thigh Muscles in Magnetic Resonance Imaging

Ezak Fadzrin Ahmad Shaubari

A thesis submitted in partial fulfilment of the requirements of
the Manchester Metropolitan University for the degree of Doctor of Philosophy

Faculty of Science and Engineering
School of Computing, Mathematics and Digital Technology

June 2018

Declaration of Authorship

I, [Ezak Fadzrin Ahmad Shaubari](#) , declare that this thesis titled, ‘Automatic Segmentation of the Human Thigh Muscles in Magnetic Resonance Imaging’ and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at Manchester Metropolitan University (MMU).
- Where any part of this thesis has previously been submitted for a degree or any other qualification at MMU or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:



Date:

19 APRIL 2018

“How foolish is a man! He ruins the present while worrying about the future, but weeps in the future by recalling the past!”

Ali ibn Abi Talib

“Acquire wisdom from the story of those who have already passed.”

Uthman ibn Affan

“Do not be an arrogant scholar, for scholarship cannot subsist with arrogance.”

Umar ibn Khattab

“Run away from greatness and greatness will follow you.”

Abu Bakr ibn Siddiq

“The cure for ignorance is to question and; there is no beauty better than the intellect.”

Muhammad pbuh

Abstract

Advances in magnetic resonance imaging (MRI) and analysis techniques have improved diagnosis and patient treatment pathways. Typically, image analysis requires substantial technical and medical expertise and MR images can suffer from artefacts, echo and intensity inhomogeneity due to gradient pulse eddy currents and inherent effects of pulse radiation on MRI radio frequency (RF) coils that complicates the analysis. Processing and analysing serial sections of MRI scans to measure tissue volume is an additional challenge as the shapes and the borders between neighbouring tissues change significantly by anatomical location. Medical imaging solutions are needed to avoid laborious manual segmentation of specified regions of interest (ROI) and operator errors.

The work set out in this thesis has addressed this challenge with a specific focus on skeletal muscle segmentation of the thigh. The aim was to develop an MRI segmentation framework for the quadriceps muscles, femur and bone marrow. Four contributions of this research include: (1) the development of a semi-automatic segmentation framework for a single transverse-plane image; (2) automatic segmentation of a single transverse-plane image; (3) the automatic segmentation of multiple contiguous transverse-plane images from a full MRI thigh scan; and (4) the use of deep learning for MRI thigh quadriceps segmentation.

Novel image processing, statistical analysis and machine learning algorithms were developed for all solutions and they were compared against current gold-standard manual segmentation. Frameworks (1) and (3) require minimal input from the user to delineate the muscle border. Overall, the frameworks in (1), (2) and (3) offer very good output performance, with respective framework's mean segmentation accuracy by *JSI* and processing time of: (1) 0.95 and 17 sec; (2) 0.85 and 22 sec; and (3) 0.93 and 3 sec. For the framework in (4), the ImageNet trained model was customized by replacing the fully-connected layers in its architecture to convolutional layers (hence the name of Fully Convolutional Network (FCN)) and the pre-trained model was transferred for the ROI segmentation task. With the implementation of post-processing for image filtering and morphology to the segmented ROI, we have successfully accomplished a new benchmark for thigh MRI analysis. The mean accuracy and processing time with this framework are 0.9502 (by *JSI*) and 0.117 sec per image, respectively.

Acknowledgements

I wish to acknowledge the many people who have helped to make this work possible. First of all I would like to express my sincere gratitude to my Director of Study, Dr Moi Hoon Yap, for introducing me to this research project, for her constant encouragement and guidance and for her consistent supervision given the circumstances. Her profound insight and splendid vision and patient, gave me a great opportunity to get into the world of computer science, considering my Medical Engineering academic background. Her enthusiasm, optimism, hard work, valuable support and advice were the greatest factors that enabled me to complete my research work, publication establishment and in thesis writing up.

I would like to thank Prof Jamie S McPhee and Prof Dr Ir Hans Degens, both from School of Healthcare Science, for being part of the supervisory team, for their consistent and excellent technical advice and support, for the access to those MRI datasets, for their contributions in publishing manuscripts, their feedbacks and input towards robust and balance technical statements and “jargons” that ideal for computing, healthcare and engineering environment, for that brief introduction to histological side of the research, for constant guidance and understanding, especially on the management and administrative parts.

I would like to thank and appreciate many people that have helped me directly and/or indirectly, inside and outside the Manchester Metropolitan University (MMU), Dr Kevin Tan, Dr John Darby, Dr Xia Han, Dr Choong-Ching Ng, Dr Daniel Leightley, Dr Muhammad Salman Haleem, Dr Adrian Davison, Dr Zubir Azhar, Dr Badrul Esa, Mr Liam Bagley, Mr Alex Ireland, Mr Mathew Piasecki, Dr Anna Leida Mölder, Ms Gemma Stringer, Mr Rohimi Tukiman, Mr Manu Goyal and everyone in *E121*. They have made my research life and experience, inside and outside the university, an unforgettable one, and many have provided me with useful thought, input and advice as well as moral and administrative support.

My special thanks to Mr Thomas Sherwood, Mr Gary Chadwick and Mr Chuo Yeh Poo for the brief collaboration, input and contribution, towards my research potentials and its flexibility. I am also grateful to have met and made many friends and associates throughout this journey and therefore, wishing the very best and dearest outcome possible, academically, spiritually, career-wise and for life in general, to all. And finally, my special appreciation and gratitude to Majlis Amanah Rakyat (MARA) for the opportunity and financial support and MMU as a whole, for making this research study possible.

Publications

This thesis is based on material from the following publications:

1. Published article(s)

- (a) E. Ahmad, M. H. Yap, H. Degens, and J. S. McPhee, “Atlas-registration based image segmentation of MRI human thigh muscles in 3-D space”, in SPIE Medical Imaging. International Society for Optics and Photonics, 2014, pp. 90 371L90 371L.
- (b) E. Ahmad, M. H. Yap, H. Degens, and J. S. McPhee, “Enhancement of MRI Human Thigh Muscle Segmentation by Template-based Framework”, in IEEE International Conference Control System, Computing and Engineering ICCSCE, 2014.
- (c) E. Ahmad, J. S. McPhee, H. Degens, and M. H. Yap, “Automatic Segmentation of MRI Human Thigh Muscle - Combination of Reliable and Fast Framework Methods for Quadriceps, Femur and Marrow Segmentation”, in International Conference on Biomedical Engineering and Technology ICBET, 2018.

2. Article(s) under review

- (a) E. Ahmad, M. Goyal, H. Degens, J. S. McPhee and M.H. Yap, “Semantic Segmentation of MRI Human Thigh Muscle”, Journal of Computerized Medical Imaging and Graphics, TBA (submitted: Dec 2017), 2018.

3. Other(s)

- (a) T. Sherwood, E. Ahmad, and M. H. Yap. “Formulating efficient software solution for digital image processing system”, in Software: Practice and Experience, 2015.
- (b) A. N. Baharum, M. H. Yap, G. Ferris, E. Ahmad and H. Degens, “Capillary detection in transverse muscle sections”, in SPIE Medical Imaging. International Society for Optics and Photonics, 2018.
- (c) M. H. Yap, M. Goyal, E. Ahmad, Fatima M. Osman, R. Marti, R. Zwiggelaar, A. Juette and E. Denton, “End-to-end breast ultrasound lesions recognition with a deep learning approach”, in SPIE Medical Imaging. International Society for Optics and Photonics, 2018.

Dedicated to A.I.M., 2M's, 2F's and my 5N's

Contents

Declaration of Authorship	i
Abstract	iii
Acknowledgements	iv
Publications	v
List of Figures	xii
List of Tables	xvii
Abbreviations	xviii
1 Introduction	1
1.1 Research Background	1
1.2 Brief Overview of Work Developed in This Thesis	4
1.3 Aims and Objectives	6
Objective 1.	6
Objective 2.	6
Objective 3.	7
Objective 4.	7
Objective 5.	7
1.4 Research Contributions	7
1.5 Thesis Outline	8
2 Research Foundation	10
2.1 Introduction	10
2.2 Human Physiology	12
2.2.1 Muscular System	12
2.2.1.1 Muscle Types	13
2.2.2 Anatomy of Thigh Muscles	14
2.2.2.1 The Anterior Thigh Muscles	15
2.2.2.2 The Medial Thigh Muscles	16
2.2.2.3 The Posterior Thigh Muscles	17

2.2.3	General Effects of Ageing in Human Physiology	17
2.3	Medical Imaging Modalities	21
2.3.1	X-Ray	21
2.3.2	Computed Tomography	22
2.3.3	Ultrasound	23
2.3.4	Magnetic Resonance Imaging	24
2.3.4.1	MR Imaging Sequences	26
2.3.5	Other Modalities	26
2.3.5.1	Positron Emission Tomography (PET)	26
2.3.5.2	Single-Photon Emission Computed Tomography (SPECT)	27
2.3.5.3	Mammography	28
2.4	Previous Works on Image Segmentation of Human Thigh Muscles	28
2.5	Discussion and Conclusion	32
3	Methodologies in Medical Image Segmentation	35
3.1	Introduction	35
3.2	Segmentation Methods	36
3.2.1	First Generation	37
3.2.1.1	Thresholding	37
3.2.1.2	Edge Detection	38
3.2.1.3	Region Growing	45
3.2.2	Second Generation	46
3.2.2.1	Clustering	46
K-means.	46
Fuzzy C-mean.	47
3.2.2.2	Watershed	48
3.2.2.3	Active Contour	50
Edge-based.	50
Region-based.	52
3.2.2.4	Level Set	55
3.2.2.5	Statistical Analysis for Shape Representation	56
Principle Component Analysis.	57
3.2.2.6	Registration	58
Feature-based.	58
Intensity-based.	61
3.2.2.7	Markov Random Field	63
3.2.3	Third Generation	64
3.2.3.1	Atlas-based	64
3.2.3.2	Shape Prior	66
3.2.3.3	Neural Network	68
3.3	Segmentation Performance Assessment	72
3.3.1	Jaccard Similarity Index	72
3.3.2	Dice Similarity Coefficient	73
3.3.3	Zijdenbos Similarity Index	73
3.4	Materials	73
3.5	Discussion and Conclusion	74

4	Semi Automatic Segmentation of MRI Human Thigh Muscles	77
4.1	Introduction	77
4.2	Manual Annotation Process	77
4.3	Methods for Semi Automatic Segmentation	79
4.3.1	Image Pre-processing	79
4.3.2	Background Removal	81
4.3.2.1	Adaptive and Automatic Grayscale Threshold	81
4.3.2.2	Convex Hull Pixel Identification	81
4.3.3	Heterogeneous Digital Line Drawing	83
4.3.4	ROI Processing	83
4.4	Experiments and Results	84
4.5	Discussion	86
4.6	Conclusion	89
5	Automatic Segmentation of MRI Human Thigh Muscles	90
5.1	Introduction	90
5.2	Data	92
5.3	Methods for Automatic Thigh MRI Segmentation	93
5.3.1	Segmentation of Components of Thigh MR Images	93
5.3.2	ROI (Quadriceps, Femur and Bone Marrow) Segmentation	95
5.3.2.1	Shape Model for Mid-scan	95
5.3.2.2	Processing of Input Image	96
	Pre-processing.	96
	Phase Congruency.	97
	Edge Enhancement Technique.	98
5.3.2.3	Transformation of Shape Model	100
5.3.2.4	Active Contour and Smoothing Process	101
5.4	Experiments and Results	103
5.5	Discussion	104
5.6	Conclusion	106
6	Combination Frameworks Leading to Automated Dataset Segmentation	108
6.1	Introduction	108
6.2	Methods for Automatic Dataset Segmentation	109
6.2.1	Semi-Automatic Segmentation	109
6.2.2	Automatic Segmentation	111
6.3	Experiments and Results	113
6.4	Discussion	116
6.5	Conclusion	118
7	Semantic Segmentation of MR Images of Human Thigh Muscle	119
7.1	Introduction	119
7.2	Deep Learning Architectures	121
7.2.1	AlexNet	122
7.2.2	VGG-16	123
7.2.3	PSPNet	123

7.3	Experiments	125
7.3.1	Datasets	125
7.3.2	Transfer Learning	126
7.3.2.1	Customized Deep CNNs (Pre-trained Models)	127
	FCN-AlexNet.	128
	FCN-VGGNet.	129
	PSPNet.	129
7.3.3	Configurations	129
	Machine 1.	129
	Machine 2.	130
7.4	Results	130
7.4.1	FCN-8s model with different solver types	132
7.4.2	Segmentation of MRI Whole-scan Dataset Trained Models on MRI Mid-scan Dataset	133
7.4.3	Post-processing	135
7.5	Discussion	137
7.6	Conclusion	138
8	Conclusion	140
8.1	Introduction	140
8.2	Research Findings	141
8.3	The Challenges	145
8.4	The Trials	146
8.5	Future Works	148
8.6	Final Remarks	150
A	OsiriX	151
A.1	3-D Viewer for Volume and Surface Rendering	151
A.2	Area Conversion	151
B	Architecture Visualization for Convolutional Neural Network	154
B.1	AlexNet Based Models	155
B.1.1	AlexNet	155
B.1.2	FCN-AlexNet	156
B.2	VGG-16 Based Models	157
B.2.1	VGG-16	157
B.2.2	FCN-32s	158
B.2.3	FCN-16s	159
B.2.4	FCN-8s	160
B.3	PSPNet	161
C	Publications	162
C.1	Published Article(s)	163
C.1.1	SPIE 2014	163
C.1.2	ICCSCE 2014	164
C.1.3	ICBET 2018	165
C.2	Other(s)	166

C.2.1	Journal of Software: Practice and Experience	166
C.2.2	SPIE 2018 - Paper 1	167
C.2.3	SPIE 2018 - Paper 2	168

Bibliography		169
---------------------	--	------------

List of Figures

1.1	Magnetic Resonance Imaging (MRI) scans from (a) Young and (b) Elderly.	2
2.1	3 different type of muscles in human body. <i>Courtesy: www.study.com</i>	15
2.2	Cross section of thigh muscles. <i>Courtesy: https://en.wikipedia.org</i>	18
2.3	Effects of ageing in human physiology 1. <i>Courtesy: https://www.flickr.com/photos/sportex/5387348258</i>	20
2.4	Effects of ageing in human physiology 2. <i>Courtesy: https://www.slideshare.net/Drraveesoni/psychopharmacology-in-elderly</i>	20
2.5	Medical imaging output characteristics of X-ray (left) and CT scan (right). <i>Courtesy: www.fda.gov</i>	23
2.6	The cutaway of typical MRI machine. <i>Courtesy: www.howstuffworks.com</i>	25
2.7	Differences of MR imaging sequences of the brain. (a) T1-weighted image. (b) T2-weighted image. (c) Flair image. <i>Courtesy: http://casemed.case.edu/clerkships/neurology/</i>	26
2.8	MR images of the spine. (a) T1-weighted image; white arrow indicates injured structure on the bone injury. (b) T2-weighted image; black arrow depicts bone oedema (feature of bone injury). <i>Courtesy: http://www.lbgmedical.com/blog/mri-a-sophisticated-technology-for-imaging-soft-tissue-pathology/</i>	27
2.9	MRI output profiles. (a) Intensity inhomogeneity. (b) Echo. (c) Overlapping of pixels (limb movements during scan). (d) Yellow arrows indicate muscles that are not part of the region of interest (ROI) but appear to attach together with the ROI. (e) Thigh MRI composition in elderly. (f) Thigh MRI composition in young subject.	33
3.1	Results of different thresholding procedures on an MR image. (a) Original image. (b) Otsu's thresholding. (c) Adaptive thresholding.	38
3.2	Edge profiles.	39
3.3	Neighbourhood pixels labelling for Sobel and Prewitt operators.	41
3.4	Results of different edge detection operators. <i>Cameraman image courtesy of Mathworks.</i>	45
3.5	Seed-based region growing segmentation. Left - Original thigh MR image. Right - region growing segmentation result after seeding point placement (yellow marker) on the muscle region.	46
3.6	<i>K</i> -means applications on 2 (top row) and 4 (bottom row) clusters of random data. First columns represent random data generations. Second columns: clusters initialization. Third columns: assignment of data points to closer cluster. Final columns: Centre (final centroid location) of each cluster.	47

3.7	Top 1-D graph - Data distribution along x -axis. Mid graph - Membership function of two clusters by K -means. Bottom graph - Membership function of two clusters by FCM. The point marked as a red dot, although belongs to B cluster, also belongs to A cluster, with membership degree valued at 0.2. <i>Courtesy: http://home.deib.polimi.it</i>	49
3.8	Left - Simulated image with two dark blobs. Right - How watershed algorithm treats image on the left, such that the dark regions are “low” and bright regions are “high”, resulting watershed line and catchment basins. <i>Courtesy: http://uk.mathworks.com</i>	50
3.9	Image segmentation by watershed. (a) Original image. (b) Watershed transformation of (a). (c) Watershed transformation with Euclidean distance transform of the binary image (a). (d) Segmentation result with desired location minima. <i>Courtesy: http://uk.mathworks.com</i>	51
3.10	Image segmentation by different mode of active contours on putamen brain’s sub-cortex MR image. First column represents different curve initializations on the image. Second column - Curve convolutions by region-based algorithm. Third column - Curve transformations by edge-based. Last column - Segmentation results by hybrid active contour. <i>Courtesy: http://www.shawnlankton.com/2007/02/cool-hybrid-segmentation/</i>	54
3.11	Lemons and CT vessel (top and bottom row, respectively) image segmentation by level set. Left column represents initial contours. Mid column - intermediate contours after certain number of iterations. Right column - final contours. <i>Courtesy: Li et al. - A Level Set Method for Image Segmentation in the Presence of Intensity Inhomogeneities with Application to MRI.</i>	57
3.12	Projection of 3-D dataset into 2-D pattern expressions using PCA. The data compresses with minimal information degradation. <i>Courtesy: http://www.geomodeling.ca/principal-component-analysis/</i>	59
3.13	Sample of affine transformation process of a hexagonal shape.	62
3.14	Registering multimodal MR images by intensity-based image registration. Image on the left represents moving image. Mid image - Fixed (template) image. Image on the right - Registered moving image with affine registration mode - Registration is done by using Matlab. <i>Courtesy: http://uk.mathworks.com/help/images/registering-multimodal-mri-images.html</i>	63
3.15	Markov model graph. (a) Grids of a simple 4-connected pixels. (b) Grids with extra connections. (c) Grids with irregular pattern. <i>Courtesy: Andrew Blake and Pushmeet Kohli - Introduction to Markov Random Fields</i>	63
3.16	Segmentation model of joint-registration method. <i>Courtesy: Wang et al. - Joint Registration and Segmentation of Neuroanatomic Structures From Brain MRI</i>	65
3.17	Image segmentation by statistical shape prior using manifold learning on right ventricular on cardiac MR image. (a), (b), (c) and (d) represent original image with respected initialization of shape prior curve. (e), (f), (g) and (h) - segmentation results, from respective top row. <i>Courtesy: http://www.litislab.fr/members-cpetitjean/shapepriorsegm/</i>	67
3.18	Schematic model for a neuron in an ANN.	68
3.19	Architecture of a simple CNN model. <i>Courtesy: https://www.clarifai.com/technology</i>	69

3.20	Dataset with 13 scans.	74
3.21	Dataset with 26 scans.	74
4.1	OsiriX layout and environment.	78
4.2	Sample of manual annotation by OsiriX.	79
4.3	MRI N3 correction of intensity inhomogeneity method by MIPAV. (a) Before and (b) After correction.	80
4.4	Adaptive and automatic grayscale thresholding for unwanted background removal. (a) Original grayscale. (b) Application of proposed threshold method. (c) Refinement result.	81
4.5	(a) A set of S random points in 2-D Euclidean space. (b) Convex hull of points in (a).	82
4.6	(a) Application of convex hull method from output of Figure 4.4(c). (b) Skin and adipose tissue removal. (c) Mask of binary blob ROIs in (b) reverted back to original grayscale value.	82
4.7	(a) Original MR image. (b) Segmented output result, where red digital marker line is the only interaction needed from the operator.	83
4.8	(a) Product of proposed MRI semi-automatic segmentation. (b) Binary image of total ROI. (c) Binary blob masks of ROI in (a).	84
4.9	(a) Manual segmentation by OsiriX. (b) Product of semi-automatic segmentation method.	86
5.1	Application of thresholding and region properties. (a) Original MR image. (b) Muscle (gray) and bone marrow regions (white).	94
5.2	Automatic segmentation general components of thigh MR image. (a) Original MR image. (b) Adipose tissue (binary). (c) Muscles, femur and marrow in grayscale. (d) Absolute muscles (binary). (e) Relative muscles area (binary). (f) Smoothed femur cortex (binary). (g) Smoothed bone marrow (binary).	94
5.3	Major primary variance modes of the shape model training dataset. (a) Training dataset of 85 aligned shape models. (b) Variation eigenmode of mean value, -2λ . (c) Mean shape of shape model. (d) Variation eigenmode of mean value, $+2\lambda$	96
5.4	Outputs of pre-processing technique to mid-scan. (a) Original mid-scans. (b) Pre-processing results.	97
5.5	Feature extractions by phase congruency. (a) Top row - Pre-processed images. (b) Bottom row - Phase congruency results.	98
5.6	Stacking of phase congruency images. (a) Top row - Application of phase congruency to images (images were taken from a 26-image dataset, with the first 7 middle images (image no. 10-16) were used and processed) . (b) Middle row - i) The mid-scan phase congruency image (image no. 13). ii) Stacking of 3 mid-scan images (images no. 12-14). iii) Stacking of 5 mid-scan images (images no. 11-15). iv) Stacking of 7 mid-scan images (images no. 10-16). (c) Bottom row - Default 64-color “parula” colormap (heatmap) of (b).	99
5.7	Image morphological methods. From left to right - sample of 5 phase congruency stacked images; object shrinking; and image skeleton process.	100

5.8	Scaling and translation process of ROI mean shape model. (a) Pre-processed MRI mid-scan image with ROI mean shape model (green boundary). (b) Major and minor axes of ellipse equivalent for both mid-scan, in red; and shape model (SM), in green. (c) SM scaling process, with s equal to 0.9. (d) SM translation to a fourth of mid-scan major axis length.	101
5.9	Segmentation results of (a) Under segmentation (due to black pixels (low energy) correspond to femur region). (b) With the application of FCM smoothing process.	103
5.10	Sample of segmentation results on variations of s . (a) Top row: $s = 1.1$, accuracy = 95.18%. (b) Bottom row: $s = 1.0$, accuracy = 90.99%.	107
6.1	Pre-processing techniques. (a) Original MRI image. Unwanted background removal by: (b) Adaptive thresholding. (c) Otsu's thresholding. (d) Bias field correction application on original image (a). Unwanted background removal by: (e) Adaptive thresholding with bias field correction. (f) Otsu's thresholding with bias field correction.	110
6.2	MRI outputs. (a) Noise patches in intensity inhomogeneity. (b) Limb movement during scans. (c) Muscles compositions in elderly and echo effects.	111
6.3	Summary of semi-automatic segmentation process.	112
6.4	The operational flow of proposed segmentation method.	113
6.5	Manual segmentation results by OsiriX.	114
6.6	Segmentation results by proposed method.	115
7.1	The output domain of a convolutional neural network (CNN) in MRI of the thigh. (a) Object classification - the task of identifying that there is a quadriceps in the image. (b) Object localization - class label including a bounding box to show the location of the object (ROI). (c) Object segmentation - class labelling and outlining the predicted region. Note: other major domain in CNN includes object detection - localization of multiple of objects (objects can be from different classes) in an image.	120
7.2	The network architecture of AlexNet [1].	122
7.3	The network architecture of VGG-16 [2].	123
7.4	The network architecture of PSPNet [3].	124
7.5	Preparation of MRI Datasets.	126
7.6	Transfer learning procedure of deep CNNs to obtain optimized weights initializations. Three fully connected layers of CNN were removed and replaced by three convolutional layers, making the pre-trained model fully convolutional.	128
7.7	Image filtering and morphological to the semantic segmentation image. (a) Ground truth label (in white) with original MR image. (b) Semantic segmentation result by FCN8-s model (for this sample: with Adam solver in WD dataset). (c) Post-processing result of (b).	136
8.1	Configuration of MR images from three different subjects, from proximal, mid-scan and distal scans.	145
8.2	Atlas-registration method. (a) Atlas (reference) image. (b) Atlas region. (c) Atlas registration. (d) Moving (targeted) image. (e) Overlay of atlas boundary on moving images. (f) Ideal atlas-registration process.	147
8.3	Procedure of isolating the "suspected area" and spline curve estimation.	148

A.1	3-D volume rendering by OsiriX for one thigh MRI dataset.	152
A.2	Corresponding dimensions of pixel (in red arrows from MATLAB) and cm (in blue arrows from OsiriX).	152

List of Tables

4.1	Summary of results for 200 processed MRI images.	85
5.1	Average segmentation accuracy for Method 1-4.	104
6.1	Average segmentation accuracy for 416 processed MRI images	116
6.2	Average annotation time per image (in seconds)	116
7.1	Mean performance index for all models on mid-scan (MD) and whole-scan (WD) datasets.	132
7.2	Mean <i>JSI</i> performance index for different solver types used in FCN-8s on mid-scan (MD) and whole-scan (WD) datasets.	133
7.3	Mean <i>JSI</i> performance index of all models on mid-scan (MD), whole-scan (WD) and all-scan (AD) datasets.	134
7.4	Mean processing (testing) time (in second) for all models to semantically segment an ROI of thigh MR image.	134
7.5	Mean <i>JSI</i> performance index of FCN-8s model on mid-scan (MD), whole-scan (WD) and all-scan (AD) datasets, without and with post-processing, respectively.	136
7.6	Average time taken for training and validation procedures for all models used (in minute) for an image. Red and green boxes depict model that was trained in Machine 1 and 2, respectively.	137
8.1	Mean performance index of research findings from Chapter 4, 5, 6 and 7, including manual “ground truth” procedure.	144

Abbreviations

1-D	One Dimension
2-D	Two Dimension
3-D	Three Dimension
Adam	Adaptive moment estimation
AI	Artificial Intelligence
ANN	Artificial Neural Network
CAT	Computerized Axial Tomography
CNN	Convolutional Neural Network
CSA	Cross Sectional Area
CT	Computed Tomography
DAC	Digital-to-Analog Converter
DC	Direct Current
DICOM	Digital Imaging and COmmunications in Medicine
DSC	Dice Similarity Coefficient
DWI	Diffusion Weighted Imaging
FCM	Fuzzy C-Mean
FCN	Fully Convolutional Neural Network
FLAIR	FLuid Attenuated Inversion Recovery
Gad	Gadolinium
GLR	Generalized Log-Ratio
GPU	Graphics Processing Unit
GVF	Gradient Vector Force
ILR	Isometric Log-Ratio
ILSVRC	ImageNet Large-Scale Visual Recognition Challenge
IMAT	Inter Muscular Adipose Tissue

JSI	J accard S imilarity I ndex
MAP	M aximum A P osteriori
MCC	M atthews C orrelation C oefficient
mIoU	m ean I ntersection o ver U nion
MR	M agnetic R esonance
MRF	M arkov R andom F ields
MRI	M agnetic R esonance I maging
NAG	N esterov's A ccelerated G radient
NHS	N ational H ealth S ervice
nms	n onmaxima supression
PCA	P rinciple C omponent A nalysis
PET	P ositron E mission T omography
ReLU	R ectified L inear U nit
RF	R adio F requency
RGB	R ed G reen B lue
ROI	R egion O f I nterest
RW	R andom W alker
SAT	S ubcutaneous A dipose T issue
SD	S tandard D eviation
SGD	S tochastic G radient D escent
SM	S hape M odel
SNR	S ignal-to- N oise R atio
SVM	S upport V ector M achine
SPECT	S ingle P hoton E mission C omputed T omography
TE	T ime of E cho
TR	T ime of R epetition
UGM	U nidirected G raphical M odel
VGG	V isual G eometry G roup
VOC	V isual O bject C lasses
WHO	W orld H ealth O rganization
ZSI	Z ijdenbos S imilarity I ndex

Chapter 1

Introduction

This chapter features the establishment of the research, the importance of the research, the goal and contributions upon its successful achievements.

1.1 Research Background

There are over 600 skeletal muscles in the human body [4]. Muscles play important roles in: 1) thermoregulation by generating heat as a product of their contraction and metabolism; 2) whole-body energy balance by utilizing and storing fatty acids, glucose and amino acids consumed as part of the diet to provide energy for movement and other cellular processes; and 3) paracrine and endocrine functions as they release growth hormones and other factors into circulation. However, the most important functions of muscles are to provide support for posture and to make movement possible by applying force while shortening to bones and joints.

The muscles of the arms and legs principally allow us to interact with and move around the environment. Of all the limb muscles, those located in the thigh are amongst the largest and most powerful. The thigh is divided into three main compartments: 1) the quadriceps muscle group on the anterior aspect; 2) the adductor muscles that bring the thigh toward the mid-line and rotate it (on the medial section); and 3) the hamstrings on the posterior aspect.

Skeletal muscles are highly adaptable to their habitual use: regular exercise will induce muscles growth, while sedentary living or other types of disuse leads to muscle atrophy.

Indeed, any disease or other long-term condition causing low muscle mass and strength can result in mobility impairments. The normal ageing process is one such condition. During ageing, physical activity levels decrease [5] and the consequent disuse alongside other biological processes such as hormonal and/or other endocrine changes contribute to muscle wasting [6]. Figure 1.1 shows a transverse plane cross section of the mid-thigh in a young (Figure 1.1(a)) and older (Figure 1.1(b)) man and it is clear that the muscles of the older man are smaller and that the thigh contained higher levels of adipose tissue.

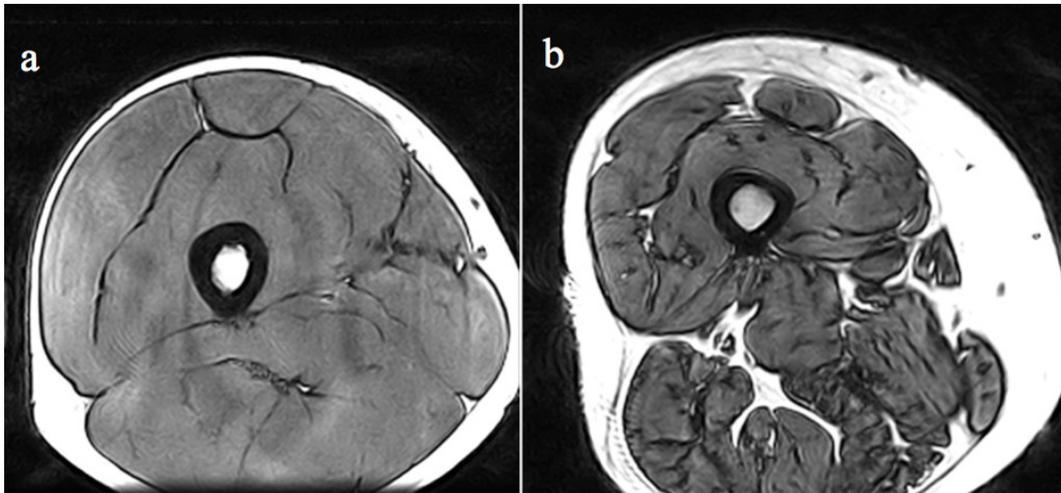


FIGURE 1.1: Magnetic Resonance Imaging (MRI) scans from (a) Young and (b) Elderly.

Technological advances in medical imaging over the past three decades have led to greater accessibility to advanced imaging techniques in clinical practice and research settings. Such technologies include magnetic resonance imaging (MRI), dual-energy X-ray absorptiometry, ultrasonography and computed tomography. They are widely used by clinicians to scan body segments to diagnose injury or disease. These imaging techniques are also used by researchers interested in studies of human anatomy and physiology in health and disease across the lifespan. These techniques produce precise, high quality distinctions between different tissue and cell types.

Over the past few years, physiologists from the School of Healthcare Science, Manchester Metropolitan University, have coordinated several large-scale research projects to investigate the causes of atrophy and weakness of leg muscles in old age and how this leads to problems in daily life activities. As part of this process, MR images were collected from around 300 people aged between 18 and 90 years¹, with about 40% of the subjects are from the age group of 18 to 40, 50% aged between 41 and 64 and the remaining 10% for the chronological age of 65 to 90. The MR images can be analyzed for muscle, fat and

connective tissue content by examining pixel areas and pixel intensity, and quantification of the results reveals differences between people and the effects of ageing or disease [7–9].

A very interesting and novel line of enquiry has revealed that leg muscles of older people begin to accumulate fatty adipose tissue. There is very little information currently available showing how adipose tissue accumulates in leg muscles. Quantification of the changes in adipose tissue with ageing coupled with the more detailed analysis of muscle biopsy samples could help to identify the regulatory mechanisms leading to loss of muscle mass and strength in older age because fatty adipose tissue can emit proteins that promote the breakdown of surrounding tissue [10].

As with any medical image, there is a need to develop analysis frameworks that accurately quantify the different elements, in the case of MRI thigh images the elements of interest are the muscle tissue, adipose tissue and femur bone. Manual segmentation of regions of interest (ROI) within slices is well established, but time consuming, laborious and prone to intra and inter-operator variability [8, 11]. Several studies have developed automated or semi-automated analysis techniques to study MRIs of brain, spinal cord, heart and other internal organs, but few have developed such techniques to study thigh muscles [9, 11–13]. The study of thigh muscles is a priority due to their importance in locomotion, mobility and metabolism and the extent of their deterioration in ageing and disease. One of the major muscle group of the thigh is the Quadriceps (muscle group consist of four muscles; vastus lateralis, vastus medialis, vastus intermedius and rectus femoris), which are all located on the anterior aspect of the thigh. By segmenting the biggest cross sectional area (CSA) of the quadriceps, which usually located within the middle-region of thigh MRI, quantification of muscle strength and muscle quality of the thigh can be measured. Estimation of the severity of a disease or an injury also possible to support further studies or assessments of the thigh [14–18].

A reliable semi-automatic or fully automatic segmentation framework is critical for large-scale studies to reduce analysis time compared to manual segmentation and improved reproducibility. However, semi-automatic or fully automatic segmentation often proves difficult due to image artefacts, noise, echo, overlapping of pixel/voxel intensities and non-uniform 2-D pixel intensity. These image deficiencies or low signal-to-noise ratio

¹For details of MRI datasets, please refer [Section 3.4](#).

(SNR) outputs are known to be affected by inherent effects of pulsed radiation on MRI radio frequency (RF) coils, gradient pulse eddy current effects from the MRI device, the subject's internal variations of electro-chemicals and biochemistry, fluid distribution and tissue density as well as issues related to body movement during scans. Fully automatic segmentation of anatomic structures of MRI thigh scans is further complicated by the potential lack of precisely defined muscle boundaries and issues related to intensity inhomogeneity or bias field across an image. Even for a well-trained and well-experienced physiologist, it can be challenging to differentiate muscles borders, especially skeletal muscles that form a same muscle unit, such as quadriceps and hamstring. This thesis primarily demonstrates methods of semi- and fully-automatically segmenting MR images of quadriceps muscles and corresponding femur cross-sections, which can then be used by researchers and clinicians to facilitate further analysis of the underlying (patho) physiology.

In recent years, numerous segmentation algorithms have been developed and applied to a small number of images, but the issue of compatibility of these algorithms to one another has not been properly documented. A popular technique for image segmentation proposed in [19], derived from the k-means method, is the fuzzy c-mean (FCM) algorithm. It performs fuzzy partitioning through iterative optimization of the objective function with update to cluster centers. However the algorithm does not incorporate any information about spatial context, which causes it to be sensitive to noise and artefacts [20]. To overcome this issue, various modifications that integrate local spatial information as constraints in the objective function were proposed in [20]. Authors in [21, 22] described a comprehensive version of the robust fuzzy clustering method for noisy image segmentation. Additionally, in [23], the authors suggest a hybrid intelligent color segmentation method, which features the combination approaches of region growing and clustering methods.

1.2 Brief Overview of Work Developed in This Thesis

The proposed semi-automatic segmentation method in this thesis is based on adaptive and automatic thresholding, followed by convex hull pixel identification for unwanted background removal, manual heterogeneous digital marker line drawing and analysis of

the ROI. The proposed algorithm is not iterative and hence has less computational complexity and uses a simple histogram concept to trigger ROI segregation. Nevertheless, optional inclusion of a hybrid method described in [23] will be considered accordingly to extract the optimum feature of segregated ROI.

Next, the thesis describes a novelty framework for automatic muscle border detection that distinguishes the quadriceps muscles from the other muscles in thigh MRI scans. At this stage, analyses were done on collections of mid-scan images (of a dataset), due to consistency in its tissue compositions. Mid-scan images are also less likely to be affected by MRI machines eddy current effect. We first measured a statistical shape model of the quadriceps muscles over a training set of ROI samples. For more robust statistical results and an average shape model that represents the best universal ROI shape, this training set is loaded with manually segmented mid-scan MR images of the thigh. Boundary localization of all manually segmented ROIs were extracted and non-parametric distance map based on Euclidean metric that represents these ROIs correspondent to background pixels were generated and piled up into a single matrix to eventually obtain the mean shape of the ROI by applying principle component analysis (PCA). Providing proper curve representation of the ROI shape model (obtained from the result of mean shape of PCA) based on information from image to segment (mid-scan image of one MRI dataset) is desirable and results in effective seeding points for optimum curve deformation process later. In doing so, the next stage covers the foundation of image transformation of ROI shape model, which are part of the rigid or affine image registration, and a subset of the more general transformations [24]. The final stage involves a combination of algorithms to extract the muscle border between quads and the hamstring.

To further amplify the significance of the research, an enhancement of MRI human thigh muscle segmentation by template-based framework is developed. This automatic segmentation framework is applied to all images in a dataset to extract ROI, based on the information of previously segmented mid-scan MR image (details of fully-automated frameworks are covered in Chapter 5 and 7). The final curve that fits ROI border is obtained and smoothed by an energy minimizing function and the application of a classification algorithm.

Finally, the application of artificial neural networks (ANNs), a concept inspired by the

human nervous system, is adapted to the automatic segmentation of the ROI in the domain of semantic segmentation. In principle, semantic segmentation in ANN is a process of identifying and analysing the role of each pixel in the image. In relation to a deep convolutional neural network (CNN), a collection of input images and its corresponding pixel-labelled images are needed to train a semantic segmentation model, where every pixel value in a pixel-labelled image represents the class label of that pixel. This procedure continues with the validation process and the final testing procedure generates the semantic segmentation results. ANN is regarded as a highly robust technique due to the minimal to no dependency on specific regulators and rules and can be operated without certain assumptions, unlike image processing approaches prior to these ANNs.

1.3 Aims and Objectives

The overall aim of the research is to design and develop an automated analysis framework of thigh MR images, with the segmentation method that is focused on the quadriceps muscles, femur and bone marrow.

Objective 1. The first objective is to evaluate the strengths and limitations of the current methodology of image analysis and segmentation in the scope of image processing and computer vision in medical imaging modalities, especially MRI. This includes examining different frameworks, statistical and probability analyses, including state-of-the-art methods, and compare the findings or proposed algorithms with current benchmarks such as pattern recognition, pixel classifier, deformable models, artificial neural network, and/or atlas-guided approaches to yielding optimum solution for automatic segmentation of MR images of the thigh.

Objective 2. The second objective is to design a reliable segmentation and quantification framework for MR image of the thigh. Our ROI is consists of quadriceps muscles, femur and bone marrow. The framework should provide accurate results in a short time with minimal user input into the segmentation task.

Objective 3. The third objective is to develop an automated segmentation and quantification framework for MRI of the thigh. The framework should provide acceptable results without any supervision.

Objective 4. The fourth objective is to propose an ROI segmentation and quantification framework for serial MR images within the same dataset (subject).

Objective 5. The fifth objective is to design a semantic segmentation framework for MR image of the thigh based on current machine learning approach of deep learning algorithm.

It is expected that with accurate characterization and localization of different tissues (muscles, adipose or pathological) in the quadriceps muscle, a dynamic (reliable, fast and precise) image analysis can be accomplished. This could also reduce the time taken for output image analysis in research, but also enhance the understanding of issues related to muscle movement, force distributions, pathological effects to human anatomical regions, and increase the reliability of medical diagnosis and patient treatment.

1.4 Research Contributions

In this thesis, several contributions related to the statistical and probabilistic analysis for image segmentation frameworks are presented. The combination of algorithms developed as part of this route have advantages of less-reliance on object boundaries while being flexible and adaptable to orientation and size of the image to be segmented, and provides a simple sequential MR images segmentation framework with no topological issues at excellent processing rate.

Our first contribution is developing a semi-automatic segmentation framework. This is interactively done by a user-defined for border delineation between ROI and the remaining thigh muscles in an MRI scan. This process combines several customized algorithms with the capability for error correction, which is highly convenient, especially for first time users or when analysing complex MRI datasets (usually MR images from old subjects).

Our second contribution is proposing the association of several approaches, including prior knowledge of the shape by the spatial reduction method of Principle Component Analysis (PCA), with optimization of localization and probabilistic analysis for muscle border detection by a dynamic shape model.

As a third contribution, a fast and accurate automatic framework was established to determine the minimal energy function for boundary detection of the ROI to series of MR images in a single MRI dataset, building upon the prior model from previously segmented mid-scan images (of the same dataset). This approach is guided by geodesic energy minimization functional of prior model as initial normalization seeds combined with modification to classification function to ensure the curve fitting remains outside or on top of zero-energy region on successive images.

Finally, the fourth contribution of this research is validating the application of artificial neural networks (in the form of deep convolutional networks) in semantic ROI segmentation domain of MRI of the thigh. The transfer learning approaches are pre-trained and personalized, so that the initialization values of the weight are applicably adjusted. This weight optimization procedure improves the general feature maps classifications of the model and in turn, delivers best performance in terms of segmentation accuracy and processing time.

1.5 Thesis Outline

The rest of the thesis is structured as follows:

1. *Chapter 2*. This chapter includes a literature review covering the physiological, anatomical and clinical interests in thigh MRI as well as the basic engineering behind medical imaging modalities and previous research into image segmentation of skeletal muscles from MRI.
2. *Chapter 3*. This chapter covers general methodologies related to image segmentation, including the *state-of-the-art* techniques, research materials and MR images and datasets acquisition.

3. *Chapter 4*. This chapter mainly describes the technicality of the semi-automatic segmentation procedure of thigh MRI and a basic description of the manual annotation process for ground truth images.
4. *Chapter 5*. This chapter describes the algorithms employed for automatic segmentation of MR mid-scan image of the thigh. The chapter starts with automatic segmentation of each tissue (component) of thigh MRI, followed by automatic segmentation of the ROI.
5. *Chapter 6*. This chapter develops a combination of frameworks that lead to automatic segmentation of the sequence of the MR image (from the same dataset as was previously segmented for mid-scan images as a prior model image).
6. *Chapter 7*. This chapter explains in depth utilization and validation of applicable transfer learning approaches on semantic segmentation of MRI of the thigh.
7. *Chapter 8*. This chapter concludes the thesis and includes the description of future works and further recommendations related to and beyond the research.

Chapter 2

Research Foundation

This chapter highlights cross combination of principal literature and technology in human physiology, medical engineering and medical image processing in computing.

2.1 Introduction

Health remains as one of the main concerns in human civilization. Health, as defined by World Health Organization (WHO) is a “State of complete physical, mental, and social well being, and not merely the absence of disease or infirmity”. It also can be defined as the general condition or state of the body, whether it is in good or poor condition.

Our environment and technology continually evolve. Major developments over the past century to motorised transportation, a shift towards increased urbanisation with the plentiful and cheap availability of calorie dense foods has transformed the way that the majority of people live, at least in westernised, developed countries. This has also had a consequence for human health and physiology, most obviously being associated with an unprecedented increase in sedentary living and progressive, year-on-year increases in obesity and related disorders such as type II diabetes mellitus and cardiovascular disease [25]. Paradoxically, the increased prevalence of these major diseases and related co-morbidities has not decreased life expectancy; the reverse is true. Advances to health and social care systems, medicine and pharmacology has led to progressive increases in life expectancy, although worryingly, people are not living healthier for longer, rather they are surviving for longer with chronic underlying disease. This has vastly increased

the costs associated with looking after older people and others in society at high risk of disease (discussed in [26]). The mean life expectancy for women in the UK at birth in 2008 was 82 and in men it was 78 years. By 2013, this had risen to 83 for women and 79 years for men. However, the mean “healthy” life expectancy had changed little, or even decreased over this same period; for women in 2008 it was 66 and in men it was 65 years, but by 2013 it had fallen to 65 in women and 64 years in men [27]. Musculoskeletal disorders are the most common chronic, disabling conditions, affecting 14% of people aged over 65 years, while heart and circulatory conditions affect 10%; respiratory conditions affect 6%; endocrine or metabolic conditions affect 6% and mental disorders affect 4% of people aged over 65 years. Incidence more than doubles in the 10 years that follow retirement, with 30% of over 75s reporting chronic musculoskeletal conditions; 32% report heart and circulatory conditions and 13% report endocrine or metabolic conditions [28].

With that in mind, in order to promote healthy culture and lifestyle that lead to the accomplishment of healthy individuals and society, awareness of recommended daily food consumptions are desirable [29], as is the application of appropriate physical activity routines.

The uses of medical equipment in diagnosing and treating patients have become essential and necessary. Advanced imaging technology, such as magnetic resonance or X-ray that are used to image the internal body and major organs, are generating vast numbers of medical images. Such equipment is readily available and accessible to healthcare practitioners for application with patients across developed countries. For example, across Europe in 2010 there were on average 10.3 MRI scanners and 20.4 Computed Tomography (CT) scanners per million population, although these numbers are underestimates because they do not fully include equipment available in the private sector [28]. In the UK National Health Service (NHS) alone during 2010, there were 41 and 76 MRI and CT, respectively, scans per 1000 population [28], which is about 2.5 million MRIs and 4.5 million CT scans in a single year. The images need to be interpreted to form a diagnosis and to inform appropriate treatment regimens. However, a criticism is often that the output images and image analysis from medical devices still lack precision, are time consuming and mostly operator dependent. They therefore lead to increases in healthcare spending. Further research is needed in order to get better signal-processing

from medical devices, whether it is for segmentation of region of interest, quantifying tissues or muscles volume, detecting localization of pathology, treatment and rehabilitation planning, computer-assisted surgery or even studying of anatomical structures.

The success of developments to improve automated image analysis could further boost the output accuracy of current techniques and enhance the time taken in processing the output images from MRI, in terms of segmentation, thereby reducing healthcare costs. This research is also expected to produce an automatic algorithm system and therefore reducing the dependability to human supervisions.

2.2 Human Physiology

By definition, physiology is a study that aims to understand the mechanisms of how living things work and emphasises body functions. Human physiology studies how our cells, tissues, organs and systems work together, how these biological tissues interact in order to maintain homeostasis. Physiology is sometimes referred to as the “science of life” that looks at living properties, from molecular basis of cell functions to the whole integrated behaviour of the entire body [30].

Anatomy on the other hand, describes and explores the structures of living things. Anatomy studies the form, while physiology investigates the function or in simple words, anatomy looks at what the living thing is, while physiology looks at what the living thing does [30].

Fundamentally, physiology and anatomy are extremely closely related disciplines and generally scholars in life science tend to study these two disciplines together.

2.2.1 Muscular System

The muscular system in general, is responsible for the movement of the human body, from as tiny as an eye blinking, involuntary or voluntary contractions of the diaphragm during breathing, blood circulation through vasoconstriction/vasodilation of blood vessels, to heart atria and ventricles contracting to eject blood, all so that movements and interactions with the environment can be sustained. Muscle contractions also produce heat, so are involved in thermoregulation and metabolism [31]. The skeletal muscle

system is under voluntary neural control and the muscles attach to bones via tendons. There are over 600 individual skeletal muscles and they collectively make up roughly half of a healthy person's body weight [32]. Each skeletal muscle is a distinct organ, composed of individual muscle fibres (roughly the diameter of a human hair) and each muscle fibre is surrounded by connective tissue and blood vessels (capillaries), as well as having a single connection to the nervous system via a motor neuron.

2.2.1.1 Muscle Types

1. *Visceral muscle.* Visceral muscle, also known as “smooth muscle”, is normally found inside of organs like the stomach, bladders, womb, intestines, and blood vessels. It is considered as the weakest of all muscle tissues, in terms of force production, however it is very effective in transporting substances, nutrients, hormones and even unwanted molecules and pathogens throughout the whole body. Visceral muscle is also known as involuntary muscle since it is controlled by the unconscious segment of the brain, meaning it cannot be directly controlled by the conscious mind. Physically, this muscle appears smooth and uniform when observed under a microscope, and this smooth appearance unambiguously contrasts with the banded physical of cardiac and skeletal muscles [33].
2. *Cardiac Muscle.* Located only in the heart, this muscle type is solely responsible for pumping blood throughout the body. Like smooth muscle, cardiac muscle tissue also cannot be controlled consciously, so it is classified as an involuntary muscle. While signals from the brain and exclusive hormones adjust the rate of contraction and rhythm, cardiac muscle stimulates itself to contract. This blood pumping power house's natural pacemaker is made of cardiac muscle tissue that encourages other cardiac muscle cells to contract. Because of its self-stimulation, this muscle type is considered to be auto-rhythmic or naturally regulated [34]. Its cells are normally striated-like array and appear to have dark and light stripes when observed under a microscope. The formation of protein fibres inside of these cells produces these dark and light striations. These patterns indicate that in general, cardiac muscle cell is much stronger than visceral muscle. Cardiac muscle cells are divided to an X or Y shaped cells firmly linked to one another by special connections called intercalated disks. These microbiological elements are built up of finger-like projections from two adjacent cells that interconnect and provide a

robust bond between the cells. The divided structure and intercalated disks allow the muscle cells to combat high blood pressures and the strain of pumping blood throughout a lifetime. These features also aid in spreading electrochemical signals rapidly from cell to cell so that the heart can beat as a unit [34].

3. *Skeletal Muscle*. It is the only voluntary muscle tissue in human body and can be controlled consciously. All physical actions that a person intentionally performs, such as writing, typing, speaking or walking, involves contractions of skeletal muscle. The typical role of skeletal muscle is to make movement possible by contraction (i.e. by moving parts of the body closer to the bone that the muscle is attached to). As the name suggest, these muscles are mostly attached to two bones across a joint, permitting those attached muscles to move parts of the respective bones closer to each other. Skeletal muscles also provide protection to bones and major blood vessels and nervous systems, contribute to posture and support and produce a major portion of total body heat [35]. Its cells were formed when many smaller progenitor cells lump themselves together to procedure long, straight, multi-nucleated fibres and its striation configuration is comparable to cardiac muscle. Some parts of these skeletal muscle fibres such as the thigh and gluteal muscle groups produce very high force and power movements [32].

Figure 2.1 shows configuration or structure of 3 muscle types and their common locality in human body.

2.2.2 Anatomy of Thigh Muscles

The muscles that enable voluntary movements of the limb are skeletal muscles and the muscles in the thigh are amongst the largest and most powerful because they are largely involved in upright locomotion, such as walking, stair negotiation and running. Anatomically, the thigh muscles are split into different sections: *Anterior muscles*, known as the quadriceps, that extend the knee joint and assist in hip flexion; *Posterior muscles*, known as the hamstrings, that flex the knee joint; *Medial muscles*, known as adductors that adduct and rotate our thighs; and *Lateral muscle*, known as abductors, that abduct the legs [35].

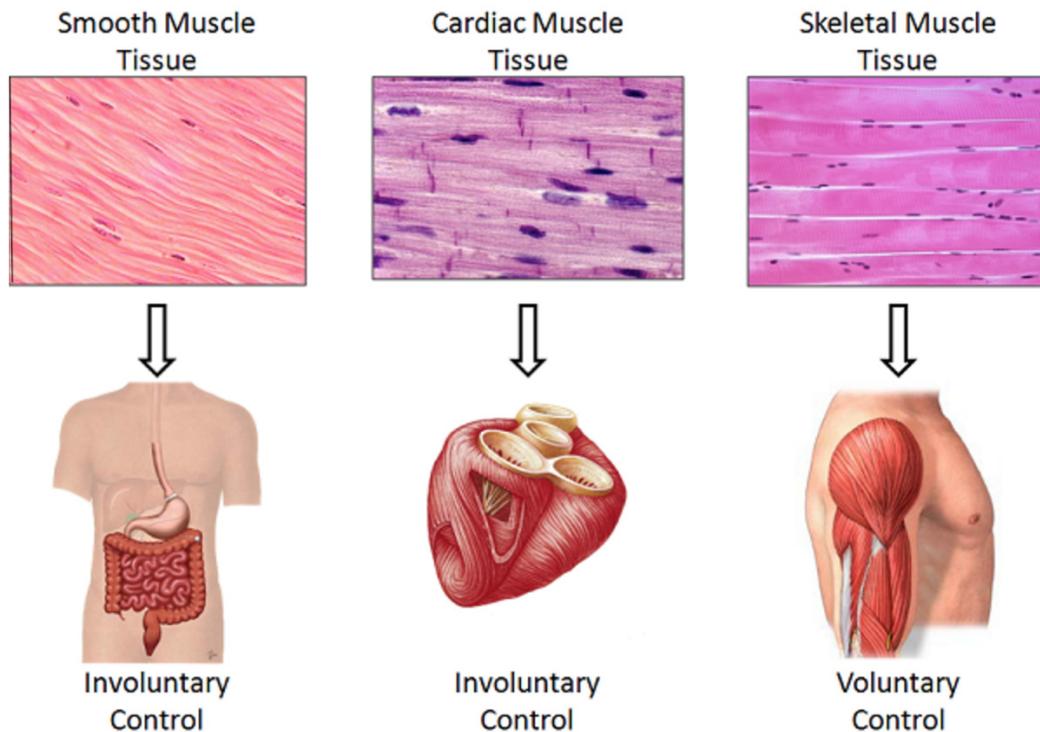


FIGURE 2.1: 3 different type of muscles in human body. *Courtesy: www.study.com*

2.2.2.1 The Anterior Thigh Muscles

The main anterior thigh muscles are the quadriceps, which includes the vastus lateralis, vastus intermedius, vastus medialis and rectus femoris. However, other anterior thigh or hip muscles include pectineus, sartorius and iliopsoas groups [35, 36].

1. *Iliopsoas*. The iliopsoas is a built up of two muscles that flex the thigh, in which one of those muscles, the psoas major, is a crucial component for posture support:
 - Psoas major originates on the twelfth thoracic and the five lumbar vertebrae and attaches onto the lesser trochanter of the femur and its contraction is regulated by the first three lumbar spinal nerves.
 - Iliacus originates on the sacrum, iliac crest, and sacroiliac ligaments and attaches onto the tendons of the psoas major and the lesser trochanter of the femur. Its contraction is regulated by the femoral nerve.
2. *Pectineus*. This muscle group originates on the pubis portion of the hipbone, termed superior ramus and attaches to femur on its pectineal line. Its contraction is regulated by the femoral nerve and allow adduction and flexion of the thigh.

3. *Quadriceps femoris*. This muscle group consists of four fibres that originate from different locations but sharing the same tendon, which attaches to the anterior tibia tuberosity via the patella. The femoral nerve regulates the contractions of all four muscle components to control knee extension.
 - Rectus femoris forms the central portion of the muscle group and originates at the anterior inferior iliac spine and just above the acetabulum of the hipbone. It is biarticular and contributes to both knee extension and hip flexion.
 - Vastus lateralis is the lateral-most muscle and originates at the linea aspera of the femur and the greater trochanter.
 - Vastus intermedius can be discovered deep to rectus femoris and vastus lateralis. It originates on the long femur shaft.
 - Vastus medialis is the most medial component of the unit and originates on the linea aspera and the line of inter-trochanteric.
4. *Sartorius*. This muscle attaches on the medial surface of the tibia and originates on the anterior superior iliac spine. The femoral nerve regulates its contraction and this allows abduction, flexion, rotation of the thigh and flexion of the leg at the knee.

2.2.2.2 The Medial Thigh Muscles

This muscle unit of the medial part of the thigh, responsible for leg rotation and bringing the thigh toward the medial or mid-line of the body [35, 36].

1. *Adductor brevis*. This muscle fibre originates on the pubis and attaches on the femur at linea aspera and pectineal line. Contraction is innervated by the obturator nerve and this permits thigh adduction.
2. *Adductor longus*. Originates on the pubis and attaches on the middle of the linea aspera of the femur. Contraction regulates by the obturator nerve and as above, this also permits thigh adduction.
3. *Adductor magnus*. Originates on the ischial tuberosity and the pubis and attaches on the femur at the gluteal tuberosity, linea aspera, and the adductor tubercle.

Innervated by the sciatic nerve and the obturator nerve and stimulates thigh adduction and assists in thigh extension and flexion.

4. *Gracilis*. This muscle originates on the pubis and attaches on the medial tibia. Innervated by the obturator nerve and this allows thigh adduction and leg flexion at the knee.
5. *Obturator externus*. This fibre is originates at the membrane of the hip bone and obturator foramen, and attaches to the femur. Allows thigh rotation by the innervation of the obturator nerve.

2.2.2.3 The Posterior Thigh Muscles

This muscle unit is also known as *the hamstring* and formed by three posterior thigh muscles and allows leg flexion and thigh extension [35, 36].

1. *Biceps femoris*. Consists of short and long fibre heads. The short head originates on the linea aspera of the femur, and the long head originates on the ischial tuberosity. These muscles attach on the fibula at the lateral side. Both short and long heads are innervated by the fibular and tibial portion of the sciatic nerve, respectively. Its roles include leg rotation and thigh flexion and extension.
2. *Semitendinosus*. Originates on the ischial tuberosity and attaches on the superior part of the medial tibia. Contractions activated by the tibial portion of the sciatic nerve and this allows thigh flexion and extension and medially rotates the leg.
3. *Semimembranosus*. Originates on the ischial tuberosity and inserts on the medial condyle of the tibia. It works in conjunction with the semitendinosus with similar innervation resource and roles as semitendinosus.

2.2.3 General Effects of Ageing in Human Physiology

Coordination and contraction of muscles are stimulated (activated) by electrical signals from the brain that travel through the central nervous system and out via motor neurons to activate individual muscle fibres very precisely. Any variations in the bones, muscles

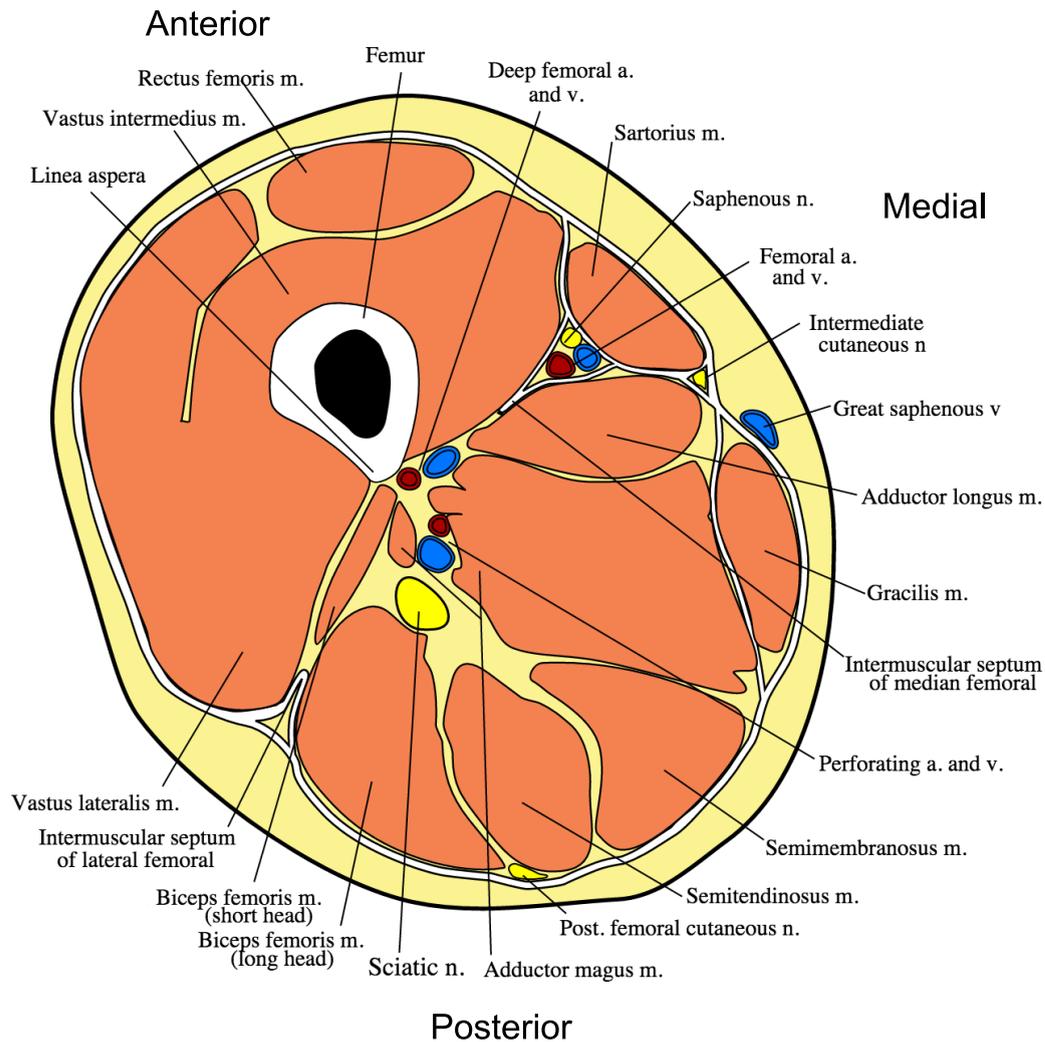


FIGURE 2.2: Cross section of thigh muscles. *Courtesy: <https://en.wikipedia.org>*

and joints due to ageing process may affect the posture shape and walking structure, typically leading to limitations to daily routine and slowing down movements [37].

With ageing, movement decelerates and may become restricted [38]. The walking pattern or gait cycle becomes shorter and slower, unsteady and with lesser arm swinging [38]. The changes of gait cycle may induce instability hence may escalate the loss of balance that may lead to falls and risk of injuries. Some elderly people have reduced reflexes. This is most often caused by declines in the muscles and tendons, as well as changes in the nerves. Metabolic changes within muscle due to decreases in capillary density and mitochondria (which are responsible for aerobic energy production) [39] contribute to lower energy levels [40], thus making the individual prematurely exhausted. Changes in hair and skin pigmentation and fragmentation are also normal [41].

Osteoporosis or porosity of bone density or bone mass is a common problem of old age, especially for post-menopausal women [42], due to declining hormones, particularly estrogen, and changed absorption of major minerals that regulate healthy bone, including proteins and calcium [42]. Bone to bone connections, such as those in the knee, hip and vertebrae that form the spine, are separated by a fluidal disk that cushions impact to reduce bone-friction. As we age, the height is reduced as the fluidity of the disks gradually decreased. The content of other critical minerals for a dynamic bone are also reduced, making each bone in the spinal cord thinner and affecting its structure and becoming compressed, curvy and bent, contributing to a slight loss of height [43]. Inhomogeneous branch of ageing bones or bone spurs is also an ageing effect and may promote tense and severe disorders due to nerve compressions on the vertebrae [44]. Joints breakdown may lead to deformity, inflammation, stiffness and pain. Such alterations in joint affect almost all elderly people. These changes range from minor stiffness to severe arthritis.

Average lean body mass also decreases, partly triggered by a loss of muscle tissue or muscle wasting/atrophy [45], but there is also a clear increase in body fat mass. Alterations to muscle composition often begin earlier for men compared to women [46].

Muscle fibres tend to shrink but not shorten, and as the result, muscle tissue is replaced slowly by fat accumulation [47]. Lost muscle tissues are also substituted with a hard and sturdy fibrous tissue and mostly visible in the hands, which may appear skinny and thin. Subsequently, muscles are less toned and unable to appropriately contract and may become stiff and inflexible with age and may lose tone, even with consistent exercise.

Reduction in muscle mass may also decrease body strength. Healthy and active ageing or “master” athletes may find that their performance improves in events that require longevity and endurance such as walking, but decreases in events that require short bursts of energy, such as sprinting [48].

Unintentional, rhythmic skeletal muscles movements (often termed muscle tremors) are also common in the elderly [49]. Elderly people who are less or not active at all may have paresthesias, a weakness or an abnormal sensations, normally tingling or pricking sense due to damage to peripheral nerves [49]. People who are unable to move on their own, or who do not stretch their muscles with exercise, may develop muscle or joint contracture syndrome, an unusual disorder of muscle or joint function. Figure 2.3 and 2.4 summarise the effects of ageing on human physiology, in general.

THE PHYSIOLOGICAL EFFECTS OF AGEING	
AGEING EFFECT	FUNCTIONAL EFFECT
Cardiovascular system	
↓ heart volume	↓ maximal SV; ↓ cardiac output
↓ capillary to fibre ratio	↓ muscle blood flow
↓ elasticity of the blood vessels	↑ total peripheral resistance; ↑ blood pressure
↓ sympathetic nerve stimulation of SA node	↓ maximal heart rate
Muscular system	
↓ muscle mass	↓ strength
↓ % of type IIa and IIx fibres	↓ strength
↓ size of motor units	↓ strength
Pulmonary system	
↓ elasticity of lung and chest wall	↑ work of breathing
↓ size of alveoli	↓ diffusion capacity; ↓ dead space
↓ number of pulmonary capillaries	↓ ventilation and perfusion
Metabolic system	
↑ % body fat	↓ joint and total body mobility; ↑ disease risk
↓ glucose tolerance	↑ risk of diabetes and heart disease

FIGURE 2.3: Effects of ageing in human physiology 1. *Courtesy:* <https://www.flickr.com/photos/sportex/5387348258>

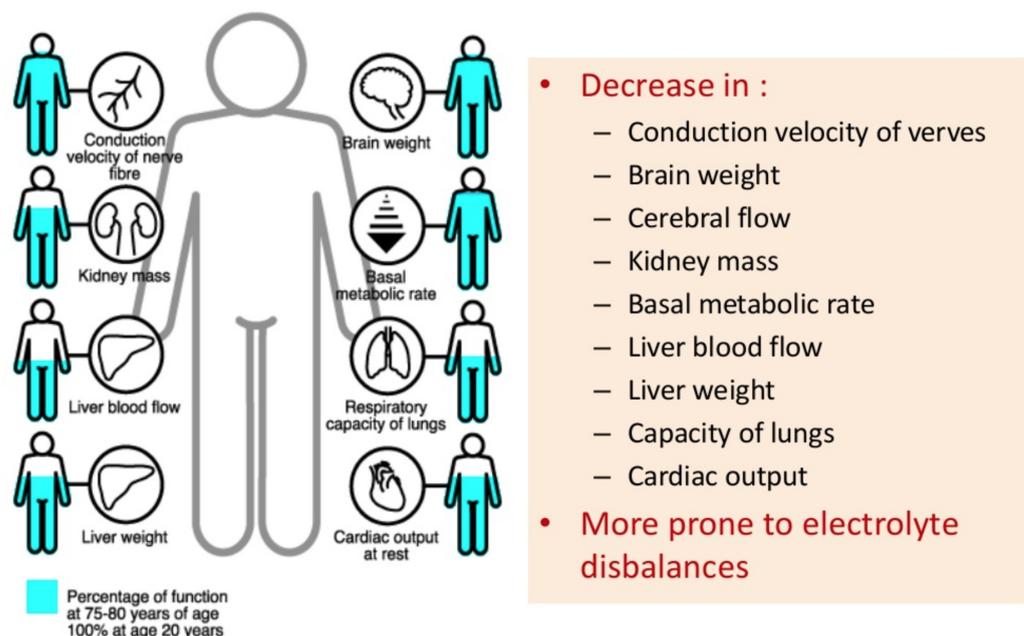


FIGURE 2.4: Effects of ageing in human physiology 2. *Courtesy:* <https://www.slideshare.net/Draveesoni/psychopharmacology-in-elderly>

2.3 Medical Imaging Modalities

The development of cutting-edge and up-to-date imaging technologies has significantly enhanced the overall quality of medical care available to both practitioners and end-users. Non-invasive imaging modalities allow a physician to make increasingly accurate diagnoses and may provide precise and reasonable modes of treatment. Modern medical imaging modalities are capable of acquiring data on internal anatomical structures, as well as mappings of physiological function.

Typical applications of medical imaging include diagnostic radiology, radiation therapy (usually applied to malignant cancerous cells), pathological study, laboratory medicine, surgery and internal organs/systems mapping. Altering energy type approaches and the acquisition techniques can obtain distinctive types of output images from one machine/scanner to another. Modalities referred to as the different modes of constructing/assembling images. Each modality has its own characteristics in medical diagnostic and treatment system. Listed below are the most common medical imaging modalities available on today's market.

2.3.1 X-Ray

Also known as projection radiography and is the first medical imaging technology available in modern medicine. X-ray was pioneered by Wilhelm Conrad Roentgen on 8 November 1895 and the first apparatus in capturing radiographic images of human internal anatomy [50]. Practitioners, who have had qualification to handle X-ray machine and appropriate training in interpreting diagnostic images, are called radiologist.

It is performed with a pulse of X-ray source on one side of the patient and a typically flat X-ray detector (photographic film or a computerized system or digital radiography detector system) on the other side. A brief duration of usually less than 1/2 second pulse of X-rays is emitted by the X-ray tube, a large fraction of the X-rays interact in the patient, and some of the X-rays penetrate through the patient and reach the detector, where an X-ray radiographic image is formed.

The distribution consistency of X-rays that enter the patient is adjusted by the degree to which the X-rays are removed from the beam by scattering and tissue density. The

attenuation properties of tissues such as bone, soft tissue and air inside the patient are different from one another, resulting in the heterogeneous distribution of X-rays that emerges from the patient [50].

Transmission imaging refers to imaging of X-ray emission from energy source (outside the body) to X-ray plate detector (the energy passes through the body and is detected on the other side of the body). Projection imaging refers to the case when each point on the image corresponds to information along straight-line trajectory through the patient.

X-ray radiographic imaging is suitable for a very wide range of medical indications, especially the diagnosis of broken bones, pathological changes in the lungs and cardiovascular disorders.

2.3.2 Computed Tomography

Computed Tomography (CT) scanner also known as Computerized Axial Tomography (CAT) scanner (or computer transmission tomography or computer tomography), is a technique with similar image acquisition as X-ray machines, where the image-capturing process is taken from X-rays that are transmitted through the body. These contain information on the body's elements in the path of the X-ray beam. By using multi-directional scanning of the object, multiple data can be collected. And this is done by configuring rotational X-ray beam sources around the object, to capture images from many different angles [51].

Although also based on the variable absorption of X-rays by different tissues, CT imaging offers a different form of imaging known as cross-sectional imaging. Figure 2.5 shows typical characteristics of X-ray and CT scan images.

A computational model of the attenuation of the matrix elements along rows and columns generate a cross-sectional image of the body. The use of computer is mandatory since the number of mathematical procedures required to construct clinically accurate and applicable images are very large. The data obtained from this computation analysis can be presented in a conventional form resulting in a 2-D CT scan image [51].

Processing power from a computer is used to put a set of regulations, in terms of timing, anode voltage and beam current. The high voltage direct current (DC) power supply

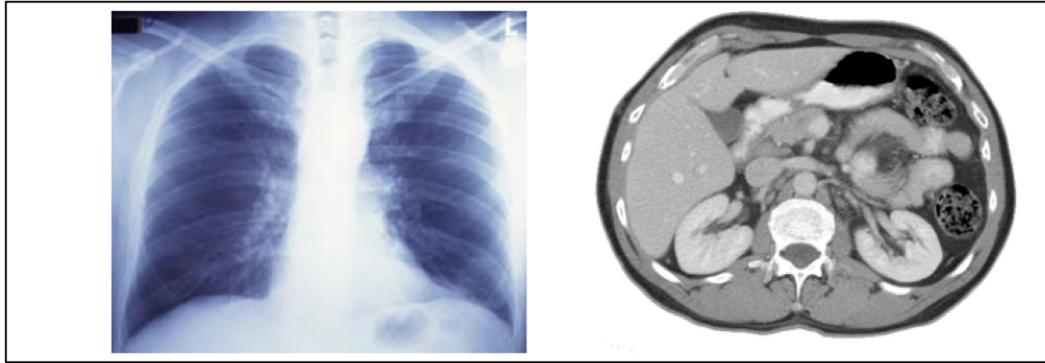


FIGURE 2.5: Medical imaging output characteristics of X-ray (left) and CT scan (right).
Courtesy: www.fda.gov

drives an X-ray tube that can be mechanically rotated along the circumference of a gantry. The patient lies in a tube through the centre of the framework. X-rays are beamed through the patient and are partially absorbed while remaining X-ray particles impinge upon several radiation detectors that are secured around the circumference of the gantry.

Reaction of the detector is directly related to the number of particles impinging on it and to the tissue density. Measurements based on data acquired from a complete scan are then made by the computer and later generates a live visual output image of a cross-sectional transverse plane of the patient on the screen. Output data also can be stored for future analysis.

Since the images from CT scanner are extremely detailed compared to output image of conventional X-rays, the computerized 2-D images can be reconstructed into 3-D mapping or model. This formation of virtual internal anatomy image would be beneficial for diagnostic and treatment, also for surgeons before, during or after operations. Nevertheless, without good medical reasons, physicians by default will try to avoid using CT scanner due to large amount of ionising radiation X-rays emission transmitted across the patient's body compared to traditional X-ray method.

2.3.3 Ultrasound

Ultrasound is another common diagnostic imaging technique by which ultrasonic signal is applied to describe the state of internal body organs. Bursts of high-frequency sonic

waves are transmitted from a “piezo-electric” or “magnetostrictive” transducer (receiver) through the skin and projected into the internal anatomy [52].

When this signal strikes a boundary between two tissues of different acoustical impedance, reflections (echoes) are returned to the transducer. The transducer transforms these reflections to an electric signal that are proportional to the depth of the boundary, which is later pre-processed (filtered), amplified and displayed on an oscilloscope. An image of the interior structure is formed based on the travelling time of total wave applied, the average sound speed transmitted and the energy concentration of the reflected waves [52].

The converter control unit receives signals of transducer position and synchronise pulse signals and generates X and Y position from reflected signals. This data is then fed to the scanner memory and stored. Digital-to-analog converter (DAC) is used to process this information and generate a real-time live image on the screen. Due to amount of noise, image processing techniques may be used to fully utilize ultrasound images [53].

2.3.4 Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is another technique for detailed soft and hard tissues image acquisition throughout the body without “ionizing radiation” like X-ray machine or CT scanner. Because of this ionizing radiation free feature, MRI has gradually become the preferred method for diagnosing anomaly or potential health issues in many different parts of the body [54]. In general, MRI produces images that can show differences between healthy and unhealthy tissues. Medical practitioners use MRI to examine major body parts such as the brain, heart, spine, joints (hip, wrist, knee, shoulder and ankle), abdomen, pelvic region, breast, blood vessels and bladder.

MRI employs powerful magnetic fields, radio waves, rapid changes in magnetic fields, and a processor to generate images [55]. Depending on application, different gradient coil systems produce a time varying, controlled spatial non-uniform magnetic fields in different directions. Similar to procedures in CT scanning, the patient is kept in this gradient field space, in a specially designed magnetically shielded room. There is also transmitter and receiving RF coils surrounding the space on which the image is to be

constructed. And respectively, the initiation of linear magnetic field gradient to the uniform magnetic field is applied to the patient [56].

The resonance frequencies of the processing nuclei depend on the positions along the direction of the magnetic field gradient. This produces a one-dimensional (1-D) projection of the structure. By taking a series of these projections at different gradient orientations using X, Y and Z gradient coils, a two- or three-dimensional (2- or 3-D) image can be acquired [57].

The slice of the image depends upon the gradient magnetic field and is controlled by a processor. This field can be positioned in three time-invariant planes of X, Y and Z. The transmitter provides the radio frequency (RF) signals and on the other hand the receiver coil picks up nuclear “resting state” magnetic resonance signals and fed to the processor. This signal is then processed and MRI scans (or images) are constructed by applying 2-D Fourier Transformation [51].

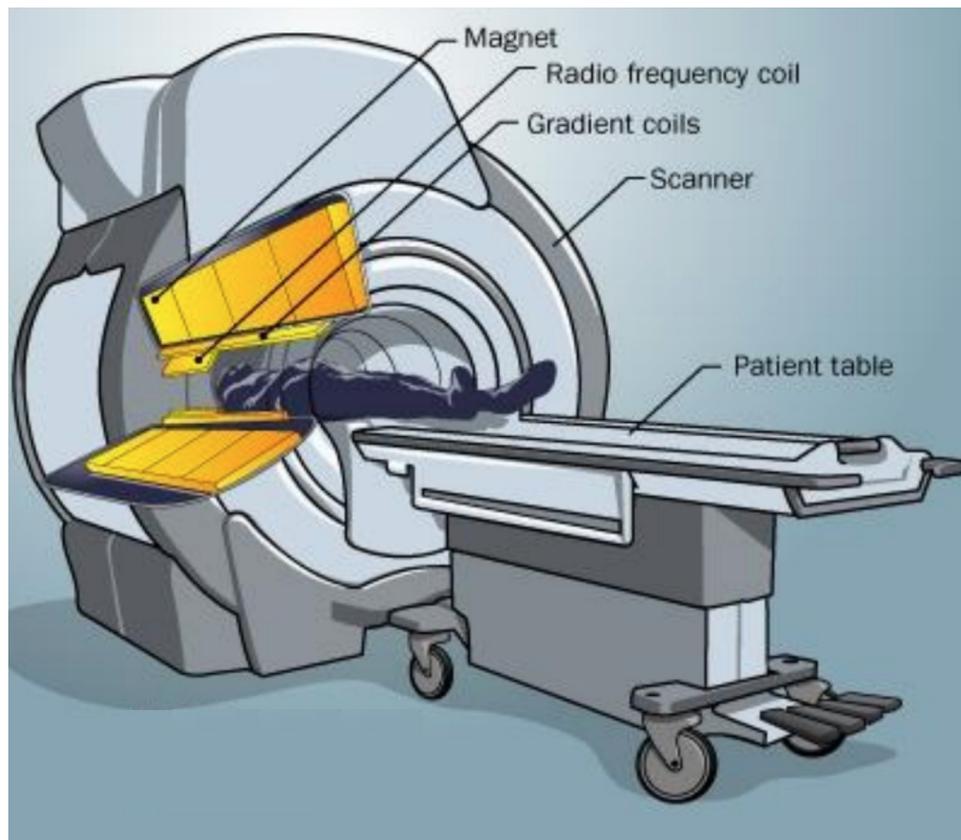


FIGURE 2.6: The cutaway of typical MRI machine. *Courtesy: www.howstuffworks.com*

2.3.4.1 MR Imaging Sequences

Image contrast is considered as the main objective in the acquisition of medical imaging, where certain contrast characteristics of anatomical structures will be emphasized before the acquisition process, depending on the output image analysis. This allows the operator to distinguish the output structures and define which tissues of the anatomy are healthy or abnormal. Without the use of any contrast enhancing agents, such as Gadolinium (Gad), image contrast in MRIs is natively superior than CT machines.

T1-weighted and *T2-weighted* scans are the most common MRI sequences, where T1-weighted MR images are produced by manipulating short pulse sequence parameters of echo time (TE) and repetition time (TR), while T2-weighted MR images are produced by using greater TE and TR times compared to T1. Other MR imaging sequences include Fluid Attenuated Inversion Recovery (FLAIR) and Diffusion Weighted Imaging (DWI). Figure 2.7 and 2.8 demonstrate the differences between T1- and T2-weighted MR images in human brain and spinal region, respectively.

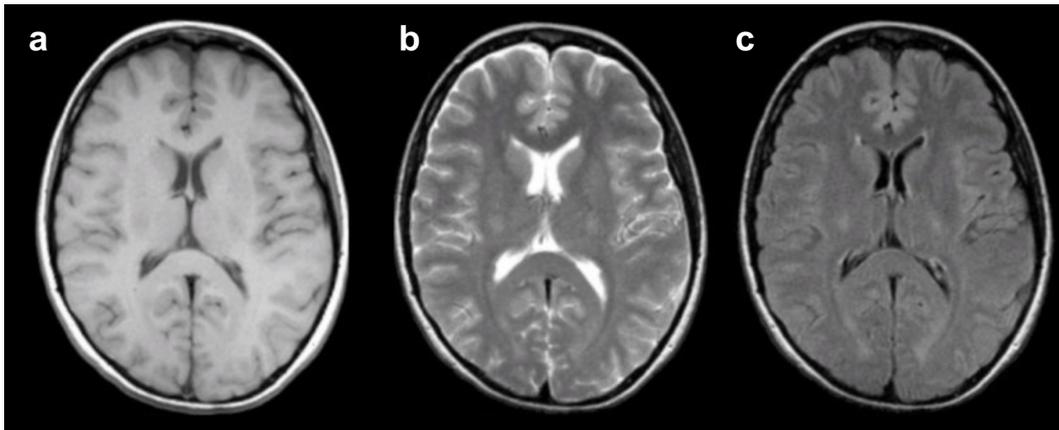


FIGURE 2.7: Differences of MR imaging sequences of the brain. (a) T1-weighted image. (b) T2-weighted image. (c) Flair image. *Courtesy: <http://casemed.case.edu/clinship/neurology/>*

2.3.5 Other Modalities

2.3.5.1 Positron Emission Tomography (PET)

This modality uses concept of coincidence detection (in the context of neurobiology) to image functional processes. Radioactive positron emitting isotope is merged with an organic substance, creating element, used as an indicator of metabolic utilization [58].

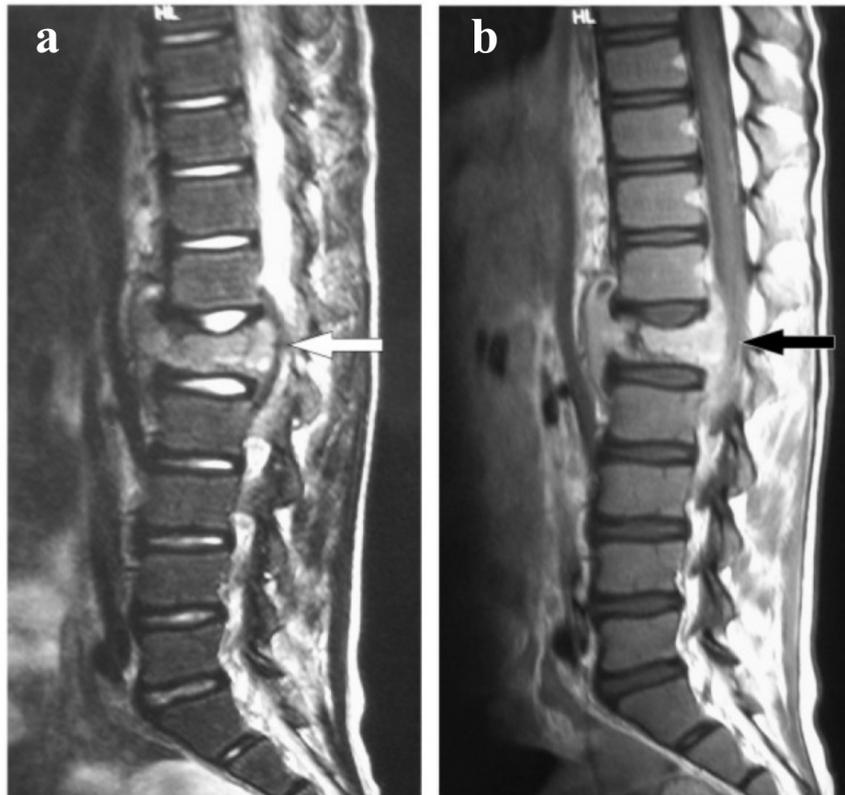


FIGURE 2.8: MR images of the spine. (a) T1-weighted image; white arrow indicates injured structure on the bone injury. (b) T2-weighted image; black arrow depicts bone oedema (feature of bone injury). *Courtesy: <http://www.lbgmedical.com/blog/mri-a-sophisticated-technology-for-imaging-soft-tissue-pathology/>*

Images of how tissues and organs are working throughout the body can be shown including irregularities such as tumour. PET images may also be viewed in conjunction with CT scan images to determine anatomical correlation. To optimize the image reconstruction that involves positron imaging, current CT or MRI scanners may incorporate PET technique, permitting PET-CT, or PET-MRI, respectively. This functional-anatomic hybrid imaging provides a very useful tool in non-invasive diagnosis and patient management.

2.3.5.2 Single-Photon Emission Computed Tomography (SPECT)

This 3-D tomographic modality is another type of nuclear imaging test that uses gamma camera from several projections to generate 3-D images. An advance dual detector head camera combined with a CT scanner, termed a SPECT-CT, provides better visualization and localization of functional organs [59]. In previously described medical

imaging modalities, energy is passed through the body and sensors convert the occurrence reactions. However, in SPECT imaging, the patient is injected with a radioisotope substance, radioactive gamma rays are then emitted through the body as the natural decaying process of this substance takes place. Gamma rays emissions are finally captured by detectors that surround the body. In other words, the patient is now the source of the radioactivity, rather than the medical imaging modality used.

2.3.5.3 Mammography

This niche imaging modality, also called mastography, uses low-dose ionizing radiation X-rays to examine the human breast and aids in the early detection of breast cancer, typically through detection at tissue microcalcifications. Mammography may also produce a missed diagnose (false negatives) of cancer. It is estimated that around 20% of the number of cancers are miss-diagnosed [60], mostly due to the cancerous malignant tissues that hidden by other dense tissue.

2.4 Previous Works on Image Segmentation of Human Thigh Muscles

Image segmentation has a long history, in which arguably, researchers at Stanford Research Institute Artificial Intelligent Group did its initial attempt in a vision project, dated back in 1970 [61]. This group used exploiting regions on image as data for image partitioning and a heuristic decision tree to trim down the explicable partition results. In 1972, an approach that was based on region growing process was described [62] and another of the earliest approaches of image enhancement and threshold selection using histogram was reported [63, 64].

For medical or clinical image segmentation in general, one of the earliest forms of digital image analysis is reported in [65], where a semi-automatic approach is carried out to outline and contain a region of interest of the left ventricle. Pattern recognition based method was applied [66] in 1973 to automatically detect tumour in radiograph images. Conclusively, according to [67], discussions on the history of medical image analysis are based on four findings:

1. First medical image analysis era, termed 2-D image analysis (pre-1980 until 1984).
2. Era of pattern recognition driven and knowledge based idea. Also the era of MRI becoming more popular and important as a source of image data (circa 1981-1990).
3. Era of the development of 3-D medical imaging (3-D MRI, SPECT and PET) and maturation of 2-D image analysis. Integration system, 3-D modelling and image analysis increasingly popular (1992-1998).
4. Advance image analysis and visualisation, focusing on solution to achieve the best accuracy, processing time and reliability (1999-present).

From 2002 until 2015, twelve segmentation attempts on the MRI of the thigh have been recorded (in year of chronicle published order), which include [7, 8, 68–77], where automation process were done by incorporating either one or a combination of these techniques: thresholding (intensity based or histogram modelling); classification (fuzzy c-means (FCM) being the popular approach or k-means); active contour; and/or region growing. These methods focusing on classifying muscles, marrow, femur, subcutaneous adipose tissue (SAT) and/or intermuscular adipose tissue (IMAT) as an individual component (or group), where after the implementation of suitable pre-processing algorithms (to remove/reduce noise or improve pixel's intensity), these individual components (or groups) can be routinely segmented.

The segmentation accuracies across all methods above are considered exceptional in general, with all methods achieved average accuracy of greater than 85% and the segmentation results can be seen statistically improved over time (2006/07 [70, 71] - 95.73%, by FCM; 2009 [7] - about 96%, by adaptive thresholding and histogram modelling; 2013 [76] - about 97%, by k -means clustering; and 2015 [77] - 96.8%, by snake active contour). Considering the application (or combination of applications) of such basic techniques, the system's average processing time across all methods are also recorded to be improved over time (2009 [8] - 52 sec per image, by FCM and active contour; 2013 [75] - about 5.21 sec per image, by region growing and 3-D intensity map; and 2015 [76] - 0.25 sec per image, by k -means clustering). Two major reasons that understandably contribute to the better mean average processing time are superior computer technology and hardware available on the market and the optimisation of algorithms and processing platforms from developers. Regardless, this simple, precise and fast automation process

may also assist significantly in quantification analysis for cross sectional area (CSA) or volumetric of MRI of the thigh, which are vital information to clinicians and researchers interested in early life development, ageing, disease and the effectiveness of rehabilitation programmes.

In 2012, an advance discrete optimization solution by a graph-based Random Walker (RW) was proposed by *Baudin et al.* [13], where a graphical model able to automatically determine appropriate seed positions with respect to different muscle classes was introduced. The process was established by combining frameworks of seeds sampling and graph edges; and Markov Random Fields (MRF) formulation that calculate the cost function form, unary potential, geodesic distance potential, and relative orientation potential of muscle structure and position. Authors reported that small muscles (within the thigh region) are prone to segmentation errors compared to the large ones and as a result, recommending the use of deformable registration approaches to alleviate the bias introduced from the rigid registration step, for future work. Next in [78], the same authors worked on the introduction of linear sub-spaces constraints within RW segmentation framework, where the novel knowledge or prior-based quadriceps segmentation was proposed, contributing to a slightly improved segmentation output accuracy, compared to [13]. The technique was developed based on Random Walker (RW) formulation, where the concatenation of all the node's (or seed's) probability is proliferated by block diagonal matrix (depending on the quantity of the label); combined with prior knowledge to the respective RW formulation above; and the design of low dimensional affine of implicit space within the spatial matrix. Results were presented in 2-D MR image cross-sectional area, with an average segmentation accuracy (by Dice coefficient) of 0.84 ± 0.08 and computational time of approximately 15 minutes per segmentation. Another effort from *Baudin et al.* including extension for the segmentation of whole muscles of thigh MRI [79], again by integration of similar RW graph-based with prior information frameworks as above, where statistical shape atlas was employed to represent prior knowledge. In addition to these frameworks, a balancing parameter was introduced to the formulation of nodal's total energy; with an addition of confidence map to adjust the influence of the contour's model. This configuration leads to an efficient iterative linear optimization and further enhanced the quality of segmentation accuracy, yielding an average segmentation accuracy (by DSC) of 0.86 ± 0.07 and processing time of around 5 minutes per segmentation. However, errors to muscle segmentation are still problematic,

especially to small muscle components (e.g. Gracilis muscle and tensor fasciae latae), due to large registration errors on the same muscle assumption.

Another fully automatic segmentation method for thigh MRI was demonstrated by *Andrews et al.* [11]. This technique provides a good accuracy for individual muscle segmentation of the thigh, by using energy minimizing probabilistic segmentation that indicates area of ambiguity. The automatic segmentation made possible with the frameworks that comprise of a probabilistic segmentation function representation of isometric log-ratio (ILR) transform; the shape space (relative locations of muscles) by principle component analysis (PCA); image alignment by using femur bone coordination as an anchor of the vector estimation; and energy construction strictly based on energy function over the shape space of an image. The results were presented in 3-D diagram and the method was capable of segmenting 11 muscles within the thigh automatically and achieved an average segmentation accuracy (by DSC) of 0.92 ± 0.03 across all images, but with undisclosed processing time per segmentation/image. A recent publication by *Andrews and Hamarneh* [80] presents comprehensive incorporation of probabilistic shape representation consists of adjacent object information and prior anatomical volume, called generalized log-ratio (GLR). The main benefit of GLR application is that it may produce a probabilistic segmentation that can be used to generate uncertainty information to aid subsequent analysis of different muscle components on thigh MRI. In other words, the GLR transformation can be designed to ensure statistical shape models capture variability in smaller structures properly. The method is designed based on four major steps. Firstly, GLR representation in the context of thigh muscle segmentation to encode details such as muscle size, adjacency information, and to train a statistical shape model over the space of GLR representations. Secondly, the introduction of pre-segmenting images technique for all components involve (fat, muscles and bone). The results are mainly used to perform alignment computation. Thirdly, random forest classifier training to assist in detecting intermuscular boundary locations. And finally, results from all steps above were integrated into a globally minimized convex energy functional by the means of primal-dual method to generate a probabilistic segmentation. As the system was mainly designed to tackle issues related to segmentation accuracy, the major drawback is that the average segmentation time took 50 ± 4.3 minutes per image to run. The average segmentation accuracy by using this method is 0.808 ± 0.074 (by Dice coefficient), which is quite promising, considering the procedure and analysis was done

to all muscles, individually, on MRI of the thigh.

In 1992, *Morrison et al.* [12] made an attempt on MRI thigh segmentation by using artificial neural network. The network model integrates a probabilistic neural network to facilitate the generation of probability estimates at each pixel for use in an iterative segmentation process. This preliminary study on MRI thigh segmentation using one of the simplest forms of neural network reported that the segmentation result is unaffected by the quality of the training dataset. Today, the concept of neural network has been revolutionised and diversely applied, including in segmentations of medical imaging. Recent research in medical imaging by the application of advance form of neural networks, represented in convolutional neural network (CNN) or fully convolutional neural network (FCN) (again, in year of chronological published order), include [81–85], which mostly covered MRI segmentations of region in brain, heart and breast.

2.5 Discussion and Conclusion

Naturally, the process of muscle wasting begins as we age (mostly due to hormonal and endocrine changes) and this deteriorating muscle tissue is replaced by the accumulation of fatty adipose tissue. The anterior thigh muscles are particularly susceptible to muscle wasting in ageing and disease, and this is of clinical significance due to these muscles playing a major role in everyday ambulation. Muscle wasting is a widely recognisable feature of ageing, but the process could be escalated at an early age due to bad dietary habits, sedentary living, and/or disorders related to metabolic syndrome or disease.

In addition, because of limitations of outputs from the medical imaging modalities (particularly the output images from MRI scanners), where images are shown in grayscale and have substantial noise, echo, overlapping of pixel/voxel intensities, various thigh tissue compositions especially conditions associated with muscle atrophy (wasting) described above, inter- and intra-slice inconsistency, indistinct muscle borders even for a well-trained physiologist, it can be challenging to differentiate muscles borders, especially skeletal muscles that form a same muscle unit, such as quadriceps and hamstring. Figure 2.9 demonstrates the MRI output profiles that defined the whole MRI database, with about 80% of the whole database is made up by images such as in Figure 2.9(f), 10% from Figure 2.9(e) and the remaining 10% from Figure 2.9(a)-(d).

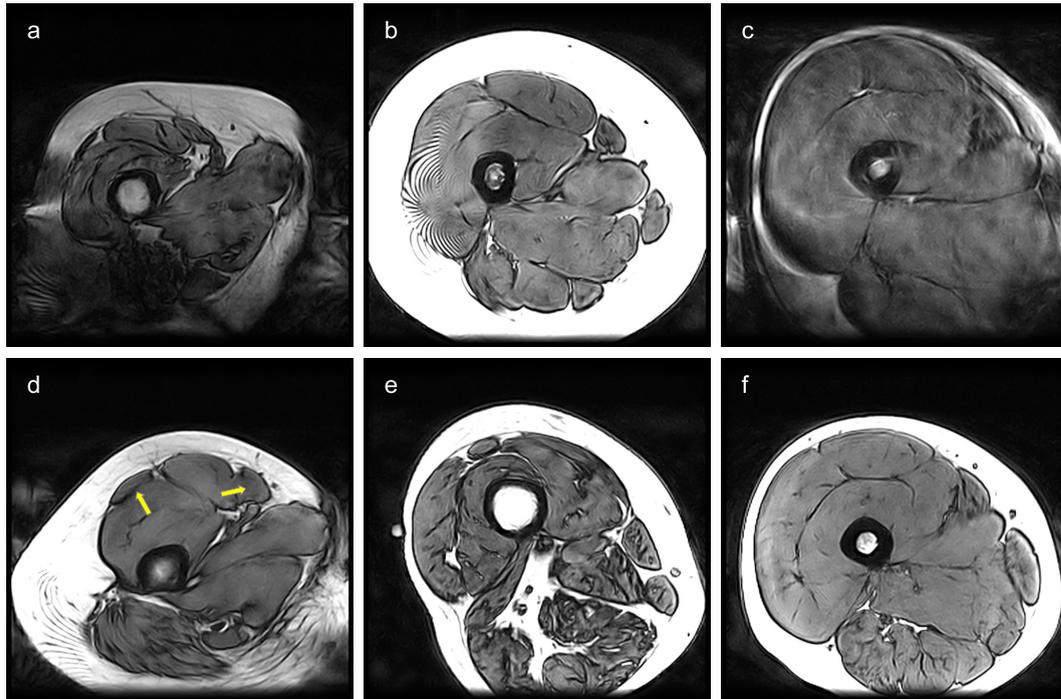


FIGURE 2.9: MRI output profiles. (a) Intensity inhomogeneity. (b) Echo. (c) Overlapping of pixels (limb movements during scan). (d) Yellow arrows indicate muscles that are not part of the region of interest (ROI) but appear to attach together with the ROI. (e) Thigh MRI composition in elderly. (f) Thigh MRI composition in young subject.

Image segmentation technique plays an important role in computer vision and particularly in analysis of medical image processing, due to its connection to the healthcare science. However, as described earlier, no single segmentation technique (or even combination of segmentation techniques) in medical imaging is capable to offer a complete solution in terms of precision, processing time and reliability. Nevertheless, approaches that are dedicated to certain applications and association of these methods usually provide better performance and results. In fact, it is important to understand that most medical image analysis efforts are fundamentally influenced and driven by the particular image datasets being utilized and the clinical or biological tasks that inspire the need for analysis of medical imaging.

In earlier medical image processing, two major procedures were commonly applied to simplify segmentation analysis. First a pre-processing phase, where it is mainly used for filtering and segmenting unwanted noise and to improved intensity inhomogeneity. The second phase usually associates with a classification process, either by interactively introducing seeding points or curves on an image or by adaptively designing an intensity or histogram model for segmenting different but apparent classes (or objects) in an

image.

At the time of writing, there are limited automatic segmentation frameworks published on MRI of quadriceps or thigh muscle components as a whole, with respect to the segmentation of individual muscle unit or muscle components in the thigh. And the template-based level-set semi-automatic quadriceps segmentation method [9] suffers from segmentation inconsistency and inaccuracy.

Since the results of this research will be critical in supporting clinical studies aiming to determine factors affecting muscle strength, structure and function in young and old age, it is crucial for the development of operator independence and accurate segmentation systems, without compromising much on processing time. We have identified that in order to achieve our goals, the ideal automatic segmentation framework, in terms of computer vision and image processing, needs to be incorporated with a combination of the following techniques: prior information, shape and probabilistic analysis, user-independence non-rigid registration.

However, the state-of-the-art methods described above, especially image processing based techniques, are not robust, due to their nature of reliance on specific regulators and rules, with certain assumptions. In contrast to artificial neural network (ANN) based algorithms, deep learning methods do not require such strong assumptions and have demonstrated superiority in region and object segmentation and classification, which suggests that the state-of-the-art innovation in MRI segmentation of human thigh muscles, may be achieved by adapting such approach. Moreover, the segmentation output performance from most recent applications of network models in medical imaging [86–89], in association with image processing technique [90] and in conjunction with the utilization of better and upgradable hardware, such as graphics processing units (GPUs), are remarkably accurate and fast.

From the literature, it may be concluded that for a development of a reliable and robust automatic thigh MRI segmentation system, it is difficult to achieve great segmentation accuracy with fast processing time in the same package. Researchers are generally advised to set the priority on the accuracy division and optimizing the processing time later.

Chapter 3

Methodologies in Medical Image Segmentation

This chapter presents the basics, fundamental and mathematical concepts behind algorithms in image processing that often linked and associated to segmentation, especially in medical imaging.

3.1 Introduction

With output images from selected medical imaging modalities (such as X-ray, CT or MRI scanners), segmentation of certain body segments or structures can be performed just using low-level image processing (such as image thresholding and classification of pixel's intensity). In general, however, segmentation is a challenging task and requires more significant human interactions (for semi-automatic segmentation effort) and sophisticated algorithms (for automatic segmentation), especially when involving analysis for region of interest (ROI) with asymmetrical structures. For example, the distribution of intensity values corresponding to one structure may overlap those of another structure, defeating intensity-based segmentation techniques, hence driving the need for pattern recognition or prior knowledge-driven techniques.

Segmentation approaches by boundary detection such as snakes and “EdgeFlow” [91, 92], are generally native algorithms that require the presence of some image features (such as an edge or intensity flow vector), along the boundary of the object of interest,

and gravitate toward that feature. These techniques are computationally fast however, naturally, these methods are sensitive to the initial curve position and may over- or under-segment through the boundary of the object of interest if the edge feature is not sufficiently salient in a certain region in the image.

Level-set based segmentation, introduced in [93, 94], involves deciphering the minimization of energy-based snake active contours problem by using geodesics or minimal distance curves [95, 96], where a curve is usually embedded as a zero level set at a higher dimensional surface. The entire surface is converged to metric minimization, outlined by the image gradient and curvature. Previous works based on level-set include extensions such as texture models [97], regularization of two shared-dimensions [98], pedal curve evolution [99] and global intensity statistics [100].

Technically, when segmenting anatomical structures of interest, establishing prior knowledge or information about the expected shape of that structure could substantially aid in the segmentation process. Both *Cootes et al.* [101] and *Wang and Staib* [102] examined a set of corresponding points through a set of training images and constructed a statistical analysis of shape variation that was later employed in the localization of the boundary. *Staib and Duncan* [103] integrate global shape information into the segmentation task by exploiting an elliptic Fourier decomposition of the boundary and insertion of a Gaussian prior on the Fourier coefficients. And finally, a shape prior technique for musculoskeletal segmentation by using shape matching and deformable models, was derived by *Gilles and Pai* [104], that was able to estimate elastic deformations of large muscle tissue.

For medical imaging, the aims of segmentation are: studies of anatomical structure; to locate and identify pathological structures or abnormalities; diagnosing (quantify, measure and compare the severity of the contracted disease); and treatment planning. In this chapter, we present the existing methodological and conceptual techniques that address the issues related to segmentation in medical imaging, including the *state-of-the-art* approaches designed to tackle muscle segmentation.

3.2 Segmentation Methods

In this section, the segmentation algorithms were categorized into three groups, each representing the difficulty level of methodological development, and often associated

with the derivation timeline of those algorithms.

3.2.1 First Generation

Algorithms within this group are considered as entry- or low-level methods. However, due to their computational simplicity, one of this group's major strengths is fast processing time.

3.2.1.1 Thresholding

Thresholding is a simple but efficient technique that distinguishes regions in an image (usually into a foreground and background), by creating binary partitioning based on its intensity. Its procedures include first determining a global intensity value, which divides the desired regions based on image histograms. Assuming a threshold value of T applied to an image $f(x, y)$, the resulting segmented image $g(x, y)$ is defined by:

$$g(x,y) = \begin{cases} 1, & \text{if } f(x,y) \geq T \\ 0, & \text{if } f(x,y) < T \end{cases} \quad (3.1)$$

Threshold value can be generated in many ways, i.e. manual interaction, adaptive method, multi-thresholding [105] or automatic thresholding (such as Otsu's [106]). Otsu's thresholding involves computing a set of histogram data, usually from a grayscale image. Then for each possible threshold value (the histogram bin number), the variance of all the bins before and the bins after that point is measured to evaluate the spread within each of the classes. As each potential threshold is evaluated, the threshold that produced the minimum intra-class variance (σ_W^2) is discovered. Statistically, this can be defined by:

$$\sigma_W^2 = W_1\sigma_1^2 + W_2\sigma_2^2, \quad (3.2)$$

where W_x is the weighting of the class given by:

$$W_x = \sum_a^b P(i), \quad (3.3)$$

where $P(i)$ is the class probability, i.e. the total number of pixels in the image divided by the number of pixels in the class. Practically, analyzing the intra-class variance σ_W^2 , is a

time consuming process. Alternatively, variance between classes σ_B^2 , can be calculated instead. The between class variance is the opposite of the intra-class variance in that the threshold that produces the maximum amount of variance is reserved, rather than the minimum. This can be measured using the following formula:

$$\sigma_B^2 = W_1W_2(\mu_1 - \mu_2)^2, \quad (3.4)$$

where μ_x is the class means level of the original image. All pixels with intensities greater than the normalized intensity value (global threshold) are then grouped into one class and all other pixels into another class, converting an intensity image to a segmented binary image.

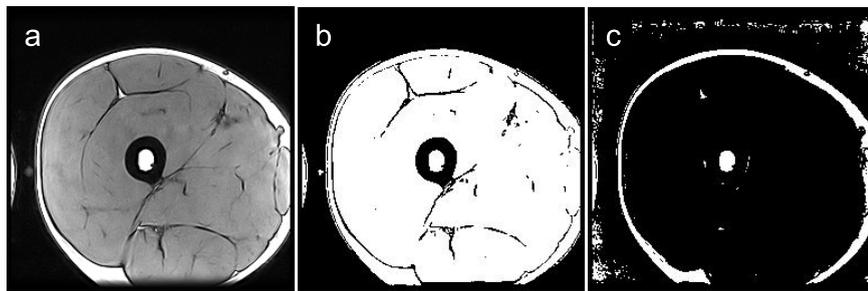


FIGURE 3.1: Results of different thresholding procedures on an MR image. (a) Original image. (b) Otsu's thresholding. (c) Adaptive thresholding.

3.2.1.2 Edge Detection

Edge detection, also known as edge tracing is a common technique in image processing for discovering the object boundaries within an image [107, 108]. It works by identifying intensity discontinuities or gaps. This method is widely applied in image segmentation and feature extraction in many areas related to medical image processing and computer vision, mainly for estimating structures and properties of objects in an image. Roberts, Sobel, Prewitt and Canny are some of the most common edge description operators used for edge detection.

Discontinuities in the image intensity can be either step discontinuities, or line discontinuities. However, step and line edges rarely emerge in ideal condition. Because of the smoothing introduced by most sensor components, sharp discontinuities rarely exist. Step edges become ramp edges and line edges become roof edges, where intensity

changes are not instantaneous but occur over a finite distance. Figure 3.2 illustrates the profile for each of these edges.

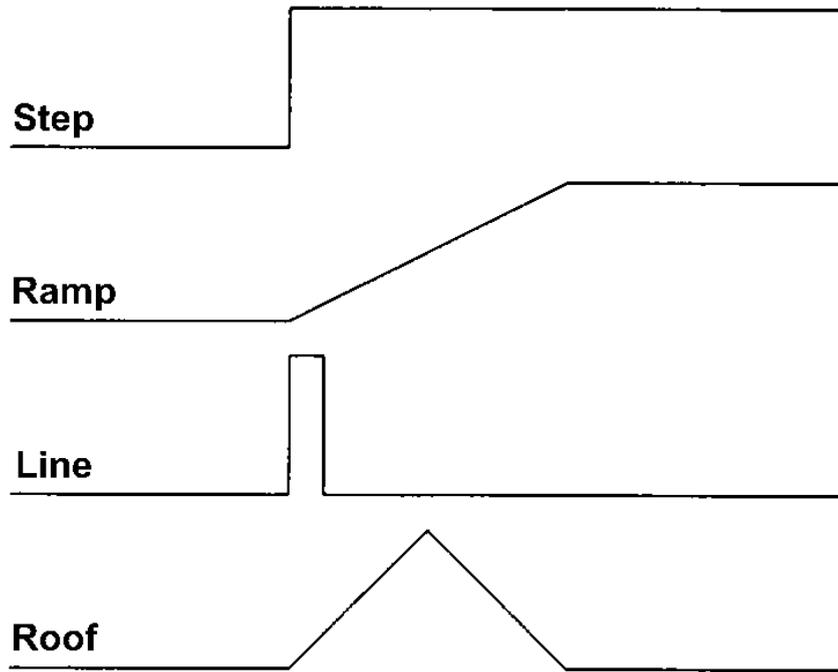


FIGURE 3.2: Edge profiles.

In one-dimensional data, a step edge is associated with a local peak in the first derivative. The gradient is a measure of change in a function, and an image can be considered to be an array of samples of some continuous function of image intensity. Theoretically, significant changes in the gray values in an image can be detected by using a discrete approximation to the gradient [109]. The gradient is the two-dimensional equivalent of the first derivative and is defined as the *vector*:

$$G[f(x, y)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\delta f}{\delta x} \\ \frac{\delta f}{\delta y} \end{bmatrix} \quad (3.5)$$

Two important properties associated with the gradient are the vector $G[f(x, y)]$ points in the direction of the maximum increase rate of the function $f(x, y)$ and the other one is the gradient magnitude, where:

$$G[f(x, y)] = \sqrt{G_x^2 + G_y^2}, \quad (3.6)$$

equals the maximum increase rate of $f(x, y)$ per unit distance in G direction. However, it is practically common to approximate the gradient magnitude by absolute values of:

$$G[f(x, y)] \approx |G_x| + |G_y|, \quad \text{or} \quad (3.7)$$

$$G[f(x, y)] \approx \max(|G_x|, |G_y|). \quad (3.8)$$

From vector analysis, the gradient direction is defined as:

$$\alpha(x, y) = \tan^{-1} \left(\frac{G_y}{G_x} \right), \quad (3.9)$$

where the angle α is measured with respect to the x -axis. Note that the gradient magnitude is independent of the edge direction and this condition is known as isotropic operators or isotropic edge detectors [110]. There are four steps for edge detection, which consist of:

1. *Smoothing.* Commonly used to improve the performance of an edge detector with respect to noise. However, there is a trade-off between edge strength and noise reduction. More filtering to reduce noise results in a loss of edge strength.
2. *Enhancement.* Emphasizes pixels where there is a significant change in local intensity values and is usually performed by computing the gradient magnitude.
3. *Detection.* Determine which edge pixels should be discarded as noise and which should be retained (usually, thresholding provides the criterion used for detection).
4. *Localization.* Edge location and orientation can be estimated with sub-pixel resolution if required for the application.

Next, some commonly used edge detection operators will be presented and note that the use of the computational approach (the inclusion of one or more of the above four steps) is differ from one operators to another.

1. First Derivative.

- (a) **Roberts Operator.** The Roberts cross operator [111] provides a simple approximation to the gradient magnitude, where:

$$G[f(i, j)] = |f[i, j] - f[i + 1, j + 1]| + |f[i + 1, j] - f[i, j + 1]|. \quad (3.10)$$

By introducing convolution masks, this becomes:

$$G[f[i, j]] = |G_x| + |G_y|, \quad (3.11)$$

where G_x and G_y are measured using the following masks:

$$G_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad G_y = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (3.12)$$

As with the previous 2×2 gradient operator, the differences are computed at the interpolated point $[i + \frac{1}{2}, j + \frac{1}{2}]$. The Roberts operator is an approximation to the continuous gradient at that point and not at the point $[i, j]$.

- (b) **Sobel Operator.** Another way to avoid having the gradient calculated around an interpolated point between pixels is by using a 3×3 neighbourhood for the gradient calculations. Consider the arrangement of pixels around the $[i, j]$ pixel as shown in Figure 3.3 below:

a_0	a_1	a_2
a_7	$[i, j]$	a_3
a_6	a_5	a_4

FIGURE 3.3: Neighbourhood pixels labelling for Sobel and Prewitt operators.

The Sobel operator [112] is the gradient magnitude computed by:

$$M = \sqrt{s_x^2 + s_y^2}, \quad (3.13)$$

where the partial derivatives are computed by:

$$s_x = (a_2 + ca_3 + a_4) - (a_0 + ca_7 + a_6), \quad (3.14)$$

$$s_y = (a_0 + ca_1 + a_2) - (a_6 + ca_5 + a_4), \quad (3.15)$$

by using a constant of $c = 2$. Like the other gradient operators, s_x and s_y can be implemented using convolution masks of:

$$s_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad s_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (3.16)$$

where emphasis is prioritize on pixels that are closer to the center of the mask.

(c) **Prewitt Operator.** The Prewitt operator [111] uses the same equations as the Sobel operator, except that the constant $c = 1$.

$$s_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad s_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad (3.17)$$

However, unlike Sobel operator, Prewitt operator does not place any emphasis on pixels that are closer to the center of the masks.

2. **Second Derivative.** The edge detectors discussed earlier computed the first derivative and, if it was above a threshold, the presence of an edge point was assumed. This results in detection of too many edge points. A better approach would be to find only the points that have local maxima in gradient values and consider them edge points. This means that at edge points, there will be a peak in the first derivative and, equivalently, there will be a zero crossing in the second derivative. Thus, edge points may be detected by finding the zero crossings of the second derivative of the image intensity [108].

(a) **Laplacian Operator.** This second derivative of a smoothed step edge is a function that crosses zero at the location of the edge and employed a two-dimensional equivalent of the second derivative [113]. The Laplacian operator function with respect to $f(x, y)$ is:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} . \quad (3.18)$$

The second derivatives along x and y directions are approximated using differential equations, and since the approximation is centred around pixel $[i, j + 1]$,

the final functions of Laplacian operator, with respect to x and y are:

$$\frac{\partial^2 f}{\partial x^2} = f[i, j + 1] - 2f[i, j] + f[i, j - 1], \quad (3.19)$$

$$\frac{\partial^2 f}{\partial y^2} = f[i + 1, j] - 2f[i, j] + f[i - 1, j]. \quad (3.20)$$

- (b) **Laplacian of Gaussian.** Therefore, it is desirable to filter out the noise before edge enhancement. Laplacian of Gaussian [108], with the combination approaches of Gaussian filtering and edge detection by Laplacian are employed, mainly for noise filtering before extracting edges. The output of this operator - $h(x, y)$, is computed by the convolution operation of:

$$h(x, y) = \nabla^2[g(x, y) * f(x, y)]. \quad (3.21)$$

By applying the derivative rule for convolution, also known as *Mexican hat* operator [114], the function simplified as:

$$h(x, y) = [\nabla^2 g(x, y)] * f(x, y), \quad (3.22)$$

where:

$$\nabla^2 g(x, y) = \left(\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} \right) e^{-\frac{(x^2 + y^2)}{2\sigma^2}}. \quad (3.23)$$

This results in a formula mathematically equivalent to the following two methods, which are: (i) Convolve the image with a Gaussian smoothing filter and compute the Laplacian of the result and; (ii) Convolve the image with the linear filter that is the Laplacian of the Gaussian filter.

3. Gaussian Edge Detection.

- (a) **Canny Edge Detector.** The Canny edge detector [109] is the first derivative of a Gaussian and closely approximates the operator that optimizes the product of signal-to-noise ratio and localization. Its algorithm is summarized by the following notation. Let $I[i, j]$ denote the image. The result from convolving the image with a Gaussian smoothing filter using separable filtering is an array of smoothed data:

$$S[i, j] = G[i, j; \sigma] * I[i, j], \quad (3.24)$$

where σ is the spread of the Gaussian and controls the degree of smoothing. The gradient of the smoothed array $S[i, j]$ can be computed using the 2×2 first derivative approximations to produce two arrays $P[i, j]$ and $Q[i, j]$ for the x and y partial derivatives:

$$P[i, j] \approx (S[i, j + 1] - S[i, j] + S[i + 1, j + 1] - S[i + 1, j])/2, \quad (3.25)$$

$$Q[i, j] \approx (S[i, j] - S[i + 1, j] + S[i, j + 1] - S[i + 1, j + 1])/2. \quad (3.26)$$

The finite differences are averaged over the 2×2 square so that the x and y partial derivatives are computed at the same point in the image [109]. The magnitude and orientation of the gradient can be computed from the standard formulas for rectangular-to-polar conversion:

$$M[i, j] = \sqrt{P[i, j]^2 + Q[i, j]^2}, \quad (3.27)$$

$$\theta[i, j] = \arctan(Q[i, j], P[i, j]), \quad (3.28)$$

where the *arctan* function takes two arguments and generates an angle over the entire circle of possible directions. These functions must be computed efficiently, preferably without using floating-point arithmetic. Next, to identify edges, the broad ridges in the magnitude array must be thinned so that only the magnitudes at the points of greatest local change remain. This process is called nonmaxima suppression (nms) [114].

$$\begin{aligned} N[i, j] &= \text{nms}(M[i, j], \text{Sector}(\theta[i, j])) \\ &= \text{nms}(M[i, j], \zeta[i, j]). \end{aligned} \quad (3.29)$$

And finally, the application of double thresholding [115] (also known as hysteresis thresholding) τ_1 and τ_2 to the nonmaxima-suppressed image $N[i, j]$, with $\tau_2 \approx 2\tau_1$, to produce two thresholded edge images $\tau_1[i, j]$ and $\tau_2[i, j]$, for finer edges extraction.

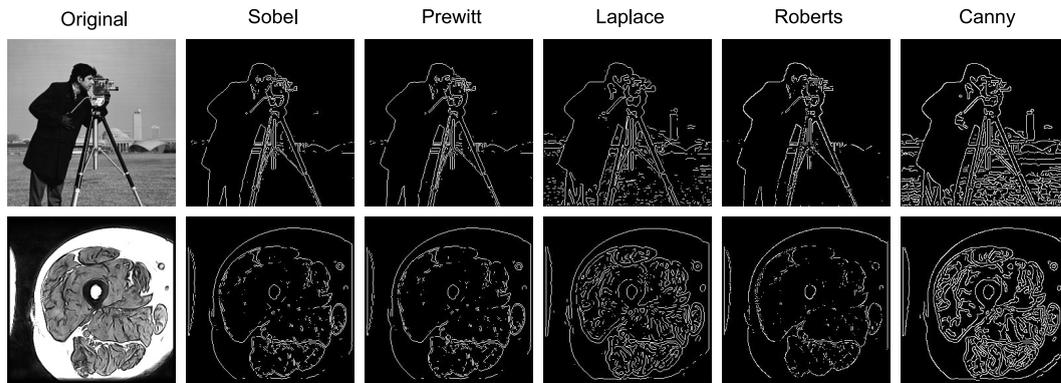


FIGURE 3.4: Results of different edge detection operators. *Cameraman image courtesy of Mathworks.*

3.2.1.3 Region Growing

Region growing [116] is another algorithm in image processing and computer vision for object extraction that is connected based on some predefined criteria or sometimes referred to as similarity constraints, such as intensity information, colour or edges in the image. This method requires a seed point that can be manually selected by an operator or adaptively placed by using prior information, and extracts all neighbouring pixels that satisfy one or more similarity constraints as above.

Region growing is useful to isolate structures in an image, especially when the regions of interest and background have defined edges, overlapping pixel intensities and separated by some partition. Like most methods, region growing also employed with a set of other algorithms, particularly to increase the efficacy of small delineation processes of simple structures, such as tumours and lesions [117].

The major drawback of this algorithm is its higher computational time. As it is considered a data parallel method, the numbers of threads change as the border expands. This usually involves evaluation of all the values over the global memory, which is time consuming and requires additional works and initiates branch divergence, restricting the potential speedup over an enhanced sequential operation. Issues related to noise however, can be dominated by using additional masks for hole and outliers filtration. Figure 3.5 demonstrates region growing application with an addition of thresholding to separate the prominent muscle mass area without splitting the image apart.

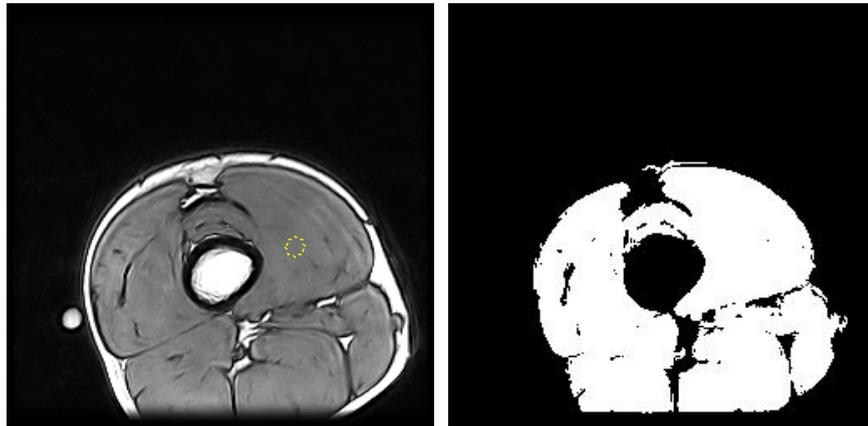


FIGURE 3.5: Seed-based region growing segmentation. Left - Original thigh MR image. Right - region growing segmentation result after seeding point placement (yellow marker) on the muscle region.

3.2.2 Second Generation

Considered as mid-level approaches and often used in conjunction with methods in the first group.

3.2.2.1 Clustering

Clustering, in general, is a process of grouping objects in an image based on information found in the data that describes the objects and their connections. The goal is that the objects within a cluster (or group) be related to one another and dissimilar from objects in other clusters. In image processing, it is considered as the most important unsupervised learning technique and also one of the most used algorithms in image segmentation [118], due to its intuitive and easy nature. Four commonly used clustering algorithms are the K-means [119], the fuzzy c-means [120], Hierarchical clustering [121] and mixture of Gaussians [122].

K-means. This clustering technique classifies data or objects by iteratively computing mean intensity for a certain number of clusters (named k clusters) and segmenting the image by associating each pixel to the appropriate cluster with the closest mean. When the assigning process is complete, the k new centroids require a re-calculation for a new binding of the same data set to the k new centroids. This looping process generates changes to the k centroid location, and iteratively occurs until no further changes

are possible [119]. This produces a separation of the objects into clusters from which the metric to be minimized can be measured by an objective function of:

$$J = \sum_{j=1}^k \sum_{i=1}^n \left\| x_i^{(j)} - c_j \right\|^2, \quad (3.30)$$

where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and cluster centre c_j , with a distance indicator of n data points from the respective cluster centres. Figure 3.6 shows 2 examples of K -means segmentation of data with 2 and 4 clusters, respectively.

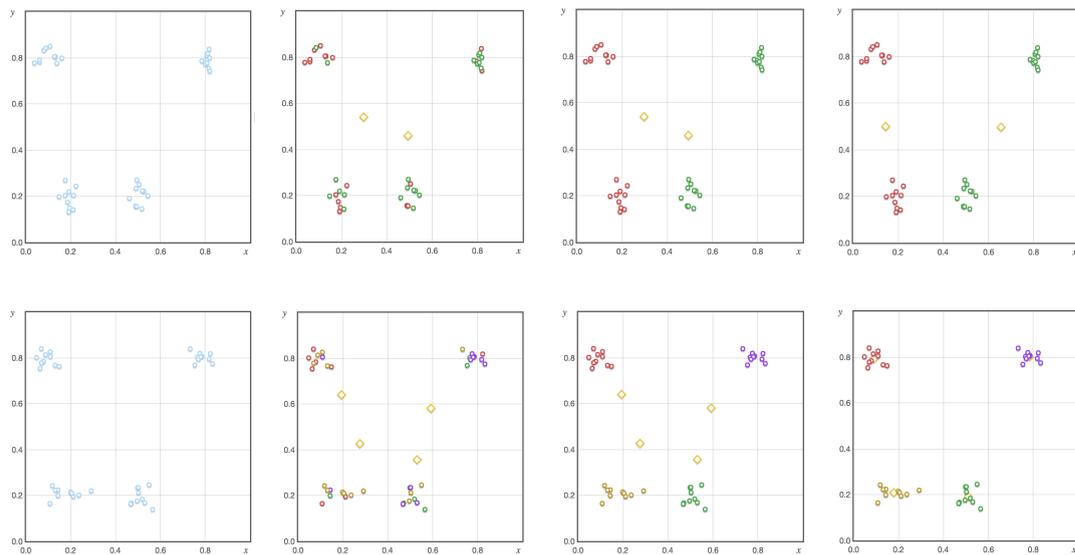


FIGURE 3.6: K -means applications on 2 (top row) and 4 (bottom row) clusters of random data. First columns represent random data generations. Second columns: clusters initialization. Third columns: assignment of data points to closer cluster. Final columns: Centre (final centroid location) of each cluster.

Fuzzy C-mean. This technique is considered as an extension of K -means clustering and lets data points to be assigned to more than one clusters [120]. The algorithms is based on minimization process of the following objective function:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, \quad (3.31)$$

where $1 \leq m < \infty$; u_{ij} is the membership degree of x_i in the cluster j ; x_i is the i th of d -dimensional measured data; c_j is the d -dimensional centre of the cluster; and $\| * \|$ is any norm expressing the similarity between any measured data and the centre. Fuzzy

clustering is processed through an optimization iteration of the objective function shown above, with the update of membership u_{ij} and the cluster centres c_j by:

$$u_{ij} = \frac{1}{\sum_{k=1}^C \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}, \quad (3.32)$$

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m \cdot x_i}{\sum_{i=1}^N u_{ij}^m}. \quad (3.33)$$

Iteration stops when $\max_{ij} \left\{ |u_{ij}^{(k+1)} - u_{ij}^{(k)}| \right\} < \varepsilon$, where ε is termination criterion between 0 and 1; and k is the iteration steps. This process converges to a local-minima or a saddle point of J_m . Figure 3.7 illustrates a simple 1-dimension data, distributed along the x -axis, with corresponding post-processing output by K -means and fuzzy c-means (FCM), respectively.

Unlike K -means, FCM offers a better solution for intersected data due to data point membership is assigned to each cluster centre and may belong to more than one cluster centre. However, unequal weight underlying factors by Euclidean distance measurements may results in an expensive number of iteration, hence a further processing time compared to K -means [120].

3.2.2.2 Watershed

Watershed is a classic region-based segmentation approach proposed by Digabel and Lantuejoul [123, 124] and later improved by Beucher and Lantuejoul [125]. The concept of this method is based on topography, where an image is viewed as a three-dimensional object. The third dimension is the height of each pixel based on its intensity value, or in this regards also known as local topology or elevation. The algorithm floods catchment basins from local minima (or sometimes referred to as hole) and build watershed lines (or divide lines) when different water sources from other basins are about to merge. The height of watershed lines are increased at the same rate as the water level increases and this process stops after the water reaches the highest point in the topology, corresponding to the pixel with maximum intensity.

Watershed is a popular method in image processing, including medical image segmentation due to its simplicity, intuitive, fast and can be paralleled [126] to further speedup

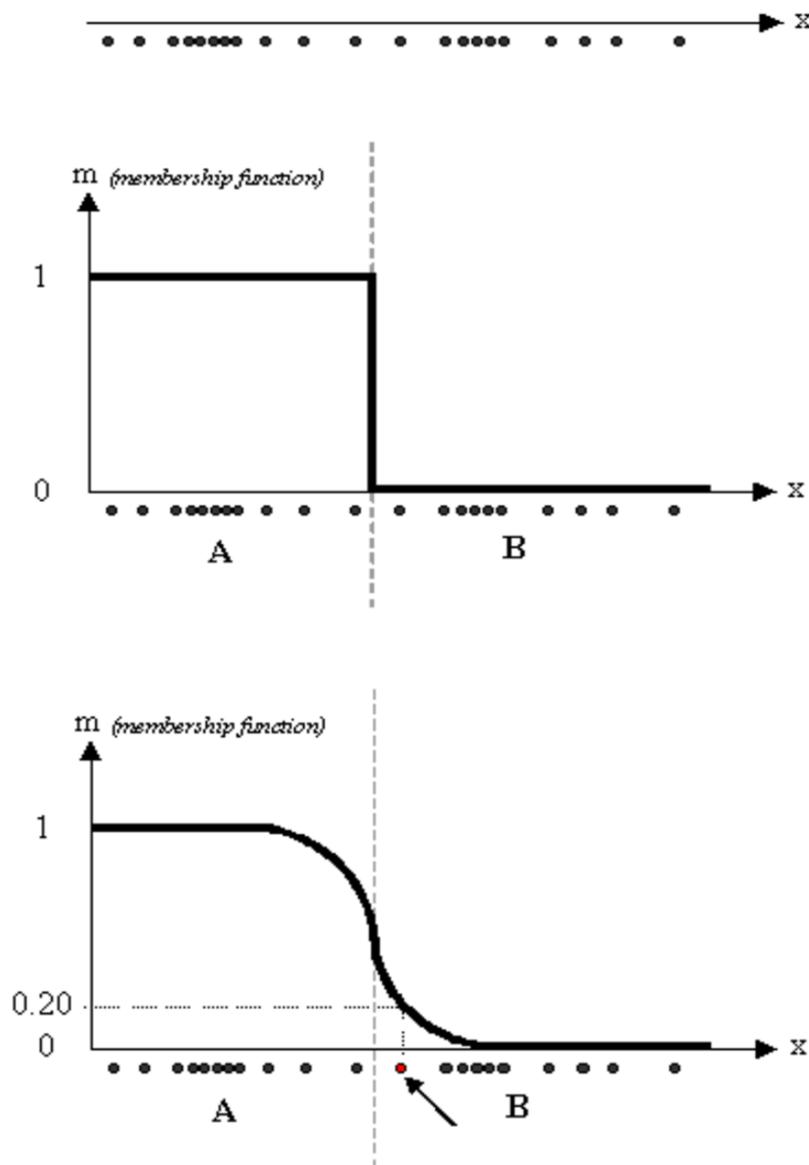


FIGURE 3.7: Top 1-D graph - Data distribution along x -axis. Mid graph - Membership function of two clusters by K -means. Bottom graph - Membership function of two clusters by FCM. The point marked as a red dot, although belongs to B cluster, also belongs to A cluster, with membership degree valued at 0.2. *Courtesy: <http://home.deib.polimi.it>*

the processing time. However, there are also certain important issues related to this method such as over-segmentation; sensitivity to noise; poor detection of areas with prominent low contrast boundaries; and poor detection of thin areas. Some adaptation, especially on medical image segmentation, have been completed to improve such issues. [127, 128] recording good segmentation performance and reducing over-segmentation errors by controlling marker (local minima) location with mathematical morphology.

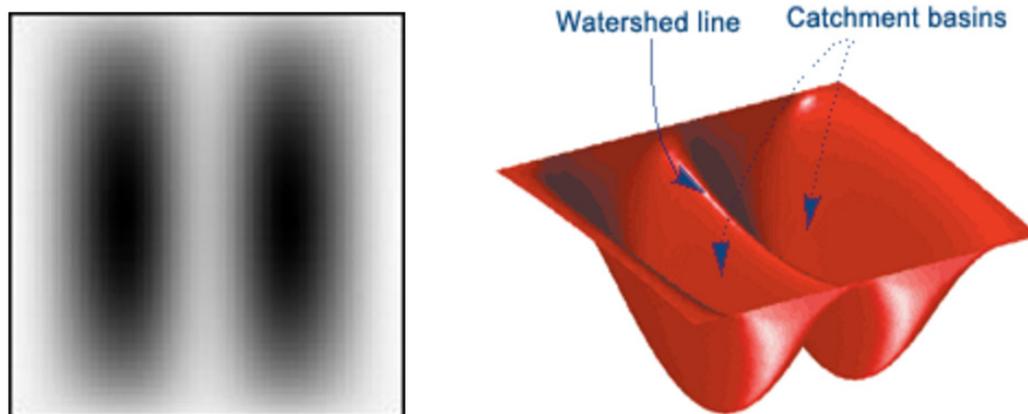


FIGURE 3.8: Left - Simulated image with two dark blobs. Right - How watershed algorithm treats image on the left, such that the dark regions are “low” and bright regions are “high”, resulting watershed line and catchment basins. *Courtesy: <http://uk.mathworks.com>*

Over-segmentation can also be minimized by adjusting the opening-closing (or erosion-dilation) of morphological gradient [129]. A new and improved segmentation algorithm of watershed transformations by prior knowledge [130] that employs different functions based on different objects used (instead of the usual application of intensity gradient) demonstrated substantial segmentation accuracy on MRI and CT images.

3.2.2.3 Active Contour

Part of a parametric deformable model family and considered to be one of the most applied algorithms in image processing [131]. This simulated closed parametric curves capable of contraction or extension, with a number of iterations over time, and adapts to certain features for a given image, by the means of energy minimizing function.

Edge-based. Introduced by *Kass et al.* [91] and also known as “snake” algorithm. This edge-based snake model represents a controlled continuity spline that can deform to match any shape, under the influence of image forces and external constraint forces. The internal spline forces add a piecewise smoothness constraint. The image features attract the snake to the salient image features, such as lines, edges and terminations.

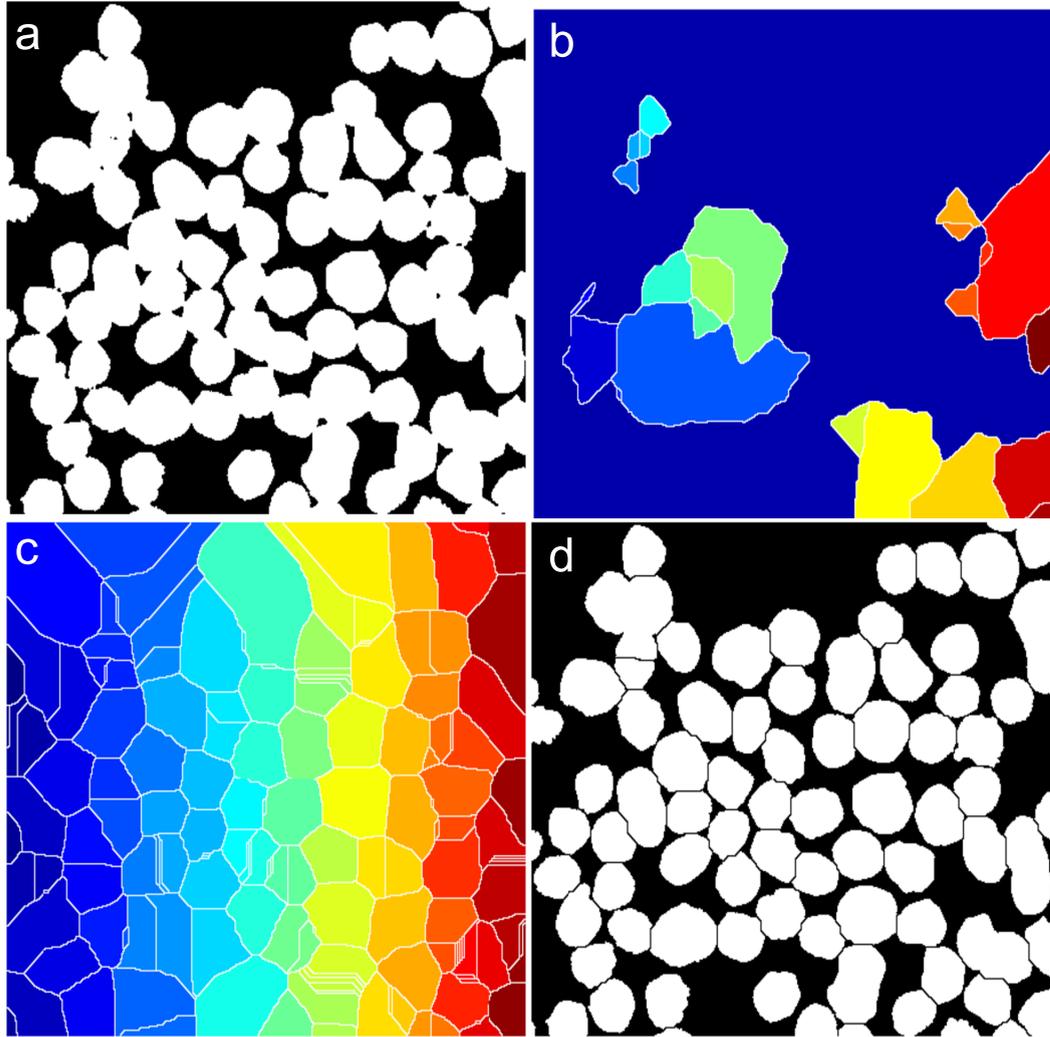


FIGURE 3.9: Image segmentation by watershed. (a) Original image. (b) Watershed transformation of (a). (c) Watershed transformation with Euclidean distance transform of the binary image (a). (d) Segmentation result with desired location minima. *Courtesy: <http://uk.mathworks.com>*

The total energy of the snake can be written as:

$$\begin{aligned}
 E_{snake}^* &= \int_0^1 E_{snake}(v(s)) \\
 &= \int_0^1 [E_{int}(v(s)) + E_{image}(v(s)) + E_{con}(v(s))] ds,
 \end{aligned} \tag{3.34}$$

where E_{int} represents the internal energy of the spline and computed from within the curve for smooth deformation, E_{image} represents the image forces that drive the curve toward the desired feature of interest, and E_{con} represents the external constraint forces. v denotes a parametric curve as:

$$v = v(s) = (x(s), y(s)), \tag{3.35}$$

where $s \in [0, 1]$ and s is subset material of an image. The internal energy can be written as:

$$E_{\text{int}} = (\alpha(s)|v_s(s)|^2 + \beta(s)|v_{ss}(s)|^2)/2 . \quad (3.36)$$

The first order term represents the stretching force (elasticity) of the snake while the second order term represents its bending force (rigidity). The values of α and β act as balancing parameters.

The image energy can later be inscribed as:

$$E_{\text{image}} = \omega_{\text{line}}E_{\text{line}} + \omega_{\text{edge}}E_{\text{edge}} + \omega_{\text{term}}E_{\text{term}} . \quad (3.37)$$

The external constraint energy E_{con} has 2 major forms; spring energy and repulsion force. The snake algorithm will iteratively deform the model and try to find the configuration with minimum total energy E_{snake}^* , which corresponds to the best fit as to the preliminary shape of the template image as a structure regulator [132].

Region-based. Although effective in generating closed parametric curves of regions of interest to an image, dependency to proximity of edges to the image energy, E_{image} of edge-based active contour yields higher sensitivity of this model to noise [131]. This may lead to inaccurate object segmentation, especially due to poor edges detection. An upgraded version of deformable active contour approach was proposed by *Blake et al.* [133] and *Mumford and Shah* [134], termed region-based active contour, with energy minimization function of:

$$E(u, C) = \frac{1}{2} \int_{\Omega} (I - u)^2 dx + \lambda^2 \frac{1}{2} \int_{\Omega/C} |\nabla u|^2 dx + \nu |C|, \quad (3.38)$$

where C is the curve, I is the image, u is the piecewise smooth function. The first term maps the similarity of u with I , the second term penalizes edges gaps of u excluding on the curve boundary C and the final term confines the C length to segment only important objects. Equation 3.38 is further improved with a simpler function by setting the limit of $\lambda \rightarrow \infty$ [134] and giving a new function of:

$$E(\{u_i\}, C) = \frac{1}{2} \sum_i \int_{R_i} (I - u_i)^2 dx + \nu |C|, \quad (3.39)$$

where u_i is a partition of image I into constant regions R_i determined by curve C . An introduction of Gaussian probabilistic model for each region with constant grayscale value by *Zhu et al.* [135] resulted in a new function of:

$$E(\{\alpha_i\}, C) = \frac{1}{2} \sum_i \int_{R_i} -\log P(I|\alpha_i) dx + \nu |C|, \quad (3.40)$$

where $P(I|\alpha_i)$ is the posterior probability by computing image I intensity values, inside region R_i given distribution parameter of α_i . A hybrid extension of edge- and region-based function for energy minimization by standard gradient descent method, termed *diffusion snake* has been introduced by *Cremers et al.* [136]:

$$E(u, C) = \frac{1}{2} \int_{\Omega} (I - u)^2 dx + \lambda^2 \frac{1}{2} \int_{\Omega/C} |\nabla u|^2 dx + \nu \int |C_s|^2, \quad (3.41)$$

by adapting equation terms on 3.38 external energy and equation 3.34 internal elastic constraint, where $\nu \int |C_s|^2 = E_{\text{int}}(v(s))$. Another region-based energy model by *Chan and Vese* [137], which is the representation of an energy minimization-based segmentation as in original fitting term of:

$$\begin{aligned} F_a(C) + F_b(C) = & \int_{\text{inside}(C)} |I(x, y) - k_a|^2 dx dy \\ & + \int_{\text{outside}(C)} |I(x, y) - k_b|^2 dx dy, \end{aligned} \quad (3.42)$$

where both variable curves (*inside/outside*) C and their dependant constants k_a, k_b are the average derivations of inside and outside image I , respectively, providing an obvious fitting term minimizer to the object's boundary C_0 , where:

$$\inf_C \{F_a(C) + F_b(C)\} \approx 0 \approx F_a(C_0) + F_b(C_0). \quad (3.43)$$

Energy curve fitting and regularization terms above were further minimized by additional Euler-Lagrange-based H_ε and δ_ε regularizing functions as $\varepsilon \rightarrow 0$, and defined as:

$$\begin{aligned}
 F_\varepsilon(k_a, k_b, \phi) = & \mu \int_{\Omega} \delta_\varepsilon(\phi(x, y)) |\nabla \phi(x, y)| dx dy \\
 & + v \int_{\Omega} H_\varepsilon(\phi(x, y)) dx dy \\
 & + \lambda_a \int_{\Omega} |u_0(x, y) - k_a|^2 H_\varepsilon(\phi(x, y)) dx dy \\
 & + \lambda_b \int_{\Omega} |u_0(x, y) - k_b|^2 (1 - H_\varepsilon(\phi(x, y))) dx dy,
 \end{aligned} \tag{3.44}$$

where ϕ is level set function represents the curve, C and $\int_{\Omega} |\nabla \chi_\omega| dx dy$ is the lower semi-continuity of the total variation. Region-based approaches possess substantial advantages over edge-based methods, especially robustness against initial curve placement, flexibility of dealing with noise in image and customizable energy minimization model [132, 138].

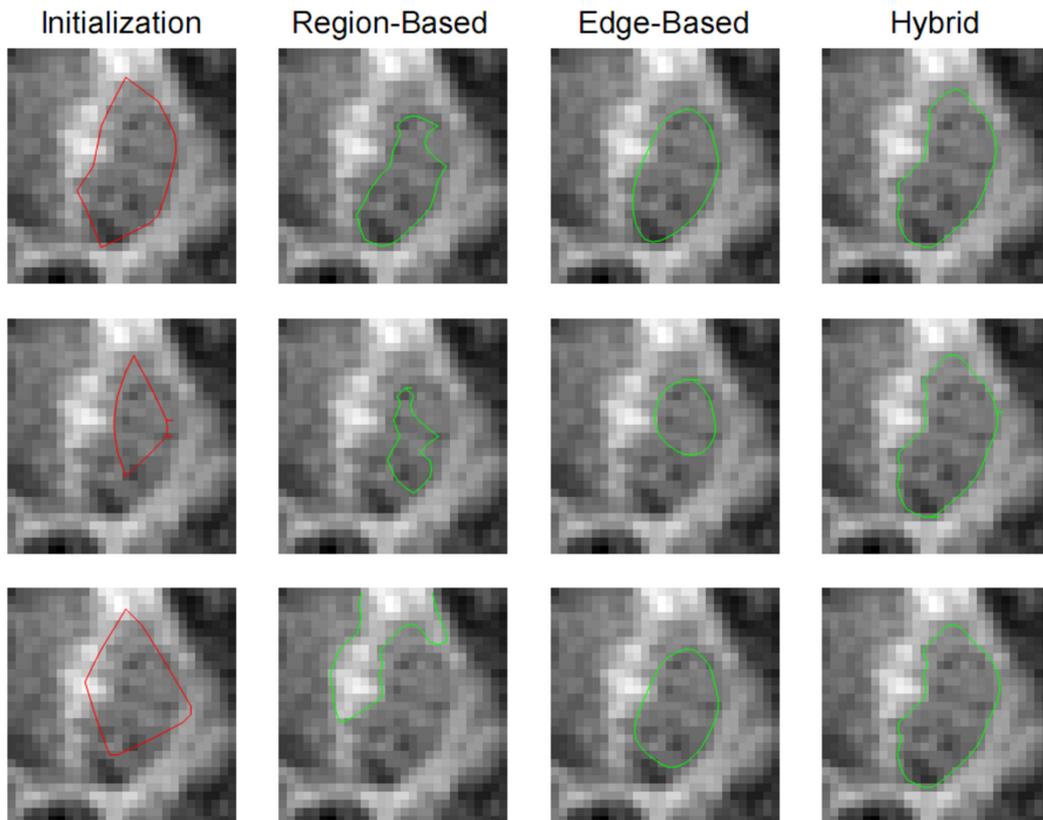


FIGURE 3.10: Image segmentation by different mode of active contours on putamen brain's sub-cortex MR image. First column represents different curve initializations on the image. Second column - Curve convolutions by region-based algorithm. Third column - Curve transformations by edge-based. Last column - Segmentation results by hybrid active contour. *Courtesy: <http://www.shawnlankton.com/2007/02/cool-hybrid-segmentation/>*

3.2.2.4 Level Set

Level set was introduced by *Dervieux and Thomasset* [139] and classified as geometric deformable models [140]. Its application is mainly to solve certain issues that active contour is unable to address and uses evolving curves C implicitly as the zero level line location of a function:

$$C = \{x \in \Omega | \phi(x) = 0\} , \quad (3.45)$$

where $\phi : \Omega \rightarrow R$, and integrated with snake's energy minimization function, with $C : [0, 1] \rightarrow \Omega$ [141, 142] to produce:

$$E(C) = \int_0^1 \alpha E_{int}(C(p)) + \beta E_{img}(I(C(p))) + \gamma E_{ext}(C(p)) dp , \quad (3.46)$$

where E_{int} controls curve smoothness, E_{img} is data term that drives the curve towards boundaries, I is the image, E_{ext} is prior term set by the user, and parametric constants of α, β, γ to harmonize the equation. This function is minimized by variational form approach of:

$$\frac{\partial C}{\partial t} = -\frac{\partial E(C)}{\partial C} = F.n , \quad (3.47)$$

where F represents speed function and n is normal vector of the curve. By joining the level set formulation to the previous equation produces:

$$\frac{\partial}{\partial t} \phi = -|\nabla \phi| F . \quad (3.48)$$

This indicates the implicit function (ϕ) progression at boundaries location. Some of the earliest applications in level set form incorporating *Geodesic Active Contours* model [93, 94] with evolution function is given by:

$$\frac{\partial}{\partial t} \phi = |\nabla \phi| \operatorname{div} \left(g(I) \frac{\nabla \phi}{|\nabla \phi|} \right) , \quad (3.49)$$

where g represents edge function that expands proportionally with the edges strength in image I . This equation represents an internal flow, which contracts the initial curve towards the discovered edges. Generally, its application is considered as a signed distance

function to the curve transformation:

$$\phi(x) = \begin{cases} 0 & , x \in C , \\ D(x, C) & , x \in C_{in} , \\ -D(x, C) & , x \in C_{out} , \end{cases} \quad (3.50)$$

where $D(x, C)$ is Euclidean distance from x to C , $C_{in} \subset \Omega$ and $C_{out} \subset \Omega$ denote the region inside and outside of C , respectively. As the function is dependable to initialization of curve C and fragile to local extrema variations, a global regional term, incorporated with energy model of [137, 142] was proposed:

$$E(\phi, \theta_1, \theta_2) = \int_{\Omega} -H(\phi) \log p(I|\theta_1) - (1 - H(\phi)) \log p(I|\theta_2) + w|\nabla H\phi| dx , \quad (3.51)$$

where $p(I|\theta_i)$ is the posterior probability to detect I by specified region model parameter θ_i . H represents Heaviside step function, where $H(x) = 1$, $x \in C_{in}$, and $H(x) = 0$. Practically, H is considered as a smooth differentiable approximation of the Heaviside distribution. θ_i denotes region parameters and can either be customized as a statistical prior knowledge properties of the objects region [143], or projected in a different optimization method [137]. This formulation was later restructured to various deviations and applications, such as multi-phase segmentation [144] and non-parametric approximation of density functions [145].

3.2.2.5 Statistical Analysis for Shape Representation

Internal organs and their inner structures for all individuals, in general, share similar shapes and profile characteristics. Statistical shape analysis is an important tool for modelling these anatomical structures, especially from medical images [146, 147]. Given a set of pre-segmented or training images from different individuals, the goal is to model the geometric variability of the anatomical structures within a class of images, in which efficient parameterization of the geometric variability of human anatomy can be offered. These models may be used as seeding points for shape deformations during image segmentation [148], valuable in understanding human development through monitoring growth and disease [149] or even for gait recognition studies [150].

Some of the most common statistical analysis approaches capable of representing or modelling shapes are Statistical Shape Models (SSM) [151], Active Shape Models (ASM)

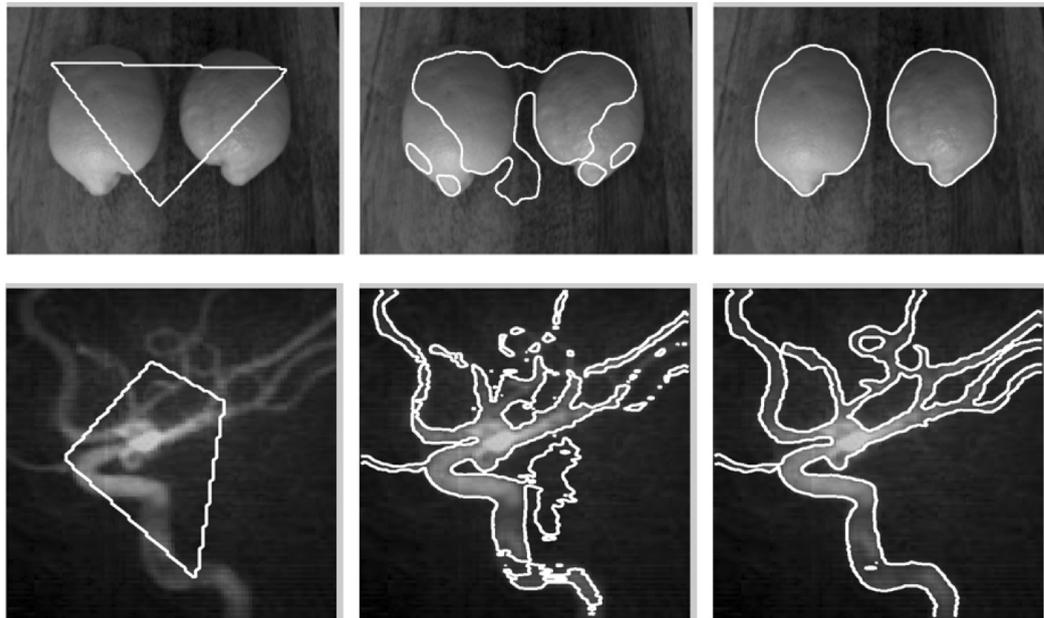


FIGURE 3.11: Lemons and CT vessel (top and bottom row, respectively) image segmentation by level set. Left column represents initial contours. Mid column - intermediate contours after certain number of iterations. Right column - final contours. *Courtesy: Li et al. - A Level Set Method for Image Segmentation in the Presence of Intensity Inhomogeneities with Application to MRI.*

[148], Active Appearance Models (AAM) [152] and Principle Component Analysis (PCA) [153]. However, in this section only PCA will be presented as other techniques employed slightly different parametric and geometric modelling approaches, suitable for other imaging modalities.

Principle Component Analysis. PCA is a very powerful and fast method for pattern recognition and dimensionality reduction (especially in computation that involves abundant of data with equal dimensions) due to one of its major characteristics of linear combination of data re-expression [154, 155]. $PX = Y$ represents a basic form of PCA as an orthogonal regression method, with linear transformation of P , and X and Y are equal to $m \times n$ matrices [156]. Another form to express data means following the above equation is a common measure of ratio of variances σ^2 or *signal-to-noise ratio (SNR)*, $SNR = \frac{\sigma_{data}^2}{\sigma_{noise}^2}$, with $SNR \ll 1$ and $SNR \gg 1$ indicates low and high precision data, respectively [156]. Covariance generalization of the resultant variance of dataset K and L summarized as $\sigma_{KL}^2 = \langle k_i l_i \rangle_i$, where i represents a single grouped observation of the k and l variables, which after the conversion of row vectors of dataset K and L becomes

[157]:

$$\sigma^2_{kl} \equiv \frac{1}{n-1} kl^T. \quad (3.52)$$

At this stage, each row and column of X is equal to x_i and \vec{X} (entire vector of X), respectively, resulting in a foundation covariance matrix of:

$$S_X \equiv \frac{1}{n-1} XX^T, \quad (3.53)$$

where S_X is a symmetric dimension matrix $m \times m$ and XX^T is ij^{th} value corresponds to x_i and x_j as per Equation 3.52. S_X is also refers as scatter matrix of:

$$S = \sum_{k=1}^n (x_k - \mu)(x_k - \mu)^T, \quad (3.54)$$

where μ is the mean vector, described as:

$$\mu = \frac{1}{n} \sum_{k=1}^n x_k. \quad (3.55)$$

Figure 3.12 demonstrates spatial minimizations by PCA to seismic volume or strata-grid attributes by using different weights and different combinations of the variables in the data sets.

3.2.2.6 Registration

Image registration is important in many aspects of functional image analysis, not only to basic concept of image stitching and alignment, satellite imaging, geospatial tagging and mapping, audio and video editing but also to medical imaging. The results of these image transformations can be used to model different registration elements of the same subject, image alignment and also as an effective tool for seeding points placement for optimum curve evolution process, especially if any statistical model analysis is involved [104, 158, 159].

Feature-based. This geometric feature-based method corresponds to physical points (point landmarks) in an image or between two images, where transformation process is measured physically according to these points.

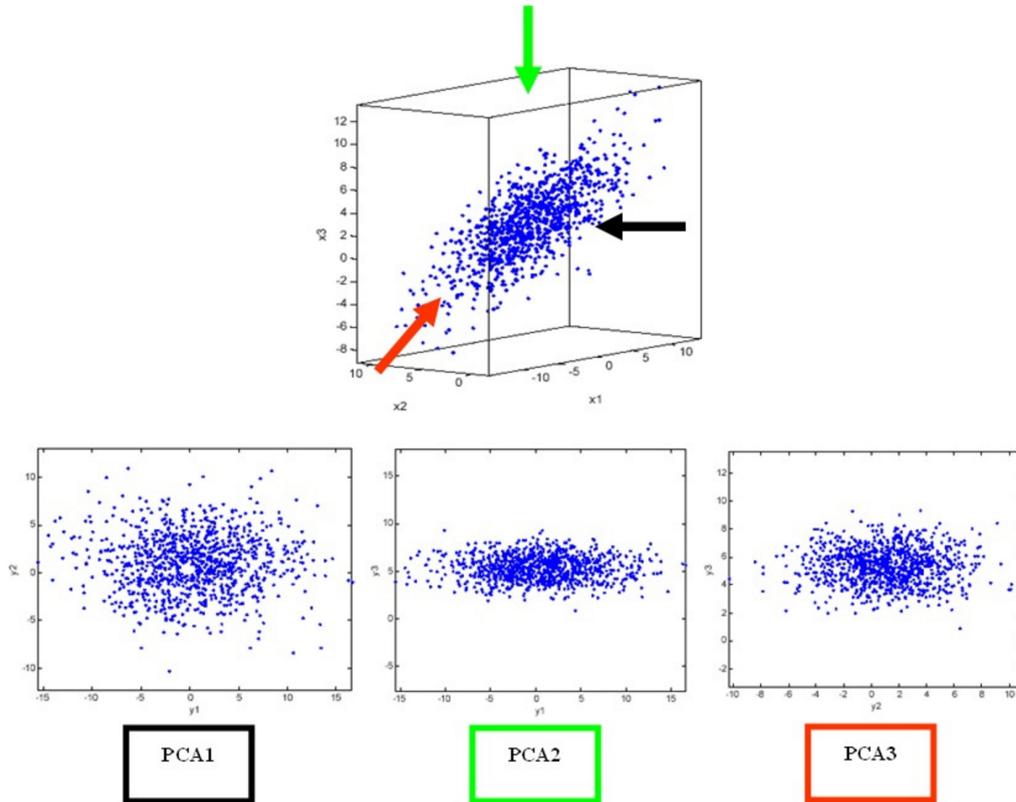


FIGURE 3.12: Projection of 3-D dataset into 2-D pattern expressions using PCA. The data compresses with minimal information degradation. *Courtesy: <http://www.geomodeling.ca/principal-component-analysis/>*

Rotation, Translation and Scaling

For each point (x_a, x_b, x_c) in an image or MRI scan, an affine mapping can be defined into the co-ordinates of another space (y_a, y_b, y_c) . This is expressed as:

$$\begin{aligned}
 y_a &= m_d x_a + m_e x_b + m_f x_c + m_x \\
 y_b &= m_g x_a + m_h x_b + m_i x_c + m_y \quad , \\
 y_c &= m_j x_a + m_k x_b + m_l x_c + m_z
 \end{aligned}
 \tag{3.56}$$

which is often represented by a simple matrix multiplication [160] of $y = mx$:

$$\begin{bmatrix} y_a \\ y_b \\ y_c \\ 1 \end{bmatrix} = \begin{bmatrix} m_d & m_e & m_f & m_x \\ m_g & m_h & m_i & m_y \\ m_j & m_k & m_l & m_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_a \\ x_b \\ x_c \\ 1 \end{bmatrix} ,
 \tag{3.57}$$

where (m_x, m_y, m_z) are arbitrary translation vectors, and $(m_d - m_l)$ is 3×3 rotation matrix [160] defined by:

$$m_{d-l} = m_a m_b m_c, \quad m_a = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha_a & -\sin \alpha_a \\ 0 & \sin \alpha_a & \cos \alpha_a \end{pmatrix}, \quad (3.58)$$

$$m_b = \begin{pmatrix} \cos \alpha_b & 0 & -\sin \alpha_b \\ 0 & 1 & 0 \\ \sin \alpha_b & 0 & \cos \alpha_b \end{pmatrix}, \quad m_c = \begin{pmatrix} \cos \alpha_c & -\sin \alpha_c & 0 \\ \sin \alpha_c & \cos \alpha_c & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where for example, m_a , rotates the image around axis a by an angle α_a . The elegance of formulating these transformations in terms of matrices is that several of them can be combined, simply by multiplying the matrices together to form a single matrix. This means that repeated re-sampling of data can be avoided when re-orienting an image. Inverse affine transformations are obtained by inverting the transformation matrix.

If a point x is to be translated by k units, then the transformation is simply: $y = x + k$ and can be inscribed in matrix as [161]:

$$\begin{bmatrix} y_a \\ y_b \\ y_c \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & k_a \\ 0 & 1 & 0 & k_b \\ 0 & 0 & 1 & k_c \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_a \\ x_b \\ x_c \\ 1 \end{bmatrix}. \quad (3.59)$$

The affine transformations described so far will generate purely rigid curvature mappings. Scaling are needed to change the size of the shape model and represent zooming in or out along the orthogonal axes, and can be represented as:

$$\begin{bmatrix} y_a \\ y_b \\ y_c \\ 1 \end{bmatrix} = \begin{bmatrix} k_a & 0 & 0 & 0 \\ 0 & k_b & 0 & 0 \\ 0 & 0 & k_c & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_a \\ x_b \\ x_c \\ 1 \end{bmatrix}. \quad (3.60)$$

A single zoom by a factor of -1 will flip an image. Two flips in different directions will merely rotate it by π radians (a rigid image transformation). In fact, any affine

transformation with a negative determinant will render the image flipped.

Let's assume two 2-D images, affine object (I') and original object I are related by registration transformation such that: $I'(x) = I(Ax + T)$, where A is the linear component matrix affine mode consists of scaling and rotation, and T is the 2-D translation vector. If both I' and I are related, then each point $I'(x_{I'}, y_{I'})$ in affine object corresponds to point $I(x_I, y_I)$ according to the matrix equation of:

$$\begin{bmatrix} x_{I'} \\ y_{I'} \\ 1 \end{bmatrix} = \begin{bmatrix} s \cos \phi & -s \sin \phi & \Delta x \\ s \sin \phi & s \cos \phi & \Delta y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_I \\ y_I \\ 1 \end{bmatrix}, \quad (3.61)$$

where s is the scale factor, ϕ is the angle of rotation and Δ is the translation in (x, y) unit. Homogeneously, for any pixel of $I'(x, y)$, it is proven that:

$$I(x, y) = I'(\Delta x + s * (x \cos \phi - y \sin \phi), \Delta y + s * (x \sin \phi + y \cos \phi)). \quad (3.62)$$

For finite disconnection of I' and I , both are assumed to be square in shape, with any given pixel area. Note that practically, in real medical images, the transformation of I' by modes of image registration introduces differences due to occlusions as some data moves into or out of the image frame. Figure 3.13 demonstrates the transformation process of a hexagonal shape, I' according to points in original hexagonal shape I .

Intensity-based. This approach match image's intensity patterns by statistical measurements, where intensity similarity between the source and the target images are computed and the transformation is adjusted until optimum similarity is registered. Measures of similarity includes squared differences in intensities, correlation coefficient, measures based on optical flow, and information-theoretic measures such as mutual information.

The sum of squared differences [162] is considered as the simplest method for intensity-based registration, which assumes that the images are similar at registration except for (Gaussian) noise. The correlation coefficient [158] assumes that matching intensities in the images have a linear correlation. These two methods are suitable for mono-modal registration where the intensity characteristics are very identical in the images.

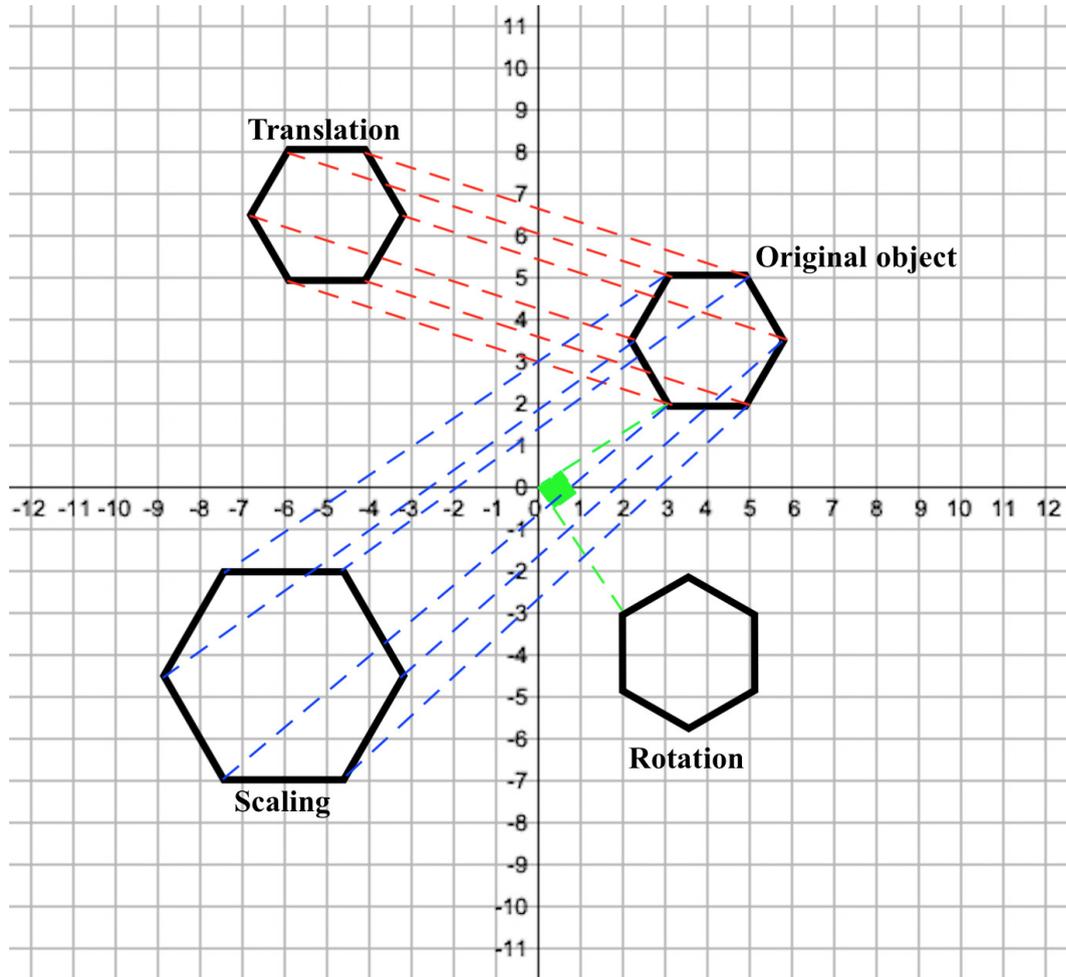


FIGURE 3.13: Sample of affine transformation process of a hexagonal shape.

For multi-modal registration, other similarity methods have been developed to accommodate issues related to weaker associations between intensities, especially on different imaging modalities. The correlation ratio [163] assumes that corresponding intensities are functionally related at registration and information-theoretic measures like mutual information [159] assume only that a probabilistic connection between landmark intensities is optimized at registration. This mutual information, MI is defined as:

$$MI(A, B) = H(B) - H(B|A), \quad (3.63)$$

where A and B are two images and correlation of Shannon's entropy [164]:

$$H(A) = \sum_{i \in A} p_i \log \left(\frac{1}{p_i} \right), \quad (3.64)$$

where, p_i is the probability of current pixel i gray value in image A . The probability p_i

can be calculated from the image's histogram. MI can be interpreted as the reduction in image B 's ambiguity, in which when the MI is high, images A and B are similar.

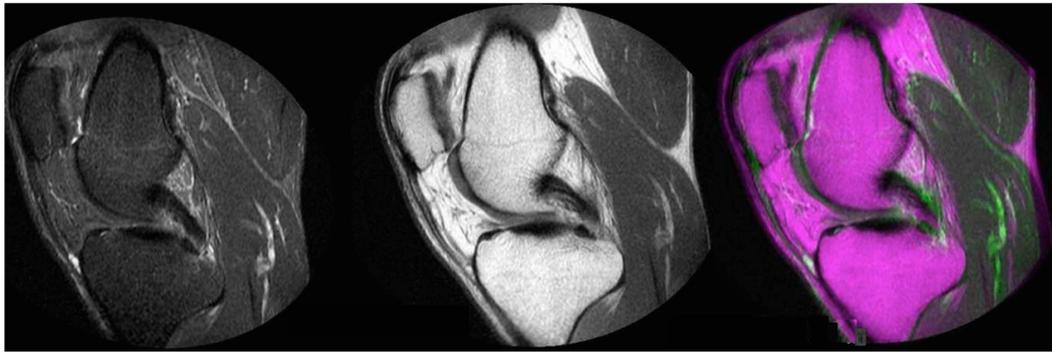


FIGURE 3.14: Registering multimodal MR images by intensity-based image registration. Image on the left represents moving image. Mid image - Fixed (template) image. Image on the right - Registered moving image with affine registration mode - Registration is done by using Matlab. *Courtesy: <http://uk.mathworks.com/help/images/registering-multimodal-mri-images.html>*

3.2.2.7 Markov Random Field

Markov Random Field (MRF) is also known as undirected graphical model (UGM) or Markov network [165]. It is a powerful method for modelling spatial continuity and can be used within image segmentation, where all pixels in the image are treated as nodes in a graph. All nodes are linked and each pixel has an authority to its neighbouring pixels and these are demonstrated in Figure 3.15. Each node contains a Markov distribution integrated with it, which consists of the probability of the pixel fitting to each class and dependable to its closest neighbours.

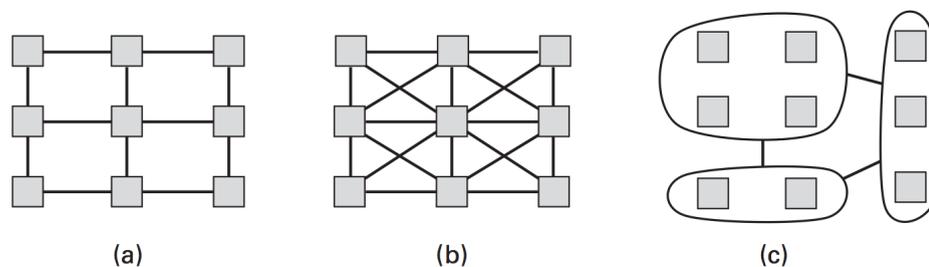


FIGURE 3.15: Markov model graph. (a) Grids of a simple 4-connected pixels. (b) Grids with extra connections. (c) Grids with irregular pattern. *Courtesy: Andrew Blake and Pushmeet Kohli - Introduction to Markov Random Fields*

In image segmentation, MRF goal is to find the segmentation (x,y) that maximizes the probability $p(x,y|z)$, where z is the observed image to be segmented [166]. (x,y) can express several different segmentation classes (x segmentation vector; y intensity

inhomogeneity vector) for each pixel and makes MRF an ideal solution for multi-label segmentation. By using Bayesian [166] rule this relationship becomes:

$$p(x, y|z) \propto p(z|x, y)p(x)p(y). \quad (3.65)$$

In this function, $p(z|x, y)$ is the probability of observing an image z given a segmentation (x, y) . $p(x, y)$ is the probability of a segmentation, and can be used to model how a segmentation result should look like. $p(z)$ is considered to be a normalization constant, and is therefore ignored in the calculations. Structures of interest can be segmented by creating different expressions for $p(z|x, y)$ and $p(x, y)$.

3.2.3 Third Generation

Third generation algorithms were developed to fulfil the growing demand for automation analysis in medical imaging (especially MR imaging and CT scanning) and corresponds with the development of current *state-of-the-art* medical imaging technology. While automation processes are clearly desirable for clinical scanning applications, its reliability in terms of processing speed and compatibility are still arguably volatile and unstable. Listed here are some of the concepts that lead to the development of *state-of-the-art* in medical image segmentation techniques.

3.2.3.1 Atlas-based

This approach employs the concept of atlas formation and registration, where atlas image is generated (interactively or statistically from manual annotation of several similar shaped objects or regions) and compiled for its properties on the structure that requires segmenting. This atlas is then used as a reference frame for segmenting the targeted images and applied in the spatial domain of the image rather than in a feature space.

Typical atlas-based approaches consider segmentation procedures as a registration problem [160]. The initial step involves discovering transformation of new image to be processed within the same dataset, which usually has similar distinguishable spatial frames as the template atlas image. This warping process may be implemented by linear transformations [167–169] or also by nonlinear transformations [170–172]. This segmentation

technique has been employed mostly in MR imaging of various brain structures [170] and spatial volume extraction [173].

[174] demonstrates a fine example of atlas-based method, in which a variational principle for joint registration and segmentation model with a given formulation of:

$$\min E(v, \tilde{C}) = \text{Seg}(I_2, \tilde{C}) + \text{dist}(v(C), \tilde{C}) + \text{Reg}(I_1, I_2, v), \quad (3.66)$$

where I_1 is atlas image containing the atlas shape C , I_2 is image to be segmented and v is the vector field. The transformation process is centred in I_2 and non-rigid deformation will be defined between the I_1 and I_2 . The first *Seg* term on the right hand side of equation 3.66 represents the segmentation functional. \tilde{C} is the boundary contour of the anticipated anatomic shape of I_2 . The second term measures the distance between the transformed atlas $v(C)$ and \tilde{C} of targeted image and the final term denotes the function of non-rigid registration I_1 and I_2 . This joint registration-segmentation model is demonstrated in Figure 3.16.

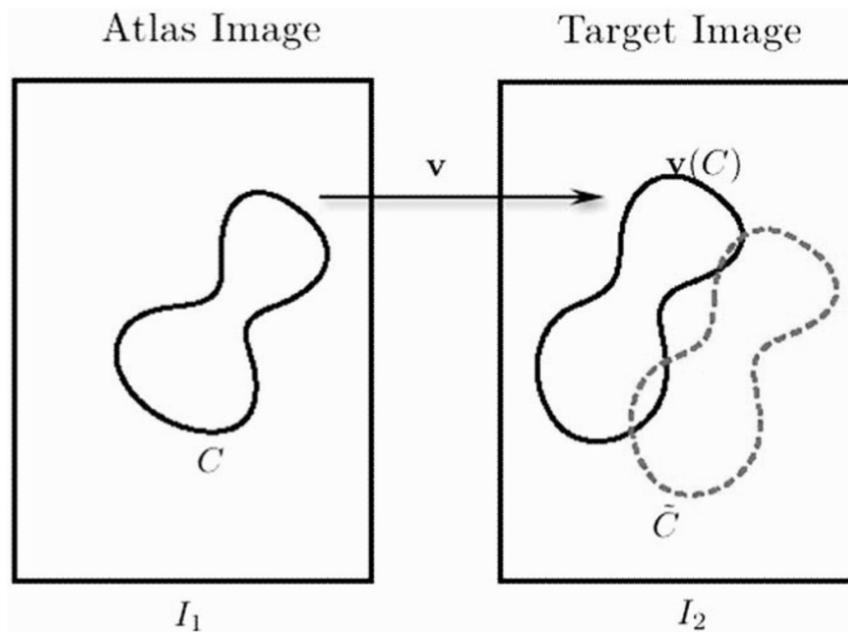


FIGURE 3.16: Segmentation model of joint-registration method. *Courtesy: Wang et al. - Joint Registration and Segmentation of Neuroanatomic Structures From Brain MRI*

The spatial matrices of an atlas image, such as feature, geometric and intensity are easily transferrable to moving or targeted image and this is one of the main advantage by applying this approach as it allows for inexpensive computational iterations and bring useful prior information to segmentation and registration tasks [175]. However,

segmenting complex structures such as cerebral cortex, often proved to be difficult and with the integration of methods like probabilistic atlases [176], the efficiency in modelling anatomical variation can be enhanced considerably. This however, requires additional processing time and interaction for data accumulation.

3.2.3.2 Shape Prior

Deformable model driven shape prior technique is arguably the most common technique in shape prior based image segmentation. The template shape boundary can be manually approximated or through statistical shape analysis from training datasets or images. Like previous atlas-based method, shape prior evolves the initial shape of the template boundary according to the image to be segmented, with restricted contour deformation protocols.

In [177], the deformation process is done locally based on image gradients and globally towards the probability of maximum a posteriori (MAP) of a shape model's position and model. [178] incorporates variational prior info into geodesic active contours, where the function is minimized when deformable model has detained high image gradients and shape prior. The integration of shape model in [177] into a reduced Mumford-Shah functional by [137] has been proposed in [179, 180], where implicit shape prior is reliant on shape vector and pose vector. [97] proposed a shape model by level-set representation and segmentation procedure by a modified geodesic active regions. And finally an integration of modified Mumford-Shah function to two statistical parametric shape models introduced in [136], for better segmentation results even with the presence of noise, occlusion or cluttered background.

Another variational model by using local and global image information with a PCA driven geometric shape prior is presented in [181], given a function of:

$$F = \beta_s F_{shape}(C, x_{pca}, x_T) + \beta_b F_{boundary}(C) + \beta_r F_{region}(x_{pca}, x_T, u_{in}, u_{out}), \quad (3.67)$$

where:

$$F_{shape} = \int_0^1 \hat{\phi}^2(x_{pca}, h_{x_T}(C(q))) |C'(q)| dq, \quad (3.68)$$

$$F_{boundary} = \int_0^1 g(|\nabla I(C(q))|) |C'(q)| dq, \quad (3.69)$$

$$F_{region} = \int_{\Omega_{in}(x_{pca}, x_T)} (|I - u_{in}|^2 + \mu |\nabla u_{in}|^2) d\Omega + \int_{\Omega_{out}(x_{pca}, x_T)} (|I - u_{out}|^2 + \mu |\nabla u_{out}|^2) d\Omega, \quad (3.70)$$

where C is the active contour, $\hat{\phi}$ is the shape function of the object of interest given by the PCA, x_{pca} is the vector of PCA eigencoefficients, h_{x_T} is an element of a group of geometric transformations parametrized by x_T (the vector of parameters), g is an edge detecting function, Ω_{in} and Ω_{out} are the inside and outside regions of the zero level set of $\hat{\phi}$, u_{in} and u_{out} are smooth approximations of the original image I in Ω_{in} and Ω_{out} and β_b , β_s , β_r are arbitrary positive constants that balance the contributions of the boundary, shape and region terms. This shape prior function advances the work from [178] with integration of shape model from [177] Mumford-Shah function in [134, 182], where it enforces the boundary to get a shape of interest, regardless of the boundary's position.

Other shape prior based image segmentation approaches include deformable template [183], superquadrics and hyperquadrics [184] and landmark-based model [148].

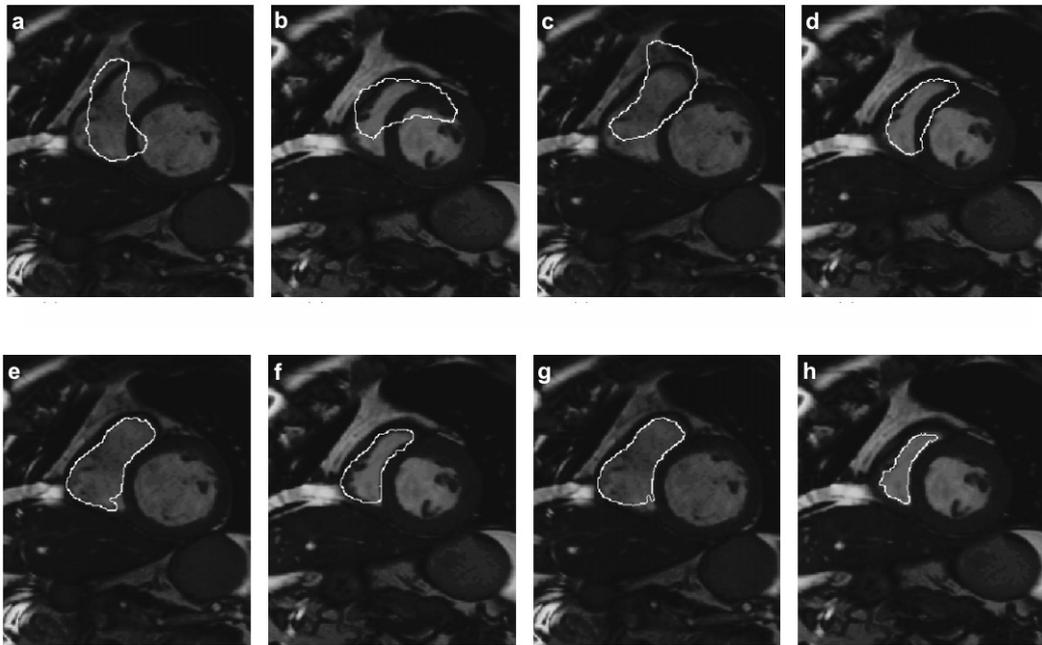


FIGURE 3.17: Image segmentation by statistical shape prior using manifold learning on right ventricular on cardiac MR image. (a), (b), (c) and (d) represent original image with respected initialization of shape prior curve. (e), (f), (g) and (h) - segmentation results, from respective top row. *Courtesy: <http://www.litislab.fr/members-cpetitjean/shapepriorsegm/>*

3.2.3.3 Neural Network

A neural network or sometimes referred to as Artificial Neural Network (ANN) is an information processing model that is inspired by how neurons in our brain, process data and learn. The fundamental element of this approach is the unique configuration of the data processing system. It is comprised of a large amount of highly unity and interconnected processing elements to solve particular problems. Like human, ANN learns by experience or example and typically designed for a specific purpose, such as data classification or pattern recognition, through a learning process [185]. Learning in biological systems involves adjustments to the synaptic networks that exist between the neurons and ANN also functions in such ways. Although the architecture and algorithm involved in modern neural network are substantially complex, any enthusiast can fairly certainly understand the foundation of operational concept, structure and function of ANN. Figure 3.18 shows the model of an artificial neuron, which forms the basis for designing a large family of ANN.

Just as in biological neurons, artificial neuron takes in some number of inputs, x_1, x_2, \dots, x_n , each of which is multiplied by its respective weight, w_1, w_2, \dots, w_n . These weighted inputs are summed together to produce $R_k = \sum_{i=0}^n w_i x_i$ [186]. In many cases, computation also includes a bias constant, b_k . The signal is then passed through an activation function f , and produce the output of $y_k = f(R_k + b_k)$. This output can be used to stimulate other neurons in the network.

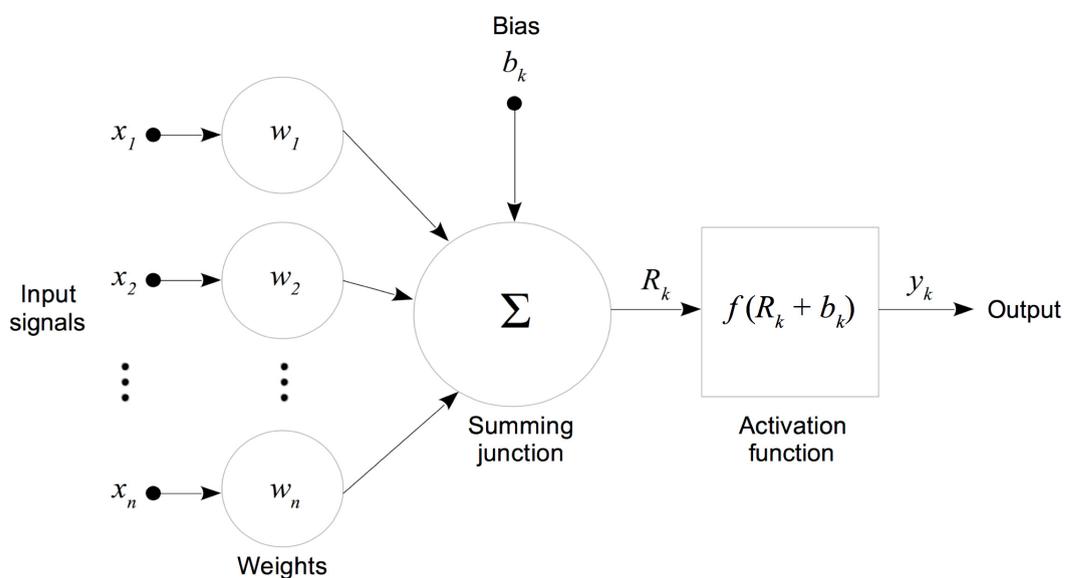


FIGURE 3.18: Schematic model for a neuron in an ANN.

Convolutional Neural Network

Convolutional Neural Networks (CNNs) are comparable to ordinary ANNs, which are supposedly constructed by artificial neurons that have learning capabilities of weights and biases. The whole network still expresses a single differentiable output probability from raw input image pixels on one end to feature extractions or classifications at the other. In short, CNN architecture in basic building block is a list of “layers” or “operations” that transform the input image into an output classification based on image’s class score [186]. Some of distinct types of layers in a CNN include convolution, rectified linear unit (ReLU), pooling (or sometimes referred to as downsampling) and fully connected (or sometimes referred to as classification layer), each of which accepts an input 3-D volume and transform it to an output 3-D volume through some dedicated functions. Figure 3.19 summarises the foundation of arguably one of the earliest and established CNN architecture by Yann LeCun [187], that later propelled the field of deep learning.

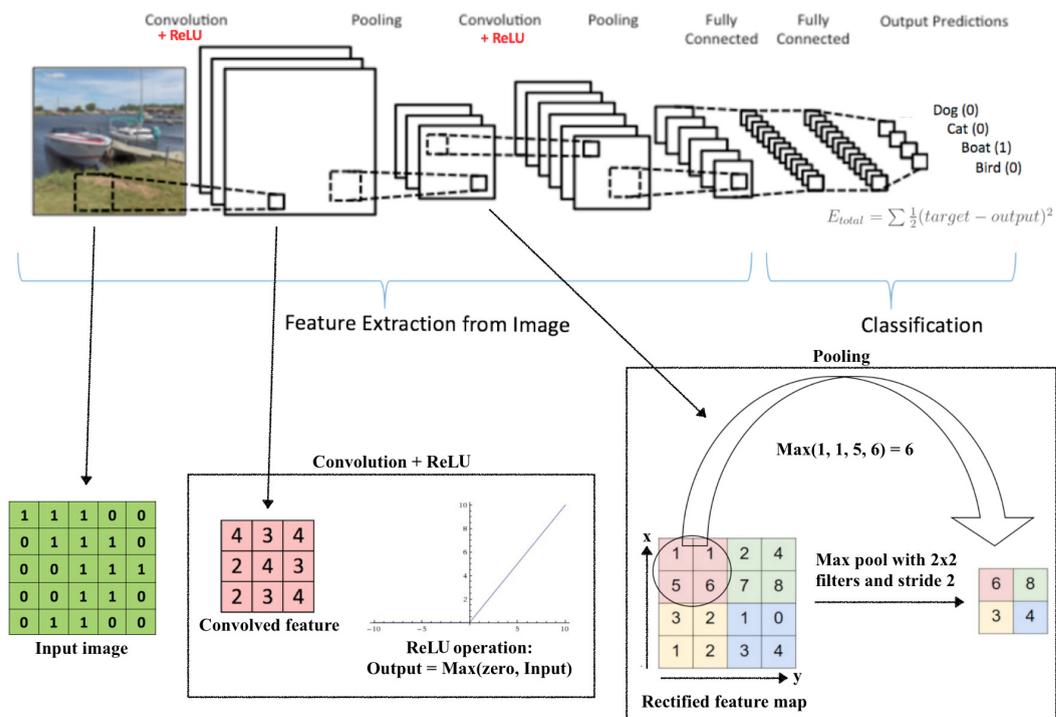


FIGURE 3.19: Architecture of a simple CNN model. *Courtesy: <https://www.clarifai.com/technology>*

Figure 3.19 illustrates a CNN model with two sets of Convolution, ReLU and Pooling layers. The second Convolution layer performs convolution procedure on the output of

the first Pooling layer using six feature detectors (or sometimes referred to as kernel maps or filters) to produce a total of six feature maps. ReLU is then applied separately on all of these six feature maps. Max Pooling was later performed individually on each of these six rectified feature maps. Together, these layers extract useful descriptions and features from input image, introduce non-linearity in the network model and reduce feature spatial while aiming to build the extracted features rather diversify to scale and translation [188]. The output of the second Pooling layer acts as an input to the Fully Connected layer, in which every neuron in the previous layer is connected to every neuron on the next layer by using classifier activation function such as softmax or Support Vector Machine (SVM), with an objective of exploiting feature-rich output from previous layers for classifying the input image into various classes based on the training dataset [188].

In general, the network model uses an input image as a form of it's training purposes by forward propagation step and then discovers the output probabilities for each class. Since weights are randomly allocated for the initial training computation, output probabilities are also random. Total error at the output layer is measured by $E_{total} = \sum \frac{1}{2} (Ideal_{output} - Actual_{output})^2$ [186]. Next, "Backpropagation" is applied to compute the gradients of the error with respect to all weights in the network and gradient descent is later engaged to update all weights and parameter values to reduce the output error. This whole process is known as training and crucial to the network model so that all of it's weights and parameters are optimized to precisely recognize patterns or classify images from the training dataset.

Solver Types

In CNNs, the solver algorithm is introduced to organize model optimization by synchronizing the forward pass and backward pass to adjust parameter updates for the weights for minimizing the loss. In other words, the learning responsibilities are distributed between the solver employed and the network model to optimize the parameter update of the weight that contribute to the minimization of loss function and gradient. Stochastic Gradient Descent (SGD) is arguably predominantly used in the methodology of training a deep learning model, mainly due to its strength of simplicity in implementation and fast processing, even for problems (or datasets) with many training patterns.

Generally, for a dataset D , the objective of optimization is the average loss over all $|D|$ data instances throughout the dataset. This is given in a formulation of [189]:

$$L(W) = \frac{1}{|D|} \sum_i^{|D|} f_W(x^{(i)}) + \lambda r(W), \quad (3.71)$$

where $f_W(x^{(i)})$ is the loss on data instance $x^{(i)}$ and $r(W)$ is a regularization term with weight λ . Practically $|D|$ can be a very large number, hence stochastic approximation is used in each solver iteration to achieve the objective and by drawing a mini-batch of $N \ll |D|$ instances, the formulation becomes:

$$L(W) \approx \frac{1}{N} \sum_i^N f_w(x^{(i)}) + \lambda r(W). \quad (3.72)$$

f_W is quantified in the forward pass, while the gradient ∇f_W is computed in the backward pass. The parameter update ΔW is produced by the solver from the error gradient ∇f_W and the regularization gradient $\nabla r(W)$.

Stochastic Gradient Descent (SGD). This solver method [190] updates the weights W by a linear combination of the negative gradient $\nabla L(W)$ and the previous weight update V_t . The learning rate α is the weight of the negative gradient. The momentum μ is the weight of the previous update.

The following equations are applied to compute the update value V_{t+1} and the updated weights W_{t+1} at iteration $t+1$, given the previous weight update V_t and current weights W_t :

$$V_{t+1} = \mu V_t - \alpha \nabla L(W_t), \quad (3.73)$$

and

$$W_{t+1} = W_t + V_{t+1}. \quad (3.74)$$

Learning rate α and momentum μ “hyperparameters” may require a manual tuning for best results, depending on the task of the image/problem to be trained. (General recommended initialization value for deep learning with SGD for: $\alpha \approx 10^{-2}$ and $\mu = 0.9$).

Adam. Proposed by *Kingma et al.* [191], this solver is conceptually similar to SGD as it is a gradient-based optimization method, which also includes an “adaptive moment

estimation" (m_t, v_t) and can be considered as a generalization of another solver type of AdaGrad [192]. The weight update formulas are:

$$(m_t)_i = \beta_1(m_{t-1})_i + (1 - \beta_1)(\nabla L(W_t))_i, \quad (3.75)$$

$$(v_t)_i = \beta_2(v_{t-1})_i + (1 - \beta_2)(\nabla L(W_t))_i^2, \quad (3.76)$$

and

$$(W_{t+1})_i = (W_t)_i - \alpha \frac{\sqrt{1 - (\beta_2)_i^t}}{1 - (\beta_1)_i^t} \frac{(m_t)_i}{\sqrt{(v_t)_i + \epsilon}}. \quad (3.77)$$

with proposed default value of $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$ (recommended by *Kingma et al.* [191]) for `momentum`, `momentum2`, `delta`, respectively in Caffe [193] platform.

Other solver types include AdaDelta [194], AdaGrad [192], Nesterov's Accelerated Gradient (NAG) [195] and RMSprop [196].

3.3 Segmentation Performance Assessment

Several numerical coefficients have been proposed as a mode of segmentation similarity comparison over the benchmark ground truth image, object or sample datasets for the past few years [197]. However, only three of the most widely used coefficients for binary data in segmentation accuracy of image processing will be presented here.

3.3.1 Jaccard Similarity Index

Also known as *Jaccard Index* or *Jaccard Similarity Coefficient*, *JSI* [198] is described as the measure between the intersection and the union of the two associate images and utilizing region property algorithms to measure the corresponding spaces.

$$JSI = \frac{|S_{gt} \cap S_s|}{|S_{gt} \cup S_s|}, \quad (3.78)$$

where S_{gt} and S_s are the manual ground truth region space and segmented region space, respectively, for a specified image, object or tissue class.

3.3.2 Dice Similarity Coefficient

Also known as *Sorensen Index*, *Sorensen-Dice Index* or *Dice's Coefficient*, *DSC* is another popular similarity comparison tool in image processing [199] and given by the formulation of:

$$DSC = \frac{2 \times |S_{gt} \cap S_s|}{|S_{gt}| + |S_s|}. \quad (3.79)$$

3.3.3 Zijdenbos Similarity Index

Another method to measure similarity or accuracy is *Zijdenbos Similarity Index (ZSI)* [200], which was derived from the kappa, κ coefficient [201]:

$$ZSI = \frac{2 \times \{S_{gt} \cap S_s\}}{\{S_{gt}\} + \{S_s\}}, \quad (3.80)$$

which results in a coefficient function that is identical to *DSC* as above.

3.4 Materials

This section covers the experimentation material for the research, where MR images of the thigh were collected locally at Manchester Metropolitan University laboratory, from 18-93 year-old men and women. All scans were acquired using the same T1-weighted Turbo 3-D sequences using a 0.25-Tesla MRI scanner (Esaote G-scan; Italy). All MRI scans consisted of serial transverse-plane slices, each with 0 mm inter-slice gap and image matrix of 256×256 (in pixels). Each MRI dataset consists of either 13 scans (of 10.0 mm to 11.3 mm thickness) or 26 scans (of 6.3 mm thickness). Figure 3.20 and 3.21 show the datasets with 13 and 26 scans each, respectively. The top left image represents the first scan followed by the serial scans next to it. These are 2 samples of typical dataset configurations that were processed in the work presented in this thesis (especially for methods analysis in Chapter 6 and 7).

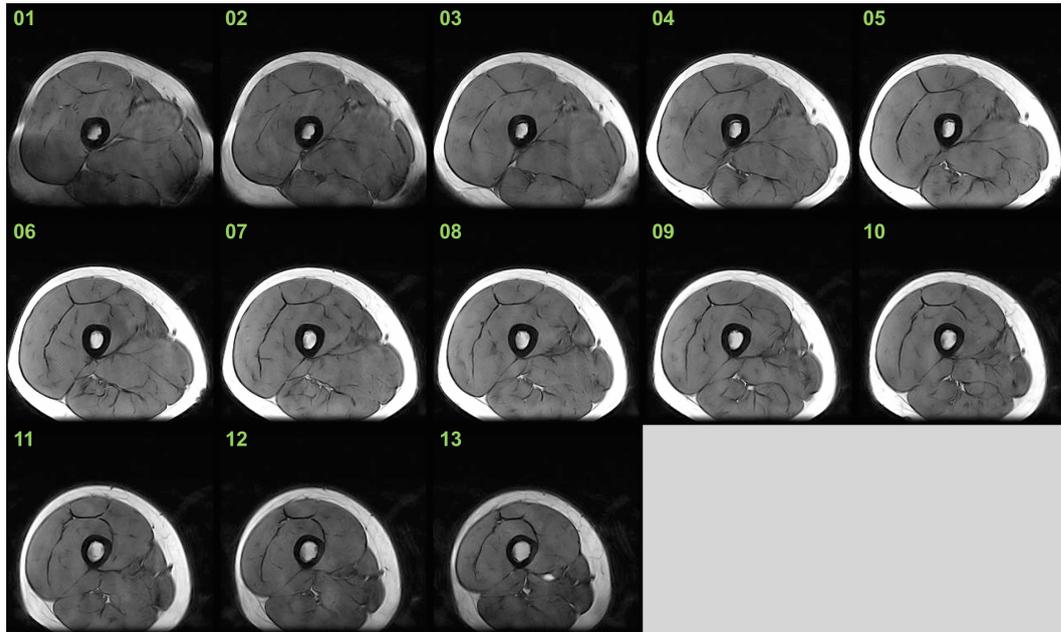


FIGURE 3.20: Dataset with 13 scans.

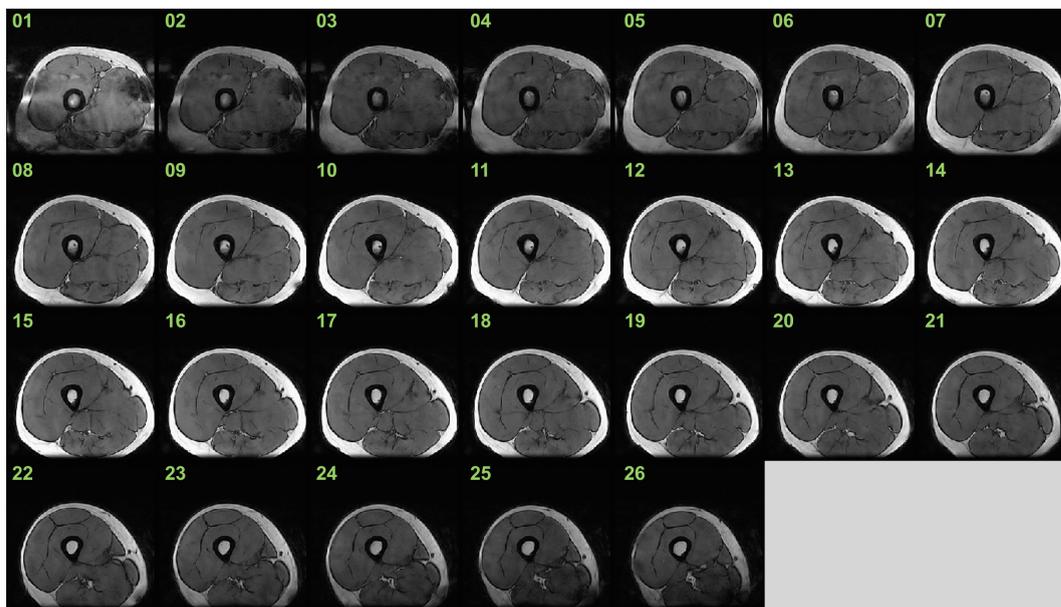


FIGURE 3.21: Dataset with 26 scans.

3.5 Discussion and Conclusion

Methodologies in medical image segmentation were classified into three groups, each indicating the level of those algorithms. The earliest group comprises of the first generation algorithms, in which these methods denote a low-level processing technique. The second group contains algorithms that examine the image's parametric or geometric

models, optimization (or minimization) methods, statistical analysis and feature extraction models. The final group is occupied by algorithms that are capable of incorporating knowledge and associates with frameworks developed towards fully automatic medical image segmentation and provides an outline for classifying the wide variety of possible methods towards automation solution. The research is focused primarily on the segmentation of MR images. However, most methods listed here may also be applied to other image types and to images from other modalities.

Thresholding and edge detection are often employed as part of a pre-processing stage in a sequence of image processing analysis. Their main limitations are that, in their simplest form, only two classes are generated, and cannot be applied to multichannel images. Furthermore, thresholding does not consider the spatial properties of an image and as a result, making it more sensitive to noise and bias field, which is a common occurrence in MR images. Both of these artefacts effectively disrupt image histogram, making segmentation more challenging.

A primary goal of statistical shape analysis is to describe the variability in any given data population of geometric objects and additionally as a mean of dimension reduction from data abundance. Principal component analysis (PCA) is considered a standard technique to accomplish this goal. However, its limitation is that it is only effective for computation of data within Euclidean vector space and futile against more complex shape representations.

Although atlas based and shape prior based segmentations are becoming less interactive and more user independent (or in a simple word automated), their dependency towards the low- and mid-level algorithms are still present. In fact, any future enhancements of such entry- and mid-level algorithms would present important developments in medical imaging as these may result in a more robust “atlas or shape memory” development towards prior knowledge, optimize computation time (as most of these methods are processed iteratively or branch divergence) and precision in segmentation accuracy, especially on complex images. The robustness and reliability of the segmentation algorithms can also be enhanced by integrating the increasingly popular Artificial Neural Networks (ANN) through Convolutional Neural Networks (CNN), Fully Convolutional Neural Networks (FCN) or other machine and deep learning models, due to its independency of strong assumptions or specific regulators.

Jaccard Similarity Index (JSI) and Dice Similarity Coefficient (DSC) are two of the most commonly used indices in image processing as a statistical validation metric to evaluate the performance, accuracy and reliability of the results of the proposed segmentation frameworks over the reference data, range from 0 (no overlap) to 1 (complete overlap). Both algorithms are sensitive to misplacement of the segmentation label. However, they are relatively insensitive to volumetric under- and over-estimations. Shape duplicity is only captured if the deviation is volumetrically high impact as a thin disparity would not result in a large deviation from one. DSC is currently more popular than the JSI and this is inauspicious because JSI is numerically more sensitive to mismatch when there is reasonably strong overlap. Measurement results from DSC look nicer because they are higher for the same pair of segmentations compared to JSI. One of the major drawback of both algorithms is that they are unsuitable for comparing segmentation accuracy on objects that differ in size [202]. Manual reference segmentations drawn by experts normally approximate ground truth, in which case user can use them as gold standard, but not as the ground truth itself (unless can be certain that it is the image of a phantom) due to the intra- and inter-observer variability of manual segmentation.

Chapter 4

Semi Automatic Segmentation of MRI Human Thigh Muscles

This chapter serves as an introductory material within 4 contributory chapters available in this thesis. Its contents are fundamentally assembled to produce a [semi-automatic segmentation framework for thigh MR images](#).

4.1 Introduction

This chapter describes a method of semi-automatically segmenting MR images of quadriceps muscles and corresponding femur cross-sections, which can then be used by researchers and clinicians to facilitate further analysis of the underlying (patho) physiology. The semi-automatic segmentation method is based on adaptive and automatic thresholding, followed by convex hull pixel identification for unwanted background removal, manual heterogeneous digital marker line drawing and analysis of region of interest (ROI).

4.2 Manual Annotation Process

This section elaborates the preparation of thigh MR images dataset for the acquisition of ground truth image. In current practice, the thigh muscle regions are segmented using a manual annotation procedure (also known as “Gold Standard”) for the means

of output image analysis. In the science of computing however, the output of this manual annotation procedure is used mainly for data comparison measurement purposes; and secondarily for procedural phase analysis of annotation (time taken for manual image annotation experiments). Consequently, a reliability assessment (intra- and inter-operator variations) can also be observed and evaluated.

The manual annotations were done by physiologists using OsiriX [203], an advanced open-source imaging software exclusively built to process MR output image in Digital Imaging and Communications in Medicine (DICOM or .dcm) format¹. Once OsiriX was installed and loaded with MRI datasets, the manual annotation can be done straight away. Figure 4.1 and 4.2 demonstrate OsiriX environment and a manually annotated MR image, respectively.

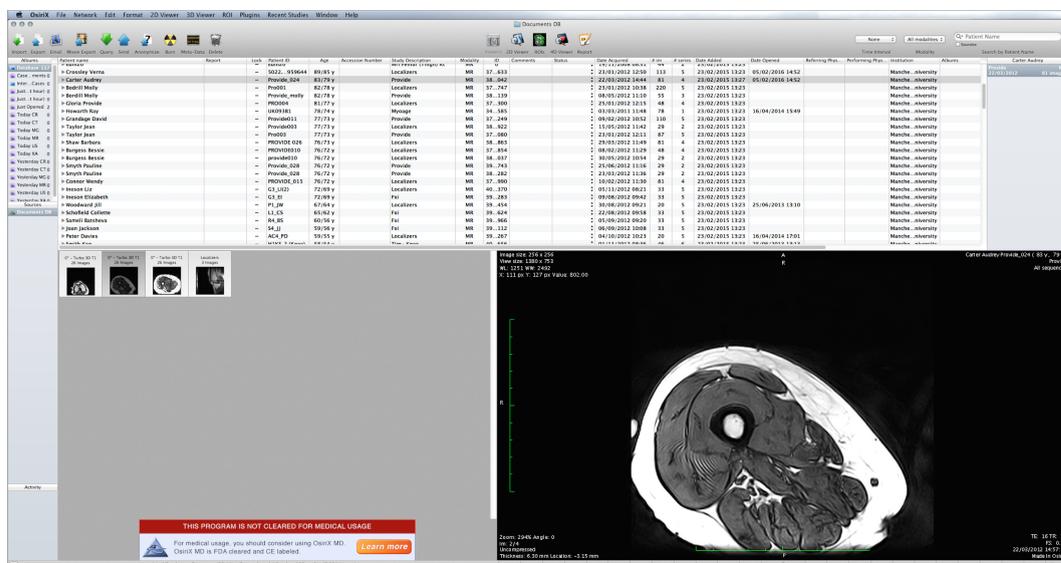


FIGURE 4.1: OsiriX layout and environment.

As can be seen on these figures, although it can be acquired and installed legitimately without a cost (freely available if it is unrelated to medical usage), the OsiriX platform and environment are considerably well designed and intuitive. DICOM formatted images are also naturally encrypted with information packages and once loaded, these information can be seen and attuned accordingly within the environment (refer Figure 4.1). Once a subject was selected, another window showing the whole MR images within dataset of the subject will be displayed. User can then scroll through individual images of this subject and start the annotation, at his/her convenient. For the best manual annotation results of the ROI, users are recommended to: firstly study about the concept

¹For DICOM images MRI datasets acquisition, please refer [here](#).

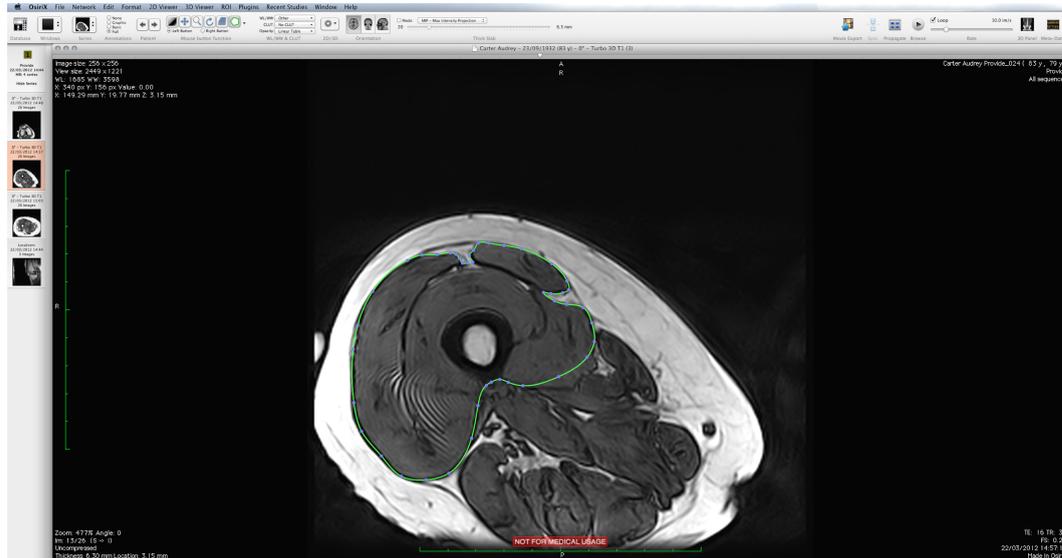


FIGURE 4.2: Sample of manual annotation by OsiriX.

of quadriceps anatomy and its general geometrical configurations and orientations; and secondly refer/analyse the previous and the next MR images in the series/dataset, prior to segmentation process.

Some other features including, 2-D Viewer (which includes subtraction, annotation, convolution filters, etc.), 3-D Viewer (which includes 3-D volume and surface rendering, 3-D endoscopy, etc.) and a section for plugins installations to unlock or explore extra OsiriX features, mostly from independent developers. For visual purposes, some of 3-D models of manually segmented dataset are demonstrated and these can be seen in Appendix A.1.

4.3 Methods for Semi Automatic Segmentation

4.3.1 Image Pre-processing

To improve image contrast and intensity distributions from different tissues and image background, noise filtering and correction of intensity inhomogeneity methods are used primarily to increase the segmentation accuracy and reliability. For images that badly affected by intensity inhomogeneity, the correction method, described as nonparametric non-uniform intensity normalization (N3) method [204], from dedicated Medical Image Processing, Analysis and Visualisation (MIPAV) application [205], is mainly used in

this proposed method for image pre-processing. N3 algorithm is pulse sequence independent, without need of prior dataset training and insensitive to pathological data that might otherwise disrupt model assumptions. To eliminate the dependence of the field estimate on tissue structure, an interactive approach is employed to estimate both the multiplicative bias field and the distribution of the true tissue intensities [204].

The algorithm of correction of intensity inhomogeneity is classically originated from this equation [204, 206, 207]:

$$I(r) = F(r).T(r) + N(r), \quad (4.1)$$

where observed image $I(r)$ consists of true image $T(r)$ modified by an inhomogeneous coil sensitivity $F(r)$ and the added noise $N(r)$. Noise has most influence on low intensity pixels and considering a noise-free case, and using the notation $\hat{I}(r) = \log[I(r)]$ the image formation model estimation becomes:

$$\hat{I}(r) = \hat{F}(r) + \hat{T}(r). \quad (4.2)$$

Non-uniformity distribution $\hat{F}(r)$ of N3 correction method is then treated by employing suitable correction strategy, field estimation, implementation details and image smoothing to acquire the exceptional pre-processed output result. Figure 4.3(a) demonstrates the badly affected non-uniformity distribution on an image and after application of N3 correction; Figure 4.3(b).

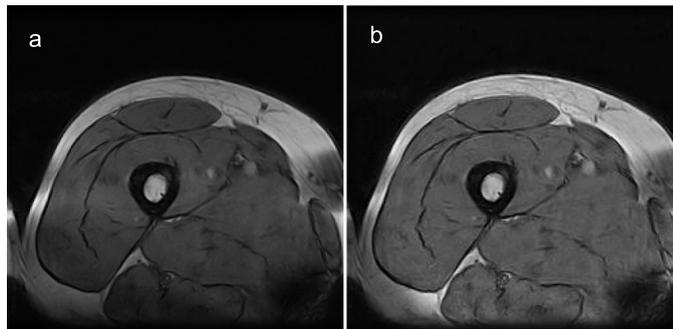


FIGURE 4.3: MRI N3 correction of intensity inhomogeneity method by MIPAV. (a) Before and (b) After correction.

Intensity inhomogeneity in images like Figure 4.3(a) can naturally be observed on up to 2 to 3 earlier and later MRI scans in a dataset, mainly due to the surface coils used in this MRI machine. Surface coils are very popular because they are a RF-receive-only coil and have a good signal-to-noise ratio for tissues adjacent to the coil. In general,

the sensitivity of a surface coil drops off as the distance from the coil increases, giving a poor quality, bias field (black patches) affected scanning results in earlier and later few MRI scans in the series (dataset).

4.3.2 Background Removal

4.3.2.1 Adaptive and Automatic Grayscale Threshold

Figure 4.3 shows prominent structure components of thigh tissues, with femur cortical layer (outer bone) and image background having low pixel intensity value, followed by muscle tissues and well-contrasted adipose tissues, bone marrow (inner bone) and intermuscular fat (or sometimes referred to as inter-muscular adipose tissue, IMAT) that have high-intensity signals. As there are only 3 apparent region intensities involved, the methods of adaptive and automatic thresholding segmentation are employed to segment different tissues of the thigh but mainly to register high intensity adipose tissue and bone marrow. This method involves conversion of the original MRI image format to grayscale, followed by adaptive and automatic grayscale thresholding based on the image grayscale histogram intensity level. Finally, the morphological method of removing any small objects within both muscle and adipose regions is implemented to refine the product, as illustrated in Figure 4.4(c).

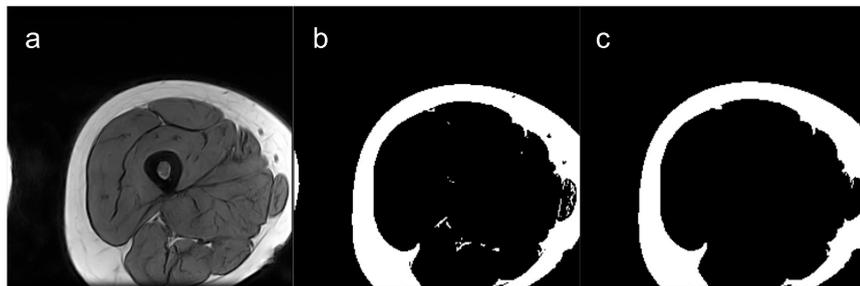


FIGURE 4.4: Adaptive and automatic grayscale thresholding for unwanted background removal. (a) Original grayscale. (b) Application of proposed threshold method. (c) Refinement result.

4.3.2.2 Convex Hull Pixel Identification

The method described in Section 4.3.2.1 will remove all high intensity pixels that are representative of adipose tissue, bone marrow and intermuscular fat. Therefore, convex hull

is used to preserve the original masks value of binary muscles, femur and intermuscular fat (if any) and removes everything outside this region.

Convex hulls are a common feature in computational geometry that are used for pattern recognition, regression, collision detection, estimation, spectrometry, cartography, topology and as a preliminary step to solve many seemingly unrelated problems [208]. The convex hull of a set S of random points in 2-D Euclidean space is defined to be the smallest convex polygon that contains all the elements of S . If you were to shrink wrap S , the shape, the plastic shrink-wrap formed around S , describes the convex hull.

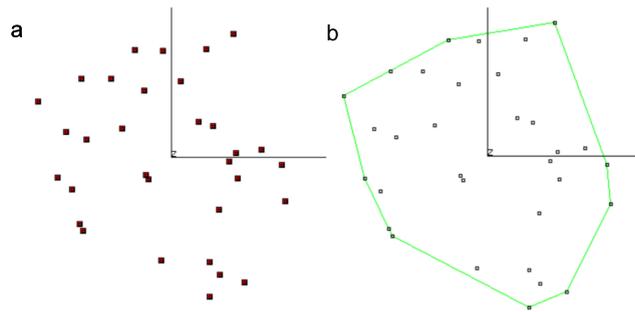


FIGURE 4.5: (a) A set of S random points in 2-D Euclidean space. (b) Convex hull of points in (a).

Following the creation of the masks with the applied convex hull algorithm on thresholded binary in Section 4.3.2.1, the unwanted pixels of skin and adipose tissue and even pixels from the adjacent thigh (Figure 4.3(a) and 4.3(b)) can be eliminated and the high intensity pixel value of bone marrow and intermuscular fat tissue (if any) is retained, as shown in Figure 4.6 below. There are some minor binary pixels corresponding to thigh skin in Figure 4.6(b) but due to the size and format restrictions, these can barely be seen.

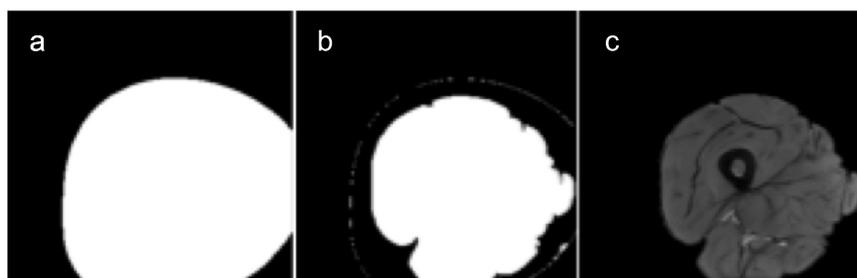


FIGURE 4.6: (a) Application of convex hull method from output of Figure 4.4(c). (b) Skin and adipose tissue removal. (c) Mask of binary blob ROIs in (b) reverted back to original grayscale value.

4.3.3 Heterogeneous Digital Line Drawing

The second stage is ROI identification. As stated earlier, this often proves difficult to automatically segment, due to the limitation of output from the MRI device where images are shown in grayscale and have substantial noise, echo, overlapping of pixel or voxel intensities, various thigh tissue compositions, inter- and intra-MRI slice inconsistency; and even for a well-trained physiologist it can be challenging to differentiate muscles borders (e.g. Figure 1.1(b)). At the time of writing [209], there are no automatic segmentation methods published on quadriceps and it appears the only semi-automatic quadriceps segmentation method [9], suffers from segmentation inconsistency and inaccuracy. Therefore, a step for the operator to interactively draw a line to delineate the border between the quadriceps and other muscles of the thigh (red line as shown in Figure 4.7(b)) is included.

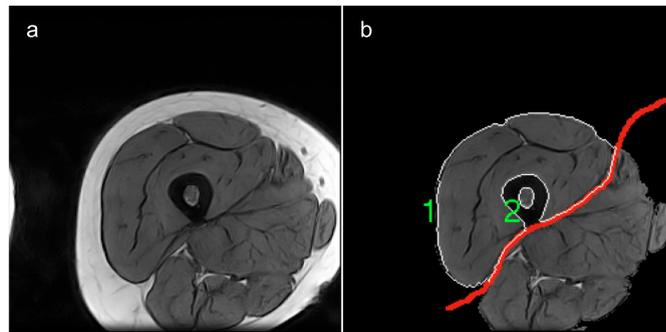


FIGURE 4.7: (a) Original MR image. (b) Segmented output result, where red digital marker line is the only interaction needed from the operator.

The manual interaction allowed us to confidently identify objective ROIs, i.e. quads muscles and femur bone, and to remove the unwanted regions (hamstring, adductors and abductor muscles).

4.3.4 ROI Processing

In the final stage, region property algorithms were developed to process ROI for each semi-automatically segmented MR image, and generate parametric data for all ROI components (cross sectional area (CSA) of quads muscles, bone marrow and bone cortical wall, in pixels).

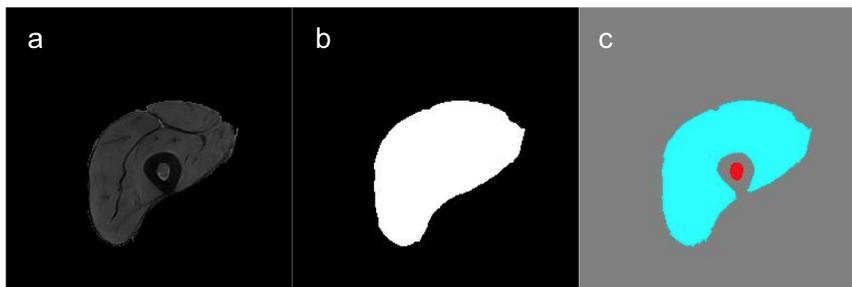


FIGURE 4.8: (a) Product of proposed MRI semi-automatic segmentation. (b) Binary image of total ROI. (c) Binary blob masks of ROI in (a).

To measure the performance of the proposed algorithms, the results obtained from this semi-automated process were compared to results from the “Gold Standard” manual segmentation (annotation) by the physiologist in OsiriX, for each associated MRI images.

4.4 Experiments and Results

Thigh MRI scans from 100 subjects (MRI output profile of Figure 2.9(f) - thigh MRI of young subjects), with 2 mid-scan slices per subject (between images no. 4-23 for dataset with 26 scans or images no. 4-10 for dataset with 13 scans) were selected from the database and semi-automatically segmented using procedures described in Section 4.3 above. The main objective is to segment the total CSA of the quadriceps muscles and femur (cortical and marrow) bone. The “Gold Standard” of manual segmentation of this ROI (named Total Area or TA) for all 200 images was carried out by a physiologist using OsiriX [203], in a procedure as described earlier in Section 4.2. To measure segmentation similarity or accuracy between output images from the semi-automated method and ground truth manual segmentation by OsiriX, a standard validation method in medical image processing/segmentation [210–212] of *Jaccard Similarity Index (JSI)* [198] is employed, as an extent of spatial overlap.

As described earlier in Chapter 3, JSI is a measurement method between the intersection and the union of the two associate images and utilizing region property algorithms to measure the corresponding spaces (for formulation please refer equation 3.78), where S_{gt} and S_{sa} are the manual ground truth region space (acquired from OsiriX) and semi-automatically segmented region space (from the developed method), respectively, for a specified tissue class. Another method widely used to measure similarity or accuracy is

Zijdenbos Similarity Index (ZSI) [200], in which the application of its algorithm varies slightly from *JSI* (again, for formulation please refer to equation 3.80).

The measurements have an average *JSI* of 0.95 (95% accuracy) and an average *ZSI* of 0.98 (98% accuracy) for all 200 images. The small difference between *JSI* and *ZSI* was likely due to the fact that both algorithms introduce different metrics to compute distances between such regions in the respective specified tissue space. However *ZSI* gave a slight increment of similarity/accuracy compared to *JSI*.

There was one poorly rescaled OsiriX manually segmented *TA* original image to the standard 256×256 pixels format used for image analysis and similarity comparison. However, even with such incompatibility and mismatching between image resolutions during accuracy measurement, *ZSI* still gave a reading of 0.86, while *JSI* gave a reading of 0.75. This suggests *ZSI* could be more tolerant/forgiving hence why the accuracy measurement restriction and robustness standard of *JSI* was preferred. There is no definite indication as to what value these indices should be for any processed image, but several papers [9, 213, 214] suggest an index value of 0.7 or above is considered accurate.

The standard deviation (SD) of output results was 0.0091 and this indicates that 68% of the segmented images fall between 0.9444 (94.44%) to 0.9627 (96.27%) of similarity/accuracy (Mean $\pm 1SD$) and 99% of the segmented images fall between 0.9261 (92.61%) to 0.9809 (98.09%) of similarity/accuracy (Mean $\pm 3SD$), which indicates accurately exceptional in segmented output images from the semi-automated method.

The research analysis mainly focussed on *TA*, which consists of quads, femur cortical and bone marrow area. However, the images were also processed in order to separately analyse the areas of the quadriceps, the cortical bone and the trabecular (marrow) bone. Measurement of area of quads and bone marrow, in cyan and red colour, respectively (such as in Figure 4.8(c) and 4.9(b)), indicate the region's CSA in pixels. CSA of femur

TABLE 4.1: Summary of results for 200 processed MRI images.

ROI	Mean <i>JSI</i>	Standard Deviation (<i>SD</i>)	Average Manual Annotation Time	Average Semi-Auto Annotation Time
<i>TA</i>	0.95	0.0091	132 sec	17 sec

cortical (the bonal hard tissue) is acquired by subtraction of TA to accumulation of quads and bone marrow.

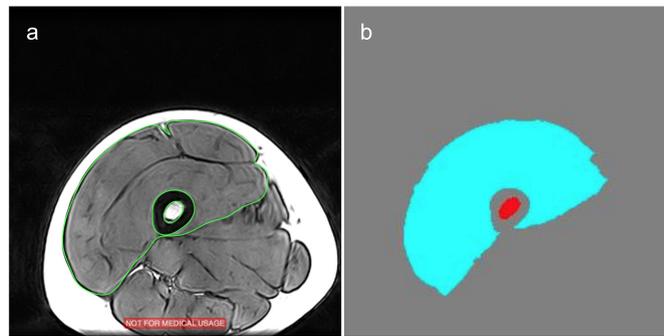


FIGURE 4.9: (a) Manual segmentation by OsiriX. (b) Product of semi-automatic segmentation method.

The average time required for manual processing of one MR image in OsiriX (in this case, thigh region) was 132 ± 42 seconds. The mean processing time required for unsupervised analysis of one slice was previously reported as 52 ± 7 seconds, without the need of any user interactions [8]. Thus, around 66% of time could be saved compared to the method in [8]. The proposed algorithms are developed to improve speed and efficiency and increase the ability to distinguish between different tissue types and generate quantitative data. The method is able to analyse (segmentation and CSA computation) the desired ROI in less than 17 seconds per image (time calculated including the manual process stated above), which represents a time saving of 87%. Three OSX (presently known as macOS) based machines were used (with different configurations of processor, memory, GPU and internal disc speed) for manual ROI segmentation of the thigh MRI by OsiriX to evaluate of any affects towards the time domain during the manual segmentation procedure. Conclusively, the differences are not perceptible and worthy of any discussion as measurements of the ROI segmentation are done through a simple OsiriX closed polygon application. Conspicuous time differences are recorded however when a complex OsiriX application such as 3-D volume rendering or 3-D surface rendering is used.

4.5 Discussion

In this chapter, a semi-automated segmentation method for analysis of MRI of human thigh muscle and bone was proposed. Compared with conventional manual analysis

of such images, the semi-automatic approach has cut the analysis time by 87% (132 seconds has been reduced to 17 seconds) and provides an exceptional segmentation accuracy of 95% (by *JSI*). This will be of use to researchers and clinicians aiming to distinguish between and determine changes over time to muscle, bone and fat tissue with development, ageing, diet and disease.

To further enhance the functionality and practicality, this method can be refined at the image pre-processing step. At the time of writing [209], MIPAV is used to universally pre-process MR images that badly affected by intensity inhomogeneity. The long processing time, due to complexity of N3 correction algorithms is the main reason as to its exclusion into the proposed method. To overcome this issue, a built-in correction of intensity inhomogeneity, based on image property, compatibility and simplicity of equation 4.1, from image-feature based automatic correction method in [206] will be included in future framework's development. The correction method of *Koivula et al.* [206] is based on the evaluation of a compensating function by using bright pixels from an image, which have low relative noise to approximate the anti-log of equation 4.2, i.e.:

$$I(r) = F(r).T(r). \quad (4.3)$$

The compensating $C(r) = \langle F(r)^{-1} \rangle$ can then be analysed when:

$$T(r) = I(r).C(r). \quad (4.4)$$

Measuring homogeneous phantom and applying a proper Gaussian averaging filter on an image are two possibilities to generate $C(r)$.

The major issue related to intensity inhomogeneity affected images (such as MRI output profile shown in Figure 2.9(a) and Figure 4.5(a)) is ineffective tissue segmentation, which corrupts the precision of final segmentation output. As part of the framework is using convex shield based methods, convex hull algorithm suffers the most under this circumstance due to unrecognised distribution of tissue intensity. Therefore, for future development, a simple but effective intensity non-uniformity correction pre-processing technique will be included in the framework.

Besides the novel approach of convex shield pixel identification for background removal, the only operator interaction, manually segmentation line drawing of *TA* inner posterior

border, is also a highlight in this chapter. Due to consistent composition of tissues in thigh MRI images, especially those in mid-scan image of a dataset used here (MRI output profile as shown in Figure 2.9(f) - thigh MRI of young subjects), this technique will be further developed with an automated segmentation application to measure important variables on MRI scans, not just to MRI of the thigh, but also to any MRI scans with similar configurations or properties.

Another important reason this technique is preferred because it requires very little time processing, which would not have been possible had the inclusion of other state-of-the-art segmentation methods are adapted. At this stage, the technique does not require the creation of templates or shape prior for feature extraction or bundled of iterations to analyse the edges of tissue and intensity variability, which making ROI calculation smooth and efficient.

Inter- and intra-operator variability is another crucial component to be considered for further explorations in the future. Manual processing of images can be prone to inter-assessor variability and it is also the case that no two semi-automated segmentations would give an identical result, even if the same person completed the segmentation on the same image. We have begun to determine the inter-operator variability of our technique (results not reported here).

The proposed method was developed by analysing two mid-scan slices (between images no. 4-23 for dataset with 26 scans or images no. 4-10 for dataset with 13 scans) from MRI datasets obtained from 100 people (MRI output profile of Figure 2.9(f) - thigh MRI of young subjects). These MR images were selected from a much larger database that includes MRI scans from around 300 different people. By testing the method against this complete database of images, a more reliable, mature and stable semi-automated segmentation framework can be developed. However, there are some issues related to our image acquisition that must be discussed. The database of images was derived from a 0.25-Tesla MRI scanner (Esaote G-scan, Genoa, Italy), which generates a lower intensity of magnetic field and more limited scanning area compared with more common 1.5-Tesla or 3-Tesla MRI scanners. As explained earlier in Section 2.5, about 20% of the whole MRI database is made up of images such as in Figure 2.9(a)-(e), where Figure 2.9(a)-(c) are usually related to the technicality during the image acquisition. However, the MRI slices within the thigh images, especially those in the middle, where images are tend

to be less affected by bias field, are comparable to those collected from more powerful scanners. Therefore, the semi-automated analysis technique is proven of its robustness and can be used to analyse thigh images derived from any MRI scanner from 0.25-Tesla or higher. Nonetheless, processing and testing new MRI datasets that are acquired from other MRI machines are also admirable and will be considered for our future work.

4.6 Conclusion

A semi-automatic segmentation framework that separates and measures quadriceps, bone marrow and femur cortical area in MRI scans of the human thigh has been successfully developed. The method needs minimal supervision and yields accurate results, with mean accuracy performance by *JSI* of 0.95 (95%) in a short processing time (average processing time of 17 sec per image).

Further development to a reliable fully automated ROI segmentation framework would be the ultimate goal, at this point. And such developments will include a built-in correction of intensity inhomogeneity, without compromising the overall processing time and accuracy.

Chapter 5

Automatic Segmentation of MRI Human Thigh Muscles

This chapter outlines the procedure for automated segmentation of human thigh magnetic resonance images. Datasets that included sequential transverse-plane images (approximately 20 sequential images) were used to develop the framework and the images of the mid-thigh were selected to establish the segmentation algorithm because these images contain essential parametric and geometric information that represents the subsequent images of the same dataset. Hence, if the mid-thigh can be decomposed accurately, it would significantly improve the prospects of decomposing all remaining transverse images within the dataset. The segmentation framework was designed based on the integration of statistical signal processing and medical imaging algorithms and the results showed good segmentation accuracy and fast processing time.

5.1 Introduction

Skeletal muscle strength is a main component of functional capacity and when strength of leg muscles is low it can impact adversely on a person's mobility levels and quality of life. For these reasons, strength is commonly assessed in research studies into physical health [14–18]. The main interest is to understand the association between muscle strength and functional deficiencies, what causes such conditions and its mechanisms. To date, it is still not clear what causes the age associated reductions in muscle mass

and strength, but they are ascribed to an inherent ageing process [215], that is often modified by dietary deficiencies [216, 217] and/or disease [218].

Muscle strength is often presented as the maximum amount of force that someone can exert, regardless of body size or weight and naturally it is higher in those with larger muscles, but there is also variation between people in the force that can be produced per unit muscle mass and this is sometimes referred to as muscle “quality”. The muscle quality can be estimated from a simple division of maximum force into muscle cross sectional area. Measurements of skeletal muscle force are straightforward and require the patient simply to apply as much force as possible against an immovable object. The object includes a strain gauge that measures forces. However, measuring skeletal muscle mass is more difficult [219]. There are various techniques commonly used, ranging from limb-circumference which is fraught with error, through to the modern clinical X-ray or magnetic resonance imaging techniques. For any measurement of the muscle quality, it is clearly vital to include an accurate measurement of the skeletal muscle size and in this respect the MRI is the gold-standard. The aim of this chapter specifically and this thesis as a whole, is to develop a method to automatically and accurately determine muscle size from thigh MRI scans.

Magnetic resonance imaging (MRI) is considered the gold standard for assessing muscle size, and has the additional benefits that it can help in diagnosing muscular pathologies. It is non invasive, non-ionizing and can also identify fat infiltrations and bone size. Automatic segmentation techniques have been successfully developed and published [7, 8, 68–77], with some generating excellent segmentation accuracies in a short period of time.

Indeed, establishing automated system for individual muscle segmentations within skeletal muscles or within the same muscle group of the same body segment is a key challenge in medical imaging because it is technically difficult to achieve due to poor contrast between connective and muscle tissues, and sometimes none or little information of muscle borders to assist the segmentation, even within the same muscle group. With this in mind, the most commonly used approach is to manually segment the regions of interest (ROI) and this requires a very good knowledge of human anatomy. This process is operator-dependent, tedious, time-consuming and can have questionable reproducibility by novice users. The issue holding back development of automatic segmentation of

individual muscles from MR images is that individual muscles share similar MR properties and can appear indistinguishable from neighbouring muscles. So far, only a few methods targeting automatic segmentation of individual muscles have been reported [11, 13, 78–80], delivering a promising initiation for in depth muscles segmentation for future investigation.

This chapter describes the development of an automated segmentation method for MRI of the thigh, based on combination frameworks of shape model analysis, edge detection by phase congruency, dynamic transformations of shape model depending on image to segment and smoothing procedure by fuzzy c-means. Segmentation is focused on mid-scan thigh MR image of one dataset, with ROI that consists of quadriceps, femur and marrow.

5.2 Data

MRI scans of the thigh were collected from men and women aged 18-90 years [220]. All scans were collected using the same T1-weighted Turbo 3-D sequences using a 0.25-Tesla MRI scanner (Esaote G-scan; Italy). All MRI scans consisted of serial transverse-plane slices, each with 6.3 mm thickness and 0 mm inter-slice gap and image matrix of 256 x 256 (in pixels).

Generally, automatic segmentation of MRI scans often proves difficult due to image artefacts, echo and other substantial noise. The low *signal-to-noise ratio* (*SNR*) output images are affected by gradient pulse eddy current effects and inherent effects of pulsed radiation on MRI radio frequency (RF) coils from the MRI machine. Other potential elements that may further escalate this issue are the internal variations of electro-chemicals and biochemistry, fluid circulation and tissue density of the scanned subject, the lack of precisely defined muscle boundaries and issues related to intensity inhomogeneity or bias field across an image.

In addition, the sensitivity of a surface coil decreases as the distance from the coil increases, resulting in a significant amount of noise especially in the most proximal and distal scans located at the limits of the scanning coil's detection. For this reason, the mid-scan MR image was selected (image no.7 for dataset with 13 scans or image no.13 for

dataset with 26 scans)¹ as this is least affected by the disturbances and inconsistencies described above.

5.3 Methods for Automatic Thigh MRI Segmentation

This section is divided into two subsections. The first subsection covers the application of algorithms to automatically segment general components in thigh MR image, such as adipose tissue, muscle, femur, bone marrow and/or inter-muscular adipose tissue (IMAT), if any.

The second subsection highlights the major contribution of this chapter, where the proposed segmentation framework was developed by integrating the algorithms in the First, Second and Third Generations, as described earlier in Section 3.2.

5.3.1 Segmentation of Components of Thigh MR Images

The segmentation process starts with image pre-processing technique by image conversion and compatibility algorithms, followed by the removal of adipose tissue, using algorithms in Section 4.3.2, (i.e. the combination of thresholding and pixels masking by convex hull). The results are a grayscale MR image of the thigh and an image of solid binary adipose tissue as a result from thresholding.

Next, muscle, femur, bone marrow and IMAT are segregated. As there are clear distinctions in the signal intensity of these tissues, application of thresholding and region properties algorithms can distinguish between muscle and bone marrow. This is illustrated in Figure 5.1.

For subsequent steps, the image processing toolbox of morphological algorithms, such as filling holes, image subtraction, structuring element and open image, were applied. The procedures are summarized in pseudocode 1 and results of components segmentation is demonstrated in Figure 5.2.

Most, if not all of the techniques described in this subsection are combinations of algorithms within First Generation and Second Generation, as in Section 3.2 and consist

¹For DICOM images MRI dataset samples, please refer [here](#).

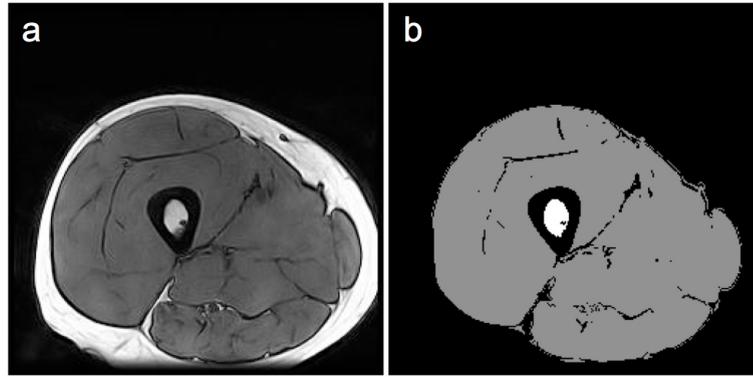


FIGURE 5.1: Application of thresholding and region properties. (a) Original MR image. (b) Muscle (gray) and bone marrow regions (white).

Algorithm 1 Automatic segmentation of major components of thigh MR image - Morphological procedure

Require: Array elements that are members of set array of muscle and bone marrow

Ensure: Original image is trimmed (without adipose tissue)

BEGIN

$component1 \leftarrow$ array elements of: *muscle and bone marrow (mbm)*

$component2 \leftarrow$ filling holes of: $component1$

$raw\ muscle \leftarrow$ image subtraction of: $component1$; $component2$

$femur\ and\ marrow \leftarrow$ open image by disk structuring element of: $raw\ muscle$

$gaps \leftarrow$ image subtraction of: $raw\ muscle$; $femur\ and\ marrow$

$marrow\ with\ gaps \leftarrow$ image subtraction of: $muscle$; mbm ; $component1$

$marrow \leftarrow$ filling holes of: $marrow\ with\ gaps$

$femur \leftarrow$ image subtraction of: $femur\ and\ marrow$; $marrow$

$muscles \leftarrow$ image subtraction of: $component2$; $femur\ and\ marrow$

END

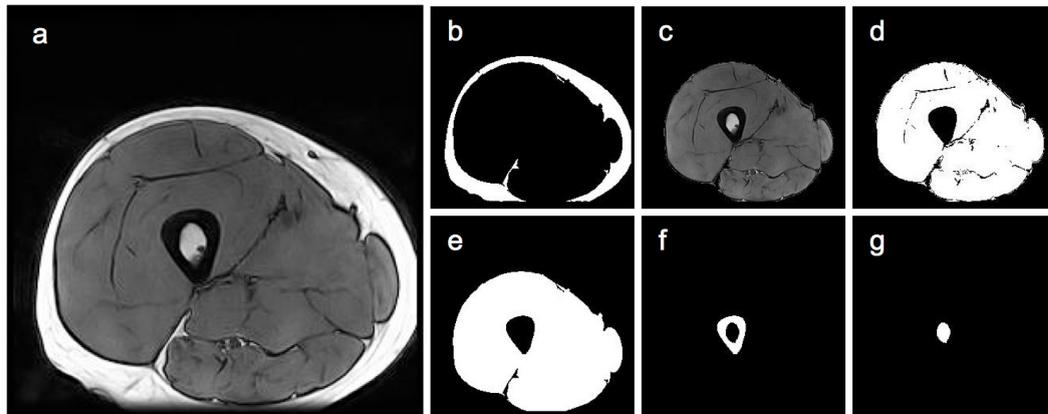


FIGURE 5.2: Automatic segmentation general components of thigh MR image. (a) Original MR image. (b) Adipose tissue (binary). (c) Muscles, femur and marrow in grayscale. (d) Absolute muscles (binary). (e) Relative muscles area (binary). (f) Smoothed femur cortex (binary). (g) Smoothed bone marrow (binary).

of algorithms that are routinely applied in earlier publications related to general thigh MRI segmentation [7, 8, 68–77], as per Section 2.4.

5.3.2 ROI (Quadriceps, Femur and Bone Marrow) Segmentation

This subsection is divided into four parts. The first part covers the construction of a shape model for the mid-scan image; second part highlights the processing techniques of the input image, which focuses on pre-processing, phase congruency and edge enhancement; followed by transformation of a shape model that is adaptable to properties of the mid-scan image; and finally curve transformation for quadriceps segmentation. The automation framework is capable of ROI segregation in a short period of time.

5.3.2.1 Shape Model for Mid-scan

The initial approach employed for automatic object segmentation covers geodesic active contours [93, 94] by integrating shape information of the ROI into the evolution process. We first compute a statistical shape model of the ROI over a training set of ROI samples. For a more robust statistical result and an average shape model that represents the best universal ROI shape, this training set is loaded with 85 manually segmented mid-scan MRI of the thigh (MRI output profile such as in Figure 2.9(f) - thigh MRI of young subjects, i.e. MRIs from which usually manifest consistent muscles composition). Boundary localization of all manually segmented ROIs were extracted and non-parametric distance map based on Euclidean metric [221] of $Ed_{ab} = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}$ that represents these ROIs correspondent to background pixels were generated and stacked together into a single matrix to eventually obtain the mean shape of the ROI by the computation of principle component analysis (PCA).

PCA is a very powerful and quick technique for pattern analysis and spatial reduction, especially for calculations that contain abundant data with identical dimensions [156]. This is true since one of its distinctive features includes linear combination of data re-expression, i.e. to define data in a different way, usually explicitly. And as a highlight, $PX = Y$ represents a basic form of PCA as an orthogonal regression method, with linear transformation of P , and X and Y are equal to $m \times n$ matrices.

Another form that best expresses data mean following the above equation is the ratio of variances σ^2 or *signal-to-noise ratio (SNR)*, $SNR = \frac{\sigma_{data}^2}{\sigma_{noise}^2}$, where $SNR \ll 1$ and $SNR \gg 1$ indicates low (much less-than) and high (much greater-than) precision data, respectively. Covariance generalization of the resultant variance of dataset K and L is

summarized as $\sigma_{KL}^2 = \langle k_i l_i \rangle_i$, where i represents a single grouped observation of the k and l variables, which after the conversion of row vectors of dataset K and L becomes equation 3.52. At this stage, each row and column of X is equal to x_i and \vec{X} (entire vector of X), respectively, resulting in a foundation covariance matrix of equation 3.53, where S_X is a symmetric dimension matrix $m \times m$ and XX^T is ij^{th} value corresponds to x_i and x_j as per equation 3.52. S_X is also refers as scatter matrix (equation 3.54), where μ is the mean vector (equation 3.55) that establishes the average form of shape model. Figure 5.3 summarizes variation modes of the shape model, following the application of PCA.

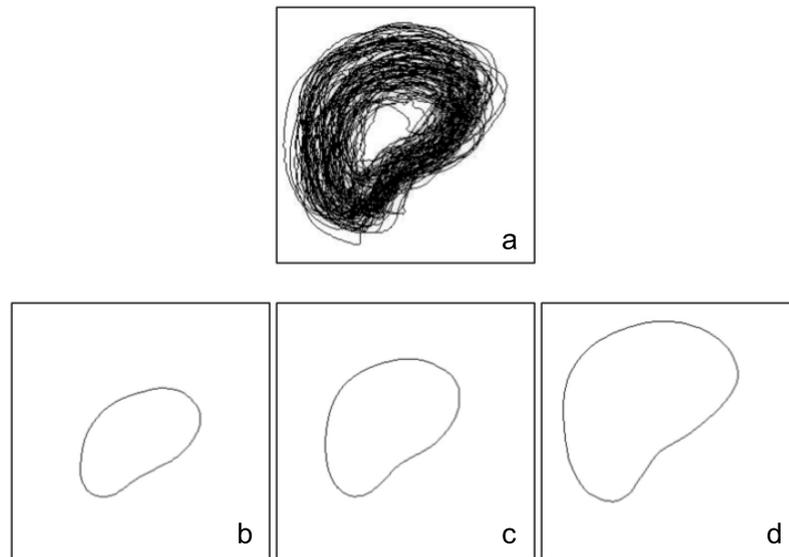


FIGURE 5.3: Major primary variance modes of the shape model training dataset. (a) Training dataset of 85 aligned shape models. (b) Variation eigenmode of mean value, -2λ . (c) Mean shape of shape model. (d) Variation eigenmode of mean value, $+2\lambda$.

5.3.2.2 Processing of Input Image

Pre-processing. Image conversion and compatibility are the two earliest image pre-processing algorithms applied, followed by combination of image thresholding and properties methods for removal of adipose tissue and skin layer that covers the thigh, without excluding the marrow or other intermuscular tissues (which both tend to share approximately similar histogram extension as adipose tissue). This prominent adipose tissue is eliminated by a simple yet efficient Otsu's thresholding method [106]. Figure 5.4 shows some mid-scan samples of before and after the pre-process stage.

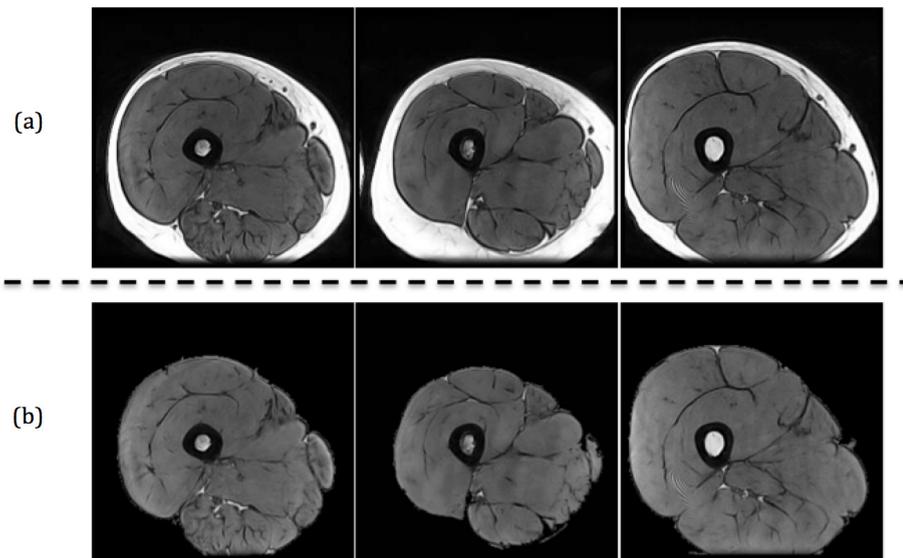


FIGURE 5.4: Outputs of pre-processing technique to mid-scan. (a) Original mid-scans. (b) Pre-processing results.

Phase Congruency. Phase congruency [222] is a low-level invariant model for feature detection (such as step edges, image corners and lines), where fundamentally, the phases of a quadrature filter responses at different scales are similar. A quadrature filter consists of even (symmetrical) section and odd (anti-symmetrical) section, both of which are the Hilbert or Riesz [223] transforms of each other. Let $u(x)$ be an even function with finite energy, and $v(x) = H(u)(x)$ be its Hilbert transform. The functions are then considered in a quadrature state. Both of these functions can be integrated as one filter by multiplying the odd section by i and adding them together, where $g(x) = u(x) + iv(x)$.

Convolving the filter with a signal, $f(x)$, generates a complex response where the even section becomes the response to the even filter and the odd section becomes the response to the odd filter, $f_A(x) = (f * g)(x)$. Phase is the argument of this complex response, where $\phi = \arg f_A(x)$ and separated from the amplitude of the response $A = |f_A(x)|$. Phase describes the shape of the signal at a point (within even and odd), while amplitude describes the strength. This independence is important as it means signals with different strengths can be compared. If the phase is the same at multiple scales, then there is likely to be signal features and this is referred to as phase congruency.

One major motivation for using phase congruency to detect image features is that it provides an absolute measure of the significance of features. Unlike edge detection

techniques, this allows one to set thresholds that are applicable across wide classes of images [224]. The other major motivation being that we are not required to make any assumptions about the luminance profile of the feature. We are simply looking for points where there is order in the frequency domain. Step discontinuities, line and roof edges are all detected. Figure 5.5 illustrates some results on mid-scans containing line features and step discontinuities of various magnitude and orientations.

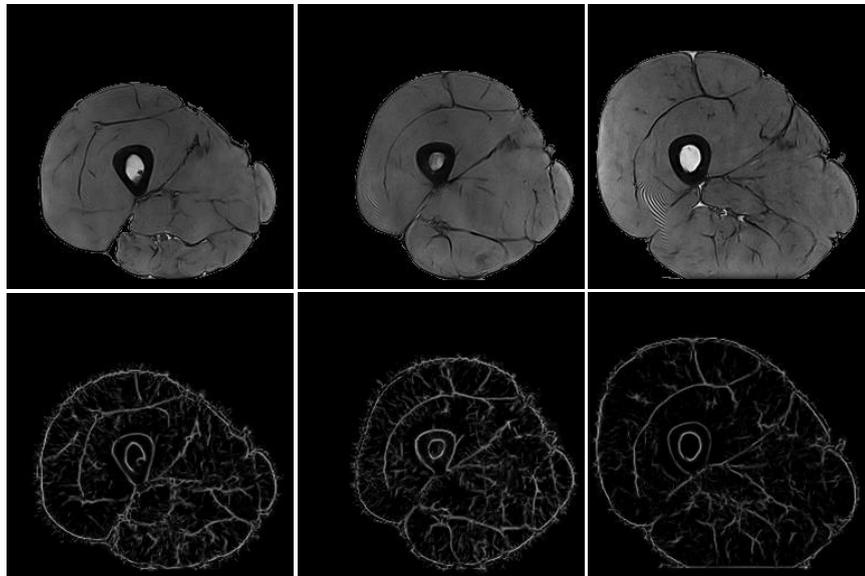


FIGURE 5.5: Feature extractions by phase congruency. (a) Top row - Pre-processed images. (b) Bottom row - Phase congruency results.

Edge Enhancement Technique. Generally, the standard MRI scanning output provides strong correlation from one image to another although the advantages are not always clear between muscle borders, even within the same scan. In order to establish noticeable borders, especially towards the border between quads and hamstring, an edge enhancement technique is included in the process. This simple stacking procedure is carried out to take advantage of the correlation between scans in one MRI dataset, by utilizing the femur as an anchor point. This non-expensive procedure ideally requires 5 mid-scan images (images no. 5-9 for 13-image dataset or images no. 11-15 for 26-image dataset) to extract influential border. Extra images to stack usually are unnecessary and practically thicken the line (border) features, due to the helical curvature of thigh muscles. Figure 5.6 shows some results on mid-scans containing line features and step discontinuities of various magnitude and orientations.

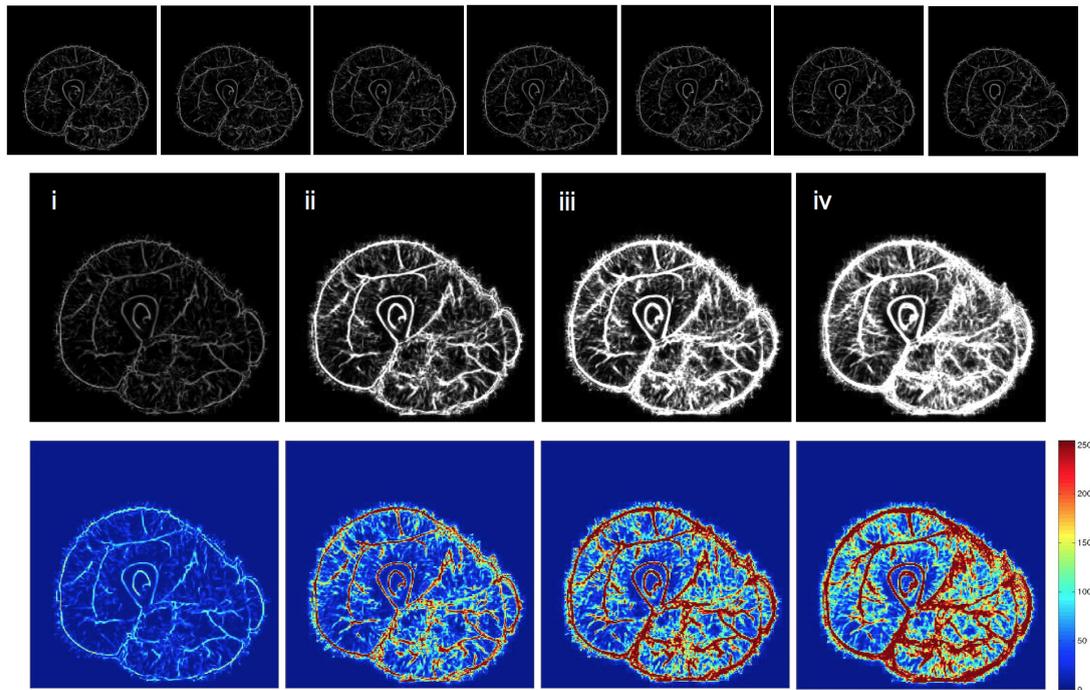


FIGURE 5.6: Stacking of phase congruency images. (a) Top row - Application of phase congruency to images (images were taken from a 26-image dataset, with the first 7 middle images (image no. 10-16) were used and processed) . (b) Middle row - i) The mid-scan phase congruency image (image no. 13). ii) Stacking of 3 mid-scan images (images no. 12-14). iii) Stacking of 5 mid-scan images (images no. 11-15). iv) Stacking of 7 mid-scan images (images no. 10-16). (c) Bottom row - Default 64-color “parula” colormap (heatmap) of (b).

Phase congruency has not been used optimally for feature detection mainly because of its response to noise. The calculation of phase congruency is ill-conditioned if all the frequency components of the signal are very small, or if there is only one (or nearly to one) frequency component present in the signal [223]. Any system that uses normalization to provide invariance to contrast and illumination must inevitably suffer from sensitivity to noise. This is crucial especially for medical or biological imaging, where precision analysis is prioritized.

This is especially an issue for images such as those shown in Figure 5.6 (iii), where 5 phase congruency images that were aligned together are visibly feature-rich and practically prone to segmentation inaccuracies. To overcome this, some image morphological techniques were included to limit the excessive features or line characteristics. The methods remove pixels so that objects without holes shrink to a point and objects with holes shrink to a connected ring; and removes pixels on the boundaries of objects without allowing objects to break apart; whilst preserving the Euler number.

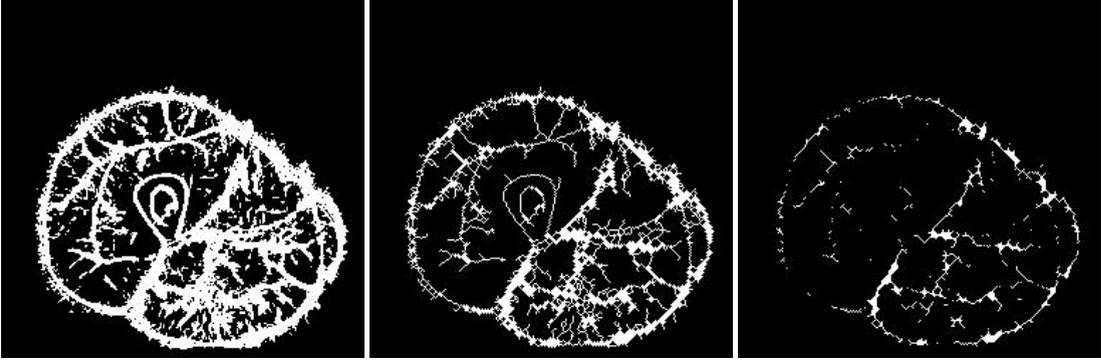


FIGURE 5.7: Image morphological methods. From left to right - sample of 5 phase congruency stacked images; object shrinking; and image skeleton process.

5.3.2.3 Transformation of Shape Model

Providing proper curve representation of the ROI shape model (obtained from the result of the mean shape of the PCA described above) based on information from image to segment (mid-scan image) is desirable and results in effective seeding points for an ideal and optimum curve deformation process later. In doing so, this section will cover the foundation of image transformations of ROI shape model employed in the process, which are part of the rigid or affine image registration, and a subset of the more general transformations.

Let's assume two 2-D images, ROI shape model SM and MRI mid-scan image I are related by registration transformation such that: $SM(x) = I(Ax + T)$, where A is the linear component matrix affine mode consists of scaling and rotation, and T is the 2-D translation vector. If both SM and I are related, then each point $SM(x_{SM}, y_{SM})$ in ROI shape model corresponds to point $I(x_I, y_I)$ according to the matrix equation of:

$$\begin{bmatrix} x_{SM} \\ y_{SM} \\ 1 \end{bmatrix} = \begin{bmatrix} s \cos \phi & -s \sin \phi & \Delta x \\ s \sin \phi & s \cos \phi & \Delta y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_I \\ y_I \\ 1 \end{bmatrix}, \quad (5.1)$$

where s is the scale factor, ϕ is the angle of rotation and Δ is the translation in (x, y) unit. Homogeneously, for any pixel of $SM(x, y)$, it is proven that:

$$I(x, y) = SM(\Delta x + s * (x \cos \phi - y \sin \phi), \Delta y + s * (x \sin \phi + y \cos \phi)). \quad (5.2)$$

For finite disconnection of SM and I , both are assumed to be square in shape, with pixel

area of 256×256 . Note that practically, the transformation of SM by modes of image registration introduces differences due to occlusions as some data moves into or out of the image frame. Figure 5.8 demonstrates the scaling and translation process of ROI shape model. This newly adapted ROI shape model will be used as a background seed mask for active contour evolution in order to segment the MRI mid-scan image later.

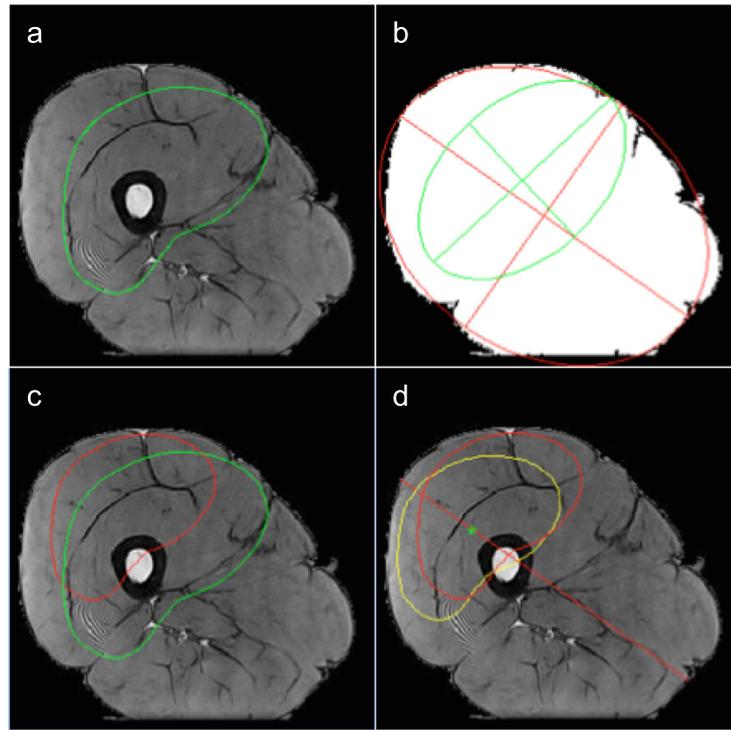


FIGURE 5.8: Scaling and translation process of ROI mean shape model. (a) Pre-processed MRI mid-scan image with ROI mean shape model (green boundary). (b) Major and minor axes of ellipse equivalent for both mid-scan, in red; and shape model (SM), in green. (c) SM scaling process, with s of equal to 0.9. (d) SM translation to a fourth of mid-scan major axis length.

5.3.2.4 Active Contour and Smoothing Process

To segment the ROI from a mid-scan image, the progress of active contour evolution was carried out both locally, based on curvature of mid-scan image, and globally to a maximum a-posteriori estimate of shape and position of the modified shape model. The initial active contour model used is a region-based energy model by *Chan and Vese* [137], which is the representation of an energy minimization-based segmentation as in original

fitting term of:

$$F_a(C) + F_b(C) = \int_{i(C)} |u_0(x, y) - k_a|^2 dx dy + \int_{o(C)} |u_0(x, y) - k_b|^2 dx dy, \quad (5.3)$$

and regularization term of:

$$\inf_C \{F_a(C) + F_b(C)\} \approx 0 \approx F_a(C_0) + F_b(C_0), \quad (5.4)$$

where both variable curve C and its dependent constants k_a, k_b are the average derivations of inside and outside image u_0 . Energy curve fitting and regularization terms above were further minimized by additional Euler-Lagrange-based H_ε and δ_ε regularizing functions as $\varepsilon \rightarrow 0$, and defined as:

$$F_\varepsilon(k_a, k_b, \phi) = \mu \int_{\Omega} \delta_\varepsilon(\phi(x, y)) |\nabla \phi(x, y)| dx dy + v \int_{\Omega} H_\varepsilon(\phi(x, y)) dx dy + \lambda_a \int_{\Omega} |u_0(x, y) - k_a|^2 H_\varepsilon(\phi(x, y)) dx dy + \lambda_b \int_{\Omega} |u_0(x, y) - k_b|^2 (1 - H_\varepsilon(\phi(x, y))) dx dy. \quad (5.5)$$

Apparent under segmentation results near the region of the ROI that corresponds to the posterior section of femur (in black pixels) is one of the disadvantages that associated with active contour as the curve transformation always try to find the configuration with the least total energy of:

$$E_{snake}^* = \int_0^1 E_{snake}(v(s)) ds = \int_0^1 [E_{int}(v(s)) + E_{image}(v(s)) + E_{con}(v(s))] ds. \quad (5.6)$$

This issue was overcome by a smoothing process of femur extraction by fuzzy c-means clustering (FCM) [20], where the intensity of the extracted femur is adjusted and overlapped with the original image, to preserve precise segmentation. Figure 5.9 demonstrates the result of segmentation, without and with the application of the FCM process as above, respectively.

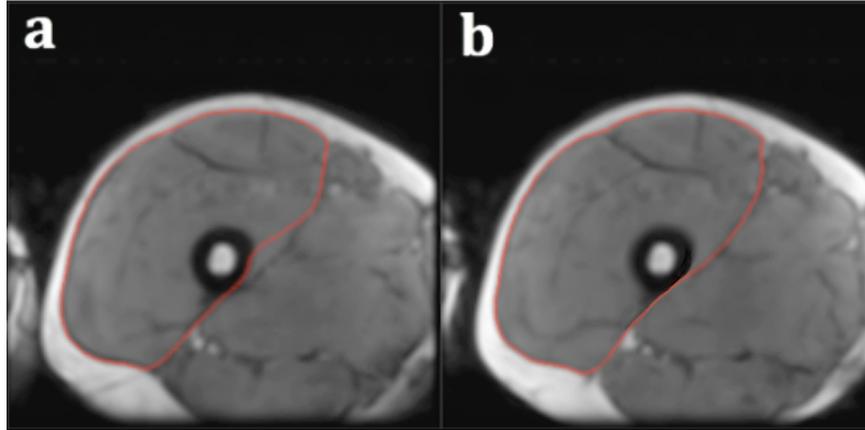


FIGURE 5.9: Segmentation results of (a) Under segmentation (due to black pixels (low energy) correspond to femur region). (b) With the application of FCM smoothing process.

5.4 Experiments and Results

Fifty-one 2-D mid-scan thigh MR images of between 18-90 year-old people were used to test the effectiveness of the combination frameworks as described above. The main intention was to segment the ROI of quadriceps, femur and marrow as precise as possible. Segmentation accuracy was quantified by *Jaccard Similarity Index (JSI)*:

$$JSI = \frac{|A_{gt} \cap A_s|}{|A_{gt} \cup A_s|}, \quad (5.7)$$

where A_{gt} indicates the area of “ground truth” of ROI manual segmentation by using OsiriX, an open-source imaging software for DICOM extension file format; and A_s indicates segmented ROI area by the proposed system. The similarity of the output results from manual segmentation and the proposed system for a corresponding mid-scan is measured by the intersection and union in Equation 5.7.

Four experiments, each of which is a diversity or an extension from one another, were performed to validate the analysis. The first experiment (Method 1) was a segmentation system as proposed in Subsection 5.3.2, but without integrating phase congruency on the input mid-scan image. Deformation of active contour of this system operates directly on the original pre-processed mid-scan, with a fixed scaling factor s , of 1.1 for the transformation of shape model.

The second experiment (Method 2) was an extension of Method 1, with the integration of phase congruency on the input mid-scan image. Active contour performs on the 5

stacked phase congruency mid-scan images, with a fixed scaling factor s , of 1.1 for the transformation of the shape model. The third experiment (Method 3) was performed exactly like Method 2, but the end segmentation result was overlaid as a new curve representation of the shape model before another active contour process on initial pre-processed mid-scan image was executed. The final experiment (Method 4) was our proposed segmentation system; an extension of the algorithms in Method 3, but with the smoothing treatment of fuzzy c-means on the femur region of original pre-processed mid-scan image. Table 5.1 summarizes the segmentation accuracy for all methods as described above.

TABLE 5.1: Average segmentation accuracy for Method 1-4.

Segmentation Method	Mean <i>JSI</i> (%)	Standard Deviation (SD)
Method 1	84.21	0.1203
Method 2	76.81	0.1502
Method 3	80.21	0.1213
Method 4	85.26	0.0915

5.5 Discussion

The main reason for experimenting with mid-scan MR images for this proposed automatic segmentation framework is because these images contain essential parametric and geometric information that represents the subsequent images of the same dataset. Therefore, if the analysis of the mid-thigh images can be deciphered accurately, it would significantly improve the prospects of interpreting all remaining transverse images within the dataset.

The inclusion of phase congruency on the input mid-scan image and smoothing function by fuzzy c-means (FCM) clustering on original pre-processed mid-scan image before another active contour process was executed clearly benefits this proposed segmentation framework. The improvement is not substantial, but enough to further explore other edge detection, edge enhancement, clustering, classification and or hybrid algorithms in the future.

Phase congruency generally associates with extraction of significant features in images. This edge detection algorithm is independent of differences in contrast and illumination

of an image, and reflects the characteristic of the image in the frequency domain. Edge-like features have many of their frequency components in the same phase and this concept is almost identical to coherence, except that it implements to different wavelength functions. FCM works by classifying membership to each data location corresponding to each cluster center on the basis of distance between the cluster center and the data location. The closer the data to the cluster center, the greater its membership towards that particular cluster center. In this case, as there are 3 distinctive regions in the mid-scan image, the extraction of hard tissue of femur bone (corresponds to the ring-shaped and black region in the center of the mid-scan) was highly effective by the application of FCM. The double active contour implementations, where the initial segmentation result (from 5 stacked phase congruency images) was overlaid as a new curve representation of the shape model before another active contour process on original pre-processed mid-scan image was executed, is required, as the initial segmented contour holds significant value of the mid-scan geometrical information.

Method 2 provides the least segmentation accuracy mainly due to the non-inclusion of smoothing treatment of FCM and double active contour implementations, as precise information of femur bone can be reproduced by FCM and robust contour that represent the quadriceps of the image to segment can be estimated by the execution of initial active contour, respectively. The application of second active contour also essential to compromise any loss of data that associate to the border of quadriceps of image to segment during edge enhancement technique and as a cheap alternative (in time domain) compared to other complex statistical or iteration-dependent algorithms, which evidently benefited Method 4 (with the incorporation of both steps - 85.26% in Method 4 versus 76.81% in Method 2). Highest value in Standard Deviation also indicates that the segmented images by Method 2 have the biggest divergence from the mean value of Method 2, which again validate the distinction of both FCM and double active contour implementations during the segmentation process.

The average manual annotation time per MRI slice (by OsiriX) was 132 seconds. In our previous works, the segmentation time was reduced to 17 seconds per image (semi-automated [209]); and further improved to just 3 seconds per image (automated dataset segmentation [220]). Average segmentation time per image by the proposed method was around 22 seconds, a distance margin compared to [220] but advantageous nevertheless as the method proposed here is fully-automated without any user interventions. The

processing time of proposed method also is a significant improvement compared to manual annotation by OsiriX (exactly 6 times faster) and comparable to the semi-automated work in [209]. Region-based active contour approach employed in the proposed framework has substantial advantages over edge-based methods, especially robustness against initial curve placement, flexibility of dealing with noise in image and customizable energy minimization model [138]. In the future, optimized values of its dependent constants k_a and k_b will be experimented to enhance the computations of average derivation of inside and outside of shape model's placement, which hypothetically may improve segmentation accuracy and reduce segmentation time.

Additional experiments focusing on the scaling factor s , of shape model was performed by tackling issue related to its placement. A fixed s value of 1.1 was used throughout all experiments (Method 1-4). However, with supervised adjustments of s value within 1.0 to 1.3, depending on the size of the mid-scan image, the average segmentation accuracy of Method 1 was successfully improved slightly from 84.21% to 86.52%. Hypothetically, this altered Method 1 may have similar precision effects on Method 4. Therefore for future developments, automated scaling factor of the shape model will be established into overall framework to optimize the segmentation accuracy. As far as the topic concerned, the procedure will ideally involve automation process of shape model's orientation (Δ) and rotation angle (ϕ) also. Figure 5.10 demonstrates the effect of scale factor s , in the segmentation framework.

5.6 Conclusion

A reliable and good precision of automatic segmentation method for mid-scan thigh image has been successfully developed by the integration of algorithms in Subsection 5.3.2. The method is fully automated and requires no interactions from the user and offers good performance in terms of segmentation accuracy of the ROI (quadriceps, femur and marrow) and processing time. However, with the replacement of a fixed value of scaling factor s , together with the applications of more dynamic and adaptive computation of orientation Δ , and rotation angle ϕ , of the shape model, the proposed segmentation method may have been a much more optimized and robust framework. These will be explored in our future investigations.

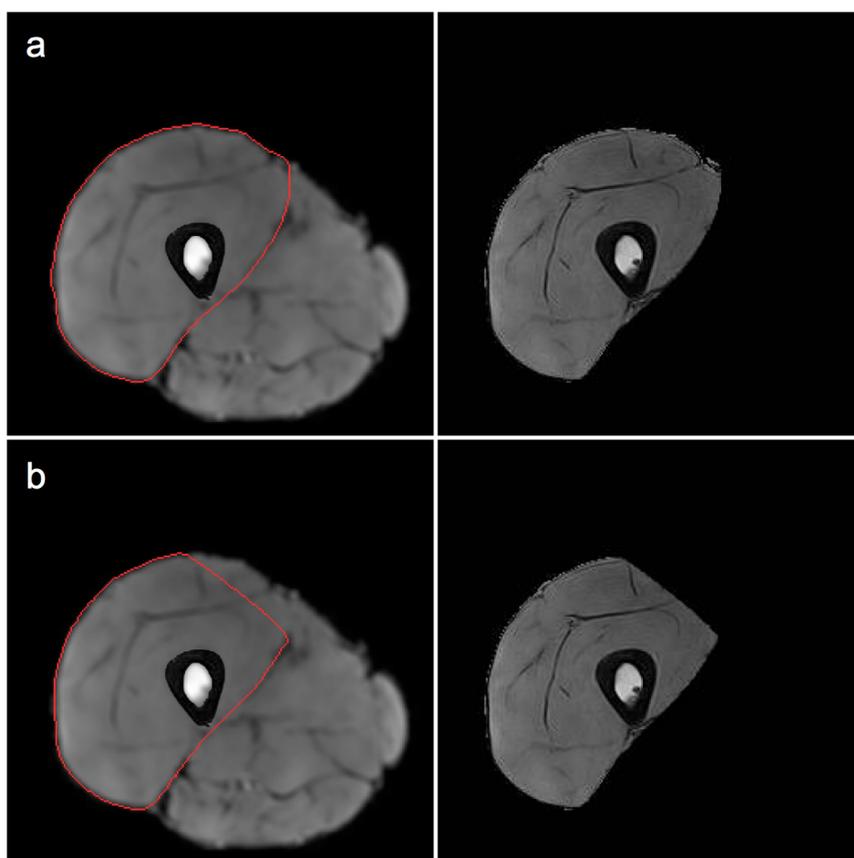


FIGURE 5.10: Sample of segmentation results on variations of s . (a) Top row: $s = 1.1$, accuracy = 95.18%. (b) Bottom row: $s = 1.0$, accuracy = 90.99%.

Chapter 6

Combination Frameworks

Leading to Automated Dataset

Segmentation

In Chapter 4, we have covered the framework for [semi automatic segmentation](#) of an MR image. In previous chapter, we also have covered the framework for [automatic segmentation](#) of an MR image. Now we will discover [the automatic segmentation framework for all images in one MRI dataset of the thigh](#).

6.1 Introduction

Image segmentation of anatomic structures is often an essential step in medical image analysis. At the time of writing [220], a variety of segmentation methods have been proposed, but none provides comprehensive evidence of efficacy, precision and robustness automatic segmentation of the thigh. In magnetic resonance (MR) images of the thigh, the segmentation is complicated by factors, such as artefacts (e.g. intensity inhomogeneity and echo) and inconsistency of soft and hard tissue compositions, especially in muscles from older people, where accumulation of intermuscular fat is greater than muscles in young person. In this chapter, combination frameworks that lead to a segmentation enhancement method for region of interest (ROI) segmentation in one dataset are demonstrated.

As discussed earlier in Chapter 1, 2 and 3, several algorithms [19–23] have been developed and widely used for basic segmentation needs in medical imaging. The application of active contour models or snakes for ROI segmentation or extraction is a common approach in computer vision and medical image analysis and to date, these techniques have been progressively developed and carefully adapted (according to datasets to be processed). Generally, the technique attempts to minimize an energy associated to the initial contour and object contour as a sum of an external or internal energy. Conventionally, active contour models can be divided into several groups, with two of the major ones being; edge-based [91, 94] and region-based [137, 182] models. Meanwhile, over- or under-segmentation commonly occurs in snakes due to its energy function sensitivity, weak or without edges and gradient variations in an image. Several researchers [177, 180, 225–228] have incorporated active contour when segmenting various kinds of images, including medical images. Therefore, to enhance the snake’s deformable model robustness or border convergence, practical customizations with other algorithms should be considered.

This chapter mainly describes the automated dataset segmentation for ROI (quadriceps, femur cortical layer and bone marrow) extractions (with minimal user interactions). The manual stage requires the user to trace a single line across the image of a single MRI slice to identify the particular ROI. This manual interaction (of framework as in Chapter 4) is re-adapted since the main goal of this framework is to achieve the exceptional ROI output segmentation as possible. Following this, a combination of automated processes uses low-, mid- and high-level image analysis algorithms (exclusively by using template-based methods) to automatically segment the ROI in serial images within the same dataset.

6.2 Methods for Automatic Dataset Segmentation

6.2.1 Semi-Automatic Segmentation

The proposed segmentation method is started by selecting a “mid-scan image” of one MRI dataset (image no.7 for a dataset with 13 scans (refer Figure 3.20); image no.13 for a dataset with 26 scans (refer Figure 3.21)). Image conversion and compatibility algorithms are part of the image pre-processing techniques that first get initiated here,

followed by an optional bias field correction method [229] that helps minimizing the effect of intensity inhomogeneity. For unwanted soft tissues (adipose tissue and skin) elimination, Otsu's [106] thresholding method is employed due to its simplicity yet powerful gray-level histogram discriminant criterion, i.e. by accurately expanding the discriminant measure of segregation of the resultant tissues in MRI of the thigh. For details formulation of Otsu's thresholding technique, please refer sub-Subsection of 3.2.1.1.

After the Otsu's thresholding, convex hull pixel identification algorithm is applied, followed by manual interactions to define muscle borders. Figure 6.1(a) shows the original MR image and the result of optional bias field correction method (on original MR image) - Figure 6.1(d), the application of convex hull with adaptive thresholding as in [209] and the application of convex hull with Otsu's thresholding (without and with the procedure of bias field correction), respectively.

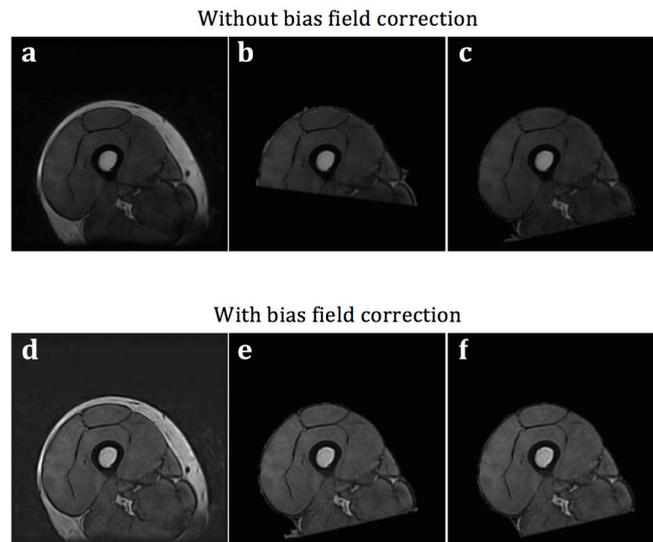


FIGURE 6.1: Pre-processing techniques. (a) Original MRI image. Unwanted background removal by: (b) Adaptive thresholding. (c) Otsu's thresholding. (d) Bias field correction application on original image (a). Unwanted background removal by: (e) Adaptive thresholding with bias field correction. (f) Otsu's thresholding with bias field correction.

Manual interactions were necessary because of the deficiencies in MRI output such as noise, echo or overlapping of pixel/voxel intensities. Such imperfections in the image of muscle tissue can be seen in a single transverse-plane MRI section (Figure 6.2) and occurred due to inherent effects of pulse radiation on MRI radio frequency (RF) coils, gradient pulse eddy current effects, internal variations of water content, local electrochemical and biochemical differences in the muscle tissue, variations in tissue density,

such as related to regional and local differences in intracellular glycogen and triglyceride storage, and inaccuracies from body movement during scanning.

Intensity inhomogeneity in images like Figure 6.2(a) can naturally be observed on up to 2-3 proximal (earlier) and distal (later) MRI scans in one MRI dataset, mainly due to the surface coils used in this MRI machine. Surface coils are very popular because they are a RF-receive-only coil and have a good signal-to-noise ratio (SNR) for tissues adjacent to the coil. In general, the sensitivity of a surface coil drops off as the distance from the center of the coil increases.

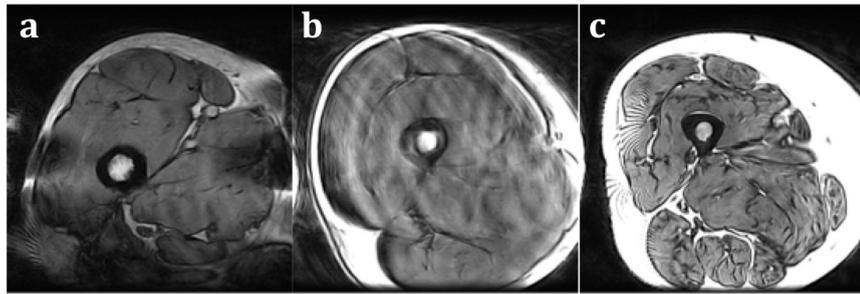


FIGURE 6.2: MRI outputs. (a) Noise patches in intensity inhomogeneity. (b) Limb movement during scans. (c) Muscles compositions in elderly and echo effects.

The final output image from semi-automatic segmentation method is basically the required ROI and template image initialization is established from this output image. Unlike [209], only parametric details from ROI boundary are registered. Figure 6.3 illustrates the process of the semi-automatic segmentation method.

6.2.2 Automatic Segmentation

The same pre-processing step of image conversion and compatibility techniques (as semi-automatic segmentation method) are again initiated and performed routinely at this early stage of the automatic segmentation method, to subsequent MR images in the same dataset as the mid-scan MR image. The influences of noise are suppressed by a Gaussian derivation before discovering the edges of the image. The following Gaussian distribution, centered at (x_0, y_0) with a standard deviation σ , is applied for this purpose.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (6.1)$$

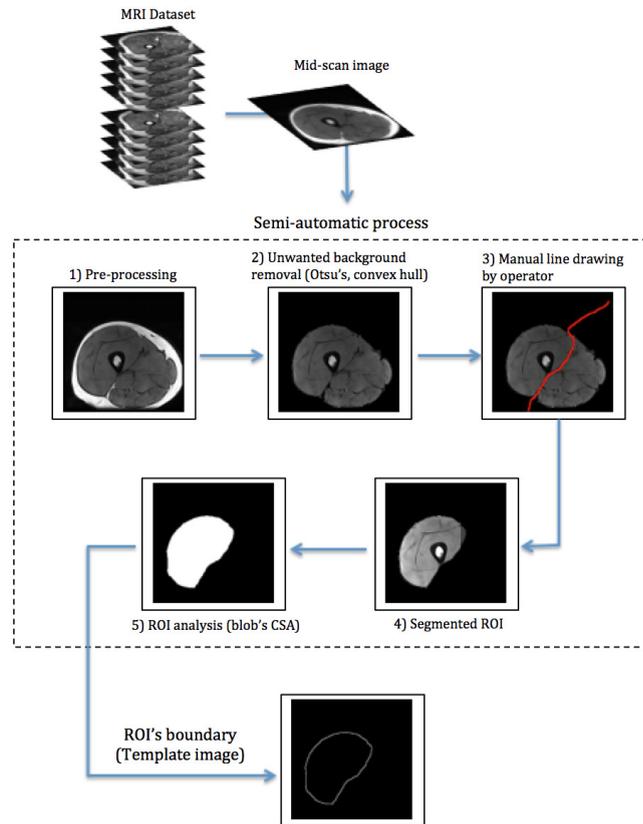


FIGURE 6.3: Summary of semi-automatic segmentation process.

The weight of each pixel point can be measured by the above function, with regards to the value of its eight nearest points and the value of sigma (standard deviation).

The curve transformation on the targeted MRI scan is done using the *Kass et al.* snake algorithm [91]. This edge-based snake model represents a controlled continuity spline that can deform to match any shape, under the influence of image forces and external constraint forces. The internal spline forces add a piecewise smoothness constraint. The image features attract the snake to the salient image features, such as lines, edges and terminations. For details formulation of the total energy of the snake, please refer sub-Subsection of 3.2.2.3 - Edge-based active contour.

Polygonal masking is later applied to extract the segmented ROI, as part of the early post-processing technique, followed by region properties for cross sectional area (CSA) quantification. The parametrical boundary information of the segmented ROI output image is eventually registered as a template and acts as a “shape memory” to segment the serial scans in the dataset. This is a continuous procedure until all MRI scans in the dataset have been processed and analysed.

Figure 6.4 outlines the diagram of the segmentation method. The new combination framework demonstrates significant improvements over the previous work and will be discussed in next section.

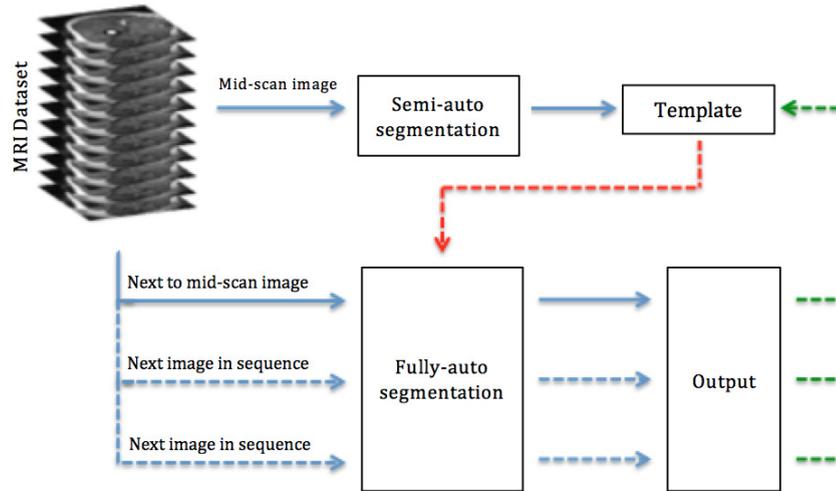


FIGURE 6.4: The operational flow of proposed segmentation method.

6.3 Experiments and Results

Seventeen MRI datasets (15 datasets with 26 scans and 2 datasets with 13 scans) with a total of 416 images were processed and evaluated by the combination framework of the segmentation methods.

The main objective of the proposed method was to segment the ROI (quadriceps, femur cortex and bone marrow) and to quantify the CSA of the ROI. To measure the accuracy of the segmented ROI, *Jaccard Similarity Index (JSI)*, a measurement for assessing segmentation accuracy (as per equation 3.78) was employed, where S_{gt} represents the area of “ground truth” or “gold standard” of the manual segmentation of ROI by using OsiriX and S_{sa} represents segmented ROI area by the proposed method. The similarity of the output results (of manual segmentation and proposed method) for a corresponding MRI scan is computed by the intersection and union of S_{gt} and S_{sa} in the given *JSI* function.

Figures 6.5 and 6.6 show examples of the ground truth manual ROI segmentation results (green border) by OsiriX and; segmentation results by the proposed method (red border), respectively, for one dataset (matching dataset as in Figure 3.21). The first four rows

depict segmentation within OsiriX and MATLAB environment, respectively followed by the next four rows, which are ROI extractions of corresponding first four rows.

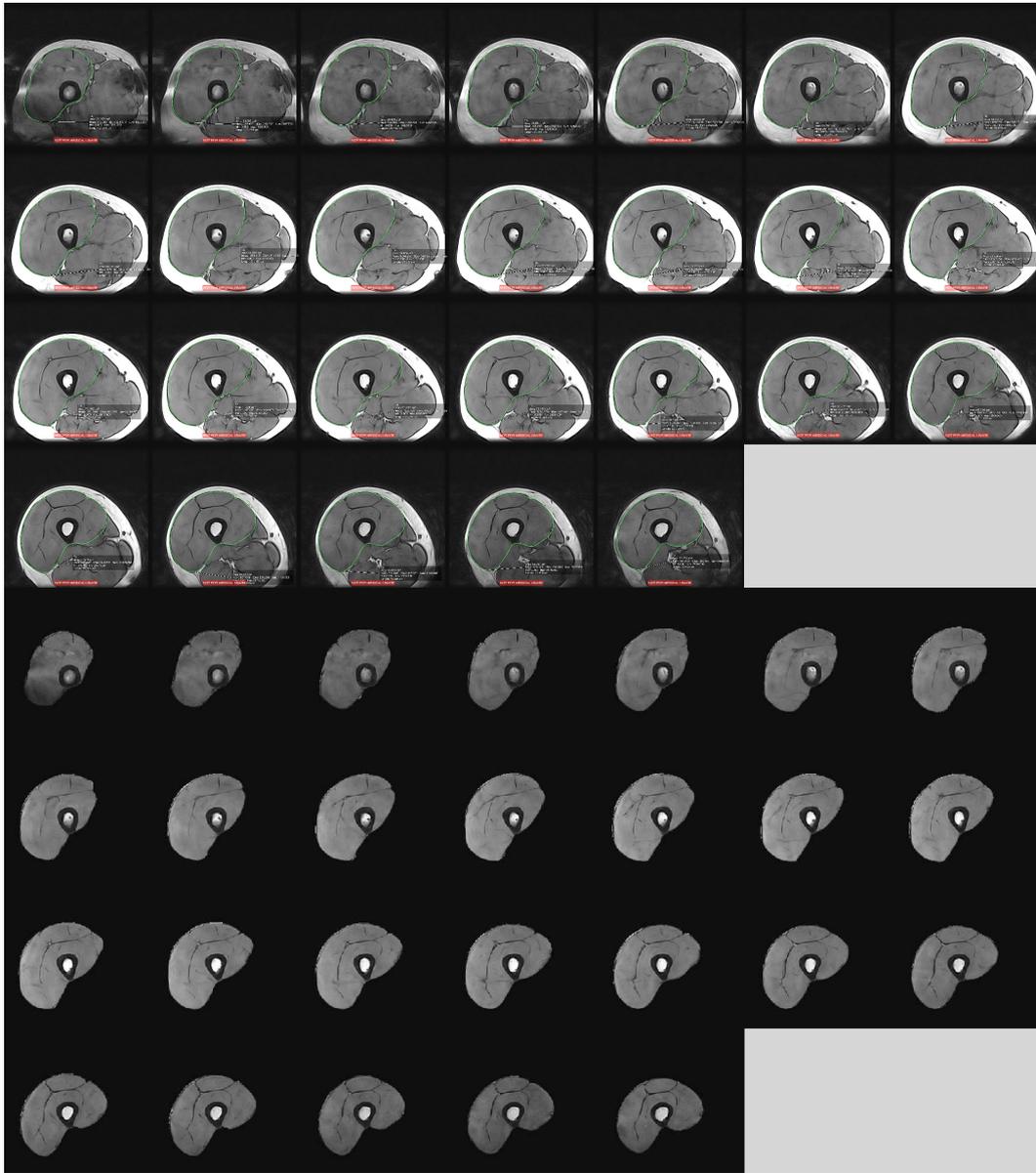


FIGURE 6.5: Manual segmentation results by OsiriX.

Overall, the mean JSI is 0.9334 (93.34% accuracy) for the 416 processed images (from 17 datasets), with a standard deviation (SD) of 0.0352. Assuming a normal distribution, this indicates that 68.3% (Mean \pm 1SD) of the segmentation results were populated between the accuracy of 89.82% and 96.86%.

The average manual annotation time for one MRI slice using OsiriX was 132 ± 42 seconds. In our previous work (Chapter 4), we were able to reduce the processing time to an average of 17 seconds per image when using a semi-automatic analysis [209]. With this

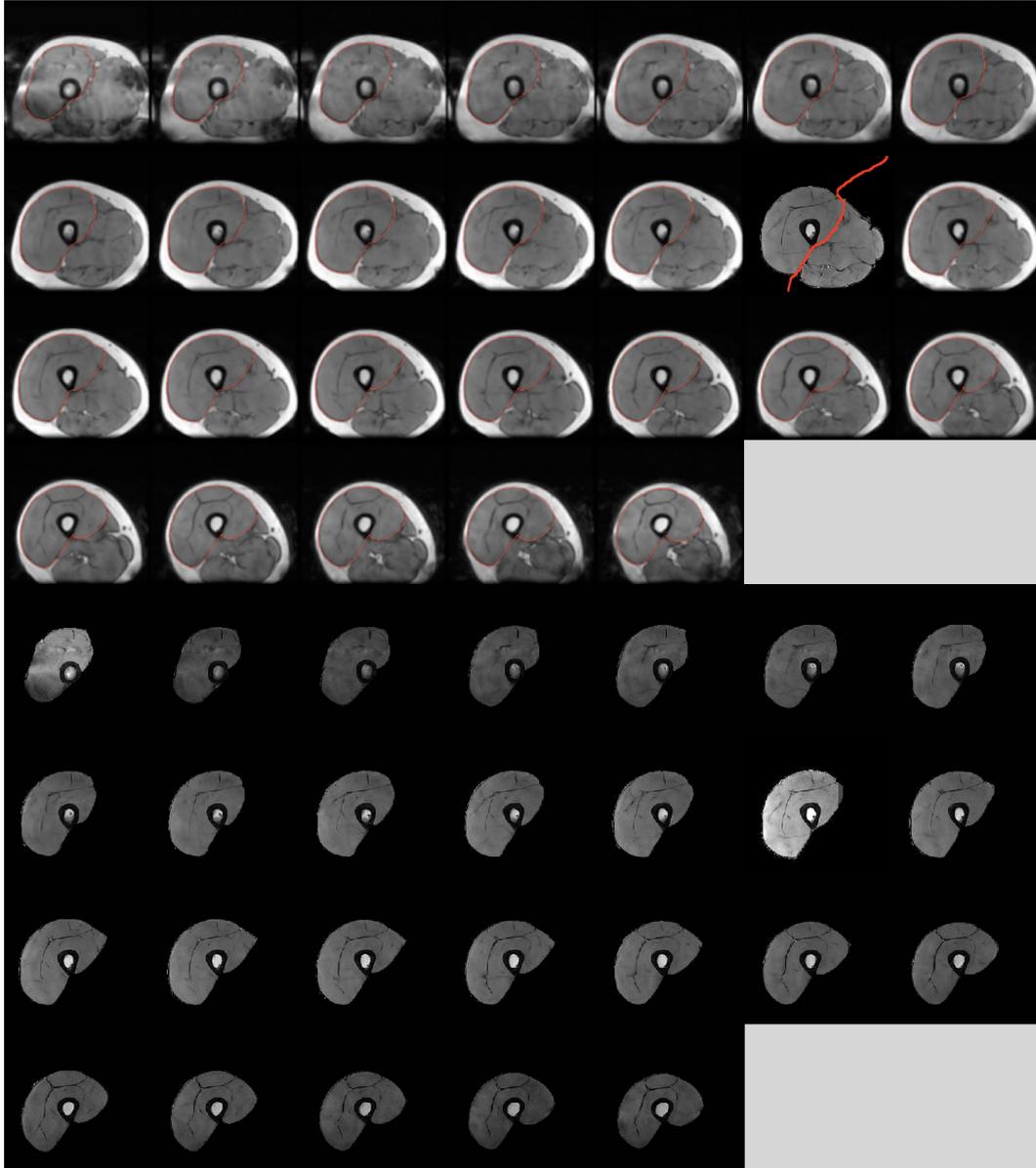


FIGURE 6.6: Segmentation results by proposed method.

method, the average annotation time has been further improved to just 3 seconds per image, a reduction in annotation time of 82.35% compared to the average processing time in [209], and 97.73% compared to average manual annotation in OsiriX. Tables 6.1 and 6.2 provide a summary of results for all assessed MRI images.

ROI was labelled as *Total Area (TA)* and consists of 3 major components: quadriceps muscles, femur cortical layer and bone marrow. Algorithms to describe region properties for ROI analysis were deployed for quantitative evaluation, for each of these components in the segmented images. The CSA of black pixels that represents femur cortex was

TABLE 6.1: Average segmentation accuracy for 416 processed MRI images

ROI	Mean <i>JSI</i> (%)	Standard Deviation (SD)
<i>TA</i> ¹	0.9334 (93.34%)	0.0352

TABLE 6.2: Average annotation time per image (in seconds)

Manual	Semi-Auto [209]	Proposed method [220]
132	17	3

obtained by subtracting the *TA* blob and the accumulation CSA of quads (in grey pixels) and bone marrow (in near white pixels).

6.4 Discussion

A new automatic segmentation method to process and analyse multiple transverse-plane images in one MRI dataset has been presented in this chapter. The proposed combination framework provides an average segmentation accuracy of 93.34%, which is similar to the level of agreement between two different manual assessors and between two different days when the same manual assessor examines the same images [209]. Moreover, the analysis of MRI muscle scans with the new method took just 2.27% of total processing time needed with conventional manual segmentation [209].

A built-in correctional method of intensity inhomogeneity or bias field, based on the partition matrix of fuzzy c-means [229], is only an optional procedure in the method's pipeline, as bias field rarely affects mid-scan slices from any single MRI dataset, opposite to the earlier and later 2 to 3 scans in a dataset, which are usually affected by bias field. Insignificant accuracy differences to the segmentation results also indicate that the active contour curve transformation performance is unaffected by the bias field, even when bias field correction was applied to the subsequent images in one MRI dataset. However, when needed, clustering functions will assemble the image data into object classes, estimate the margins of the classes and adjust the varying illumination artefacts. And with added c-programming based MATLAB execution code by *Ahmed et al* [229], this correctional method provides a quick solution for bias field correction, thus not affecting the overall processing time.

¹ *Total Area* - Consists of quadriceps, femur cortex and bone marrow.

To establish the curve internal deformable energy (equation 3.37), the spline energy of the first order term controlled by the stretching force (α - elasticity) was used alongside the second order term controlled by its bending force (β - rigidity). The weights of α and β in the snake model in this scenario have to be adjusted according to the global image gradient for all datasets in the database to ensure spline continuity and smoothness in membrane transformations. To achieve consistent and exceptional results for automatic segmentation, all mid-scan images in the 17 processed datasets have been pre-segmented by the semi-automatic segmentation method as described in Chapter 4. The experiments suggested that the efficiency (segmentation accuracy) of the automatic segmentation method is dependent on the image quality of the mid-scan image being segmented. This is true since the curve transformations are dependent on the initial geometric information (of the template image from semi-automatic segmentation product) as the prior shape regulator.

A smoothing procedure to the femur region by fuzzy c-means (FCM) clustering [20] that was adapted on Chapter 5's framework is also included in the overall automation process. It is performed by tuning down the intensity of the extracted femur and the result of intensity adjustment to the femur is then overlaid with the original image, before the application of automatic segmentation. The procedure is necessary so that over- or under-segmentation outcome can be avoided, mainly due to the converging factor to the least total energy in an object/image that naturally associated with snake or active contour E_{snake}^* algorithm.

Once the segmentation is done, the framework provides the CSA of TA and could therefore benefit scientists who require a quick estimate in the size of the thigh in studies of muscle wasting and disease [230]. Measuring muscle size will have relevance for conditions such as metabolic syndrome (diabetes, high blood pressure, obesity) and/or analysing physiological relationships of the subject to muscle disorders (muscle atrophy and muscle wasting).

6.5 Conclusion

A reliable and accurate combination framework that leads to an automatic dataset segmentation method has been developed, with added benefits of elements (quadriceps, femur cortex, bone marrow) extraction and ROI image analysis (parametric cross-sectional area). The method requires only one simple manual interaction to segment the mid-scan MR image and automatically segments serial images in the same MRI dataset by the customized framework of snake active contour curve transformation, based on template geometric information as the shape regulator.

The method needs minimal supervision and offers exceptional performance in terms of segmentation accuracy and processing time. However, the overall development of the method is still at its infancy and there is a need to develop a more robust segmentation system.

In the future, our main aim is to replace the semi-automatic segmentation method in this combination framework to an automatic mid-scan MR image segmentation method. Based on early investigations, combination frameworks of statistical data training for shape analysis algorithms (such as prior shape/positional information or deformable part model), partial differential based algorithms, fast marching method, image registration based technique (atlas-registration) and artificial neural network (ANN) application (through the emerging of convolutional neural network (CNN) and its many variational models) appear to be five major techniques with promising solution to the ROI automatic segmentation of the mid-scan MR image. Further developments in conjunction to the above methods will be adapted and explored in future work.

Chapter 7

Semantic Segmentation of MR Images of Human Thigh Muscle

This chapter presents the application of deep learning techniques for semantic segmentation of ROI of MRI human thigh muscle.

7.1 Introduction

In digital image processing and computer vision analysis, semantic segmentation is an attempt to partition meaningful parts that belong to the same object class of an image together [231] and is mostly applied in object detections [232–234]. Classifying images based on pixel-wise (or pixel) classification, leads to the same outcome but is regarded in a slightly different nuances in the terminology [235, 236]. Both semantic segmentation and pixel-wise classification algorithms are principally trying to achieve the same goal of analysing the role of each pixel in the image. In relation to the deep convolutional neural networks, a collection of images and its corresponding pixel-labelled images are needed to train a semantic segmentation model, where every pixel value in a pixel-labelled image represents the class label of that pixel. For example, in a grayscale magnetic resonance (MR) image in its simplest form (with black-coloured pixels represent the background, gray-coloured pixels correspond to soft muscle tissue and white-coloured pixels to bone marrow or adipose tissues), this could be represented by three class labels. RGB images

in contrast, could have more class labels (depending on labelled classes), for object classification and network training purposes.

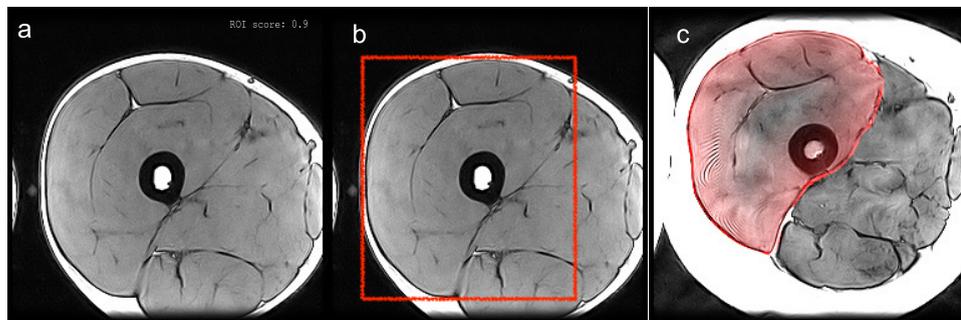


FIGURE 7.1: The output domain of a convolutional neural network (CNN) in MRI of the thigh. (a) Object classification - the task of identifying that there is a quadriceps in the image. (b) Object localization - class label including a bounding box to show the location of the object (ROI). (c) Object segmentation - class labelling and outlining the predicted region. Note: other major domain in CNN includes object detection - localization of multiple of objects (objects can be from different classes) in an image.

Deep learning is a set of algorithms within an artificial neural network (ANN) structure, inspired by the neurological system and a subfield or branch of machine learning. Over the past few years, deep learning has been established as one of the emerging paradigms in the evolution of Artificial Intelligence (AI) [237], Big Data [238] and other analytic architectures [239, 240]. Deep learning is used by major global brands such as Google, Amazon, Facebook and Nvidia in many of their applications such as voice, image and pattern recognition algorithms and engaged in self-making or logical decisions based on data or input from the user [241–244]. Its approach to AI is showing improvements in the autonomous and unsupervised learning systems, which in turn, revolutionize many industries and bring significant shifts in society through developments in healthcare analysis and treatment [245–247]. The application of deep learning is extended to skin analytics for the detection of abnormal skin lesions such as skin cancer and diabetic foot ulcer [248–250].

Recently, the deep learning has provided various end-to-end solutions in the semantic segmentation of abnormalities such as breast cancer, brain tumour, skin lesions, foot ulcers etc. in medical imaging [85, 248, 250]. As mentioned in Chapter 2, there are several studies in semantic segmentation using deep learning on MRI of the brain and heart, but only one attempt has been made on MRI thigh segmentation [12]. The model integrates a probabilistic neural network to facilitate the generation of probability estimates at each pixel for use in an iterative segmentation process. This preliminary

study on MRI thigh segmentation using one of the simplest forms of neural network reported that the segmentation result is unaffected by the quality of the training dataset. This study only analyzed a single MR image of the thigh and suffered from tissue misclassification, thus raising questions of the validity of their application, procedure and result. At the time of writing this thesis, no other study has been made in segmentation of MR images of human thigh muscles by using the deep learning approach, neither in the form of conventional ANNs, nor popular and emerging ANNs such as convolutional neural network (CNN). Current state-of-the-art methods described in Chapter 2 and 3, especially image processing based techniques, are not robust, due to their nature of reliance on specific regulators and rules, with certain assumptions. In contrast to ANNs, deep learning methods do not require such strong assumptions and have demonstrated superiority in region and object segmentation and classification, which suggests that the robust fully automated MRI segmentation of human thigh muscles may be achieved, by adapting such approach. Moreover, the segmentation output performance from most recent applications of network models in medical imaging [86–89], in association with image processing technique [90] and in conjunction with the utilization of better and upgradable graphics processing unit (GPU), are remarkably accurate and fast.

In this chapter, five pre-trained models from three different existing deep learning architectures will be reviewed and applied to the thigh MR images for semantic segmentation tasks. The region of interest (ROI) (the same as in previous chapters), consists of quadriceps muscles, bone and marrow. The main objective of the proposed work in this chapter is to discover whether CNN could: 1) be a reliable end-to-end solution for (semantically) segmenting the ROI; and 2) achieve a significant performance improvement compared to the previous benchmark set in Chapter 4.

7.2 Deep Learning Architectures

Semantic segmentation architecture is commonly considered as a pre-trained classification network (or in general terms an encoder network). An encoder network is architecturally different to a decoder network, where the decoder's task is to semantically assign the discriminative activation maps (or features) in lower resolution. These feature maps are learned by the encoder onto the pixel-wise space in higher resolution to get a dense classification. In image classification, identifying the object of interest in the image is

the main objective of the network and this is different to semantic segmentation, where the segmentation process involves class labelling and delineating the boundaries of the object of interest. Those differences are due to the different mechanisms employed in encoder and decoder networks like skipping fully connected layers, dropout layers, etc. This section describes the foundation architecture of every deep convolutional neural network (CNN) used in this work, which consist of *AlexNet*, *VGG-16* and *PSPNet*. There are several other networks available, however during the preliminary tests, these networks are the most applicable for the configuration of MR images used.

7.2.1 AlexNet

AlexNet [1] by *Krizhevsky et al.* is arguably the one that was revitalizing ANN, revolutionizing the deep CNN and is considered as one of the most notable deep learning architecture in the field. In 2012, it won the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) with a top 5 test error rate of 15.4% and test accuracy of 84.6%, while the next best model achieved an error rate of 26.2%. AlexNet has a relatively simple layout which was made up of five convolutional layers, max-pooling layers, non-linear activation function of Rectified Linear Units (ReLUs), dropout layers and three fully-connected layers. This network is designed for classification with 1000 potential categories and is trained by using batch stochastic gradient descent, on *ImageNet* data, which are comprised of over 15 million annotated images from a total of over 1000 categories. Figure 7.2 demonstrates the AlexNet architecture.

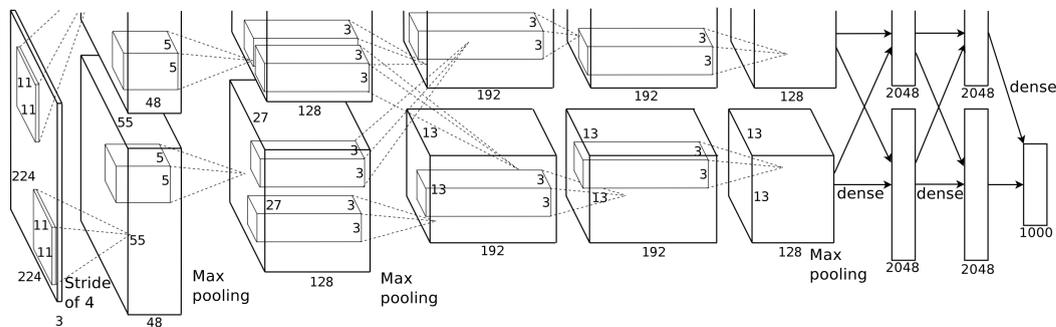


FIGURE 7.2: The network architecture of AlexNet [1].

7.2.2 VGG-16

In 2014, K. Simonyan and A. Zisserman of the Visual Geometry Group (VGG) from the University of Oxford proposed a CNN model known as *VGG-16* (also referred to as *VGGNet*) [251] and secured the first and second place in localization and classification, respectively in the ImageNet Challenge 2014, with a test error rate of 7.3% and test accuracy of 92.7% in ILSVRC 2014 submission. This 16-layer (and up to 19-layer) deep CNN utilizes small-sized 3x3 feature detectors and 2x2 max-pooling together with small stride of 1 and 2, respectively, unlike AlexNet's 11x11 kernel size. With an effective receptive field of 5x5, such hierarchical representations resulted in fewer parameters and more non-linearities of ReLU. This is true since smaller-sized kernels doubles the convolutional layers and in turn doubles the ReLU layers and the depth of the volumes, as the computations goes deeper down the network. The VGG-16 architecture is trained with mini-batch gradient descent with momentum and suitable for image recognition, localisation and classification tasks. Figure 7.3 illustrates the architecture of VGG-16.

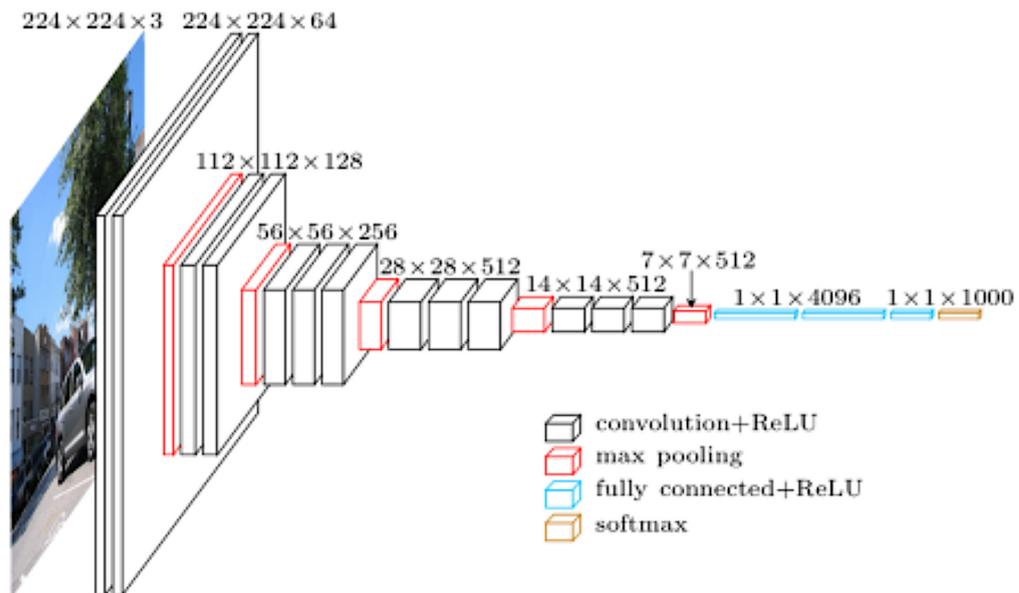


FIGURE 7.3: The network architecture of VGG-16 [2].

7.2.3 PSPNet

In 2016, Zhou *et al.* proposed a network architecture called Pyramid Scene Parsing Network (*PSPNet*) [3]. In this work, the research team from Chinese University of Hong

Kong and SenseTime Group Limited used the global context information proficiency by aggregation of a different-region-based context. The global spatial context is important since it provides suggestions on the distribution of the segmentation classes. This procedure is done by the combination of a pyramid pooling module (applying large kernel pooling layers) and the proposed PSPNet, and proved to have an effective global prior representation capable to produce good results in the scene parsing tasks and pixel-level predictions. Other key features of this architecture include: integrating dilated convolutions as a means of modifying the base architecture of Residual Network *ResNet* [252], a 152-layer network architecture proposed by Microsoft Research Asia in late 2015. The proposed method achieved the first place in ImageNet scene parsing challenge 2016, PASCAL VOC 2012 benchmark and Cityscapes benchmark. The paper also reported that PSPNet model yields accuracy of 85.4% (by mean intersection over union (mIoU)) on PASCAL VOC 2012 and accuracy of 80.2% on Cityscapes. Figure 7.4 shows the PSPNet architecture.

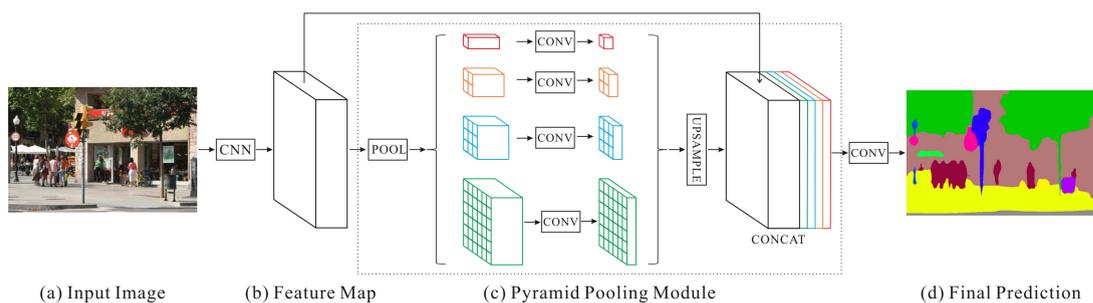


FIGURE 7.4: The network architecture of PSPNet [3].

Spatial pyramid pooling module is added after the encoder network mainly to concatenate the feature maps from modified ResNet with upsampled output of parallel pooling layers with ascending sizes of kernels (hence the name pyramid). An auxiliary loss is introduced as an input to pyramid pooling module to optimize overall learning process.

PSPNet is proposed by the motivation of tackling issues related to mismatched relationship, confusion categories and inconspicuous classes. The pyramid pooling module combines features based on four different pyramid scales. The first pyramid level in the red coloured box (as in Figure 7.4) is a global pooling layer that produces a single bin output. The feature map is separated into different sub-regions and different locations of pool representation are formed in the next pyramid level. Each pyramid pooling module holds different sizes of feature map and a 1×1 convolutional layer is applied after each

level to preserve the weight of the global feature and this results in a reduced depth of context representation to $1/N$ of the original input, where N is equal to the level of pyramid.

To get the identical original size of feature map, the low-dimension feature maps are directly upsampled via bilinear interpolation. The final pyramid pooling layer depicts global features that are concatenated from different levels of features. To train the final classifier, the main branch used softmax loss and another classifier (the res4b22 residue block)[3] is employed after the fourth stage. To assist in optimizing the learning process, the auxiliary loss (with additional weight to balance it) is used and this yields the master branch loss to take the most responsibility. For testing, the auxiliary branch is abandoned and instead only the optimized master branch is used for the final model prediction.

7.3 Experiments

7.3.1 Datasets

MRI scans of the thigh were collected from men and women aged 18-90 years [220]. All scans were collected using the same T1-weighted Turbo 3-D sequences using a 0.25-Tesla MRI scanner (Esaote G-scan; Italy). All MRI scans consisted of serial transverse-plane slices, each with 6.3-mm thickness and 0-mm inter-slice gap and image matrix of 256 x 256 (in pixels).

The datasets are separated into two groups, with the first group, being the collection of MRI mid-scans (image no.7 for datasets with 13 scans or image no.13 for datasets with 26 scans)¹, which contains 110 images. The second group consists of 1000 MR images and is a combination of serial MR images of the whole-scan dataset (all 13 or 26 scans per subject/dataset) as per Figure 3.20 and 3.21.

To train, validate and test the MR images on the customized deep CNNs, all of the original 1110 grayscale MR images are converted to an RGB format, with corresponding images manually annotated² according to the ROI. These extracted ground truth binary

¹For DICOM images MRI dataset samples, please refer [here](#).

ROI images are then converted into the Pascal VOC images (labelled as 8-bit single-channel paletted images, with 3-channel colour profile of sRGB IEC61966-2.1).

In this Pascal VOC format, 0 maps are indexed to black pixels corresponding to the background and 1 (in red) denotes the ROI (consists of quadriceps, femur and marrow). This procedure is simplified in Figure 7.5 below.

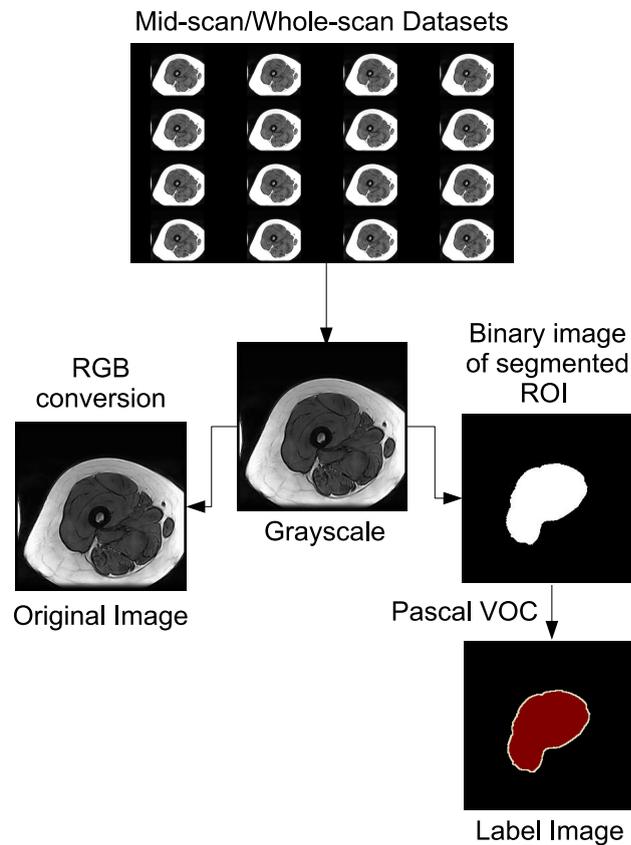


FIGURE 7.5: Preparation of MRI Datasets.

7.3.2 Transfer Learning

Having a sufficient amount of data is crucial to effectively train a deep neural network. However, in reality it is not often possible and this is why the idea of transfer learning has emerged in recent years as a new learning framework [253, 254]. Instead of starting from the beginning or rather generating random weight initializations from the beginning, transfer learning is the method that uses the weights from a pre-trained model and fine-tuning the weights with the original dataset collections of interest. This helps: 1) As we go deeper down the network of a pre-trained model that has been trained on a large

²For DICOM images manual annotation process, please refer [here](#).

dataset such as on ImageNet [255], it will detect features like edges and curves and;

2) Less computation time, especially for network training. Depending on the domain of the problem (classification, detection, recognition, etc.) and the type of data to be processed, this pre-trained model can be attuned as a feature detector. By removing the final layers of this pre-trained model and replacing it with a specific classifier, one can use the weights from the pre-trained model and emphasise more on other important layers for training purposes.

In this work, the values of the weights are initialized from a pre-trained transfer learning procedure. First, each model was trained on an ImageNet dataset, where the final layers of its fully connected layer are removed and the pre-trained weights from earlier convolutional and pooling layers are then transferred to the next model. The weights associated to these pre-trained models are subsequently used to fine-tune the customized deep CNNs for training, validation and testing the thigh MRI datasets. This transfer learning process is demonstrated in Figure 7.6.

7.3.2.1 Customized Deep CNNs (Pre-trained Models)

The customizations of deep CNNs used in this chapter are all inspired by the work of *Long et al.* [256]. Typical CNNs are comprised of basic components such as convolution, activation function (ReLU) and pooling layers. These layers operate according to local input regions and depend on relative spatial coordinates, which satisfy the following properties:

$$y_{ij} = f_{ks}(\{x_{si+\delta i, sj+\delta j}\}_{0 \leq \delta i, \delta j \leq k}), \quad (7.1)$$

where x_{ij} and y_{ij} are the corresponding layer's input and output (at location (i, j)), respectively, k is the kernel size, s is the stride and f_{ks} determines the layer type, which is either a convolution, a pooling, or a ReLU activation function, and so on for other types of layers in the network. This function operation is preserved with the composition of kernel size and stride that conform to the transformation function of:

$$f_{ks} \circ g_{k's'} = (f \circ g)_{k'l+(k-1)s', s's'} \cdot \quad (7.2)$$

Generally, traditional deep CNNs with fully connected layers are only capable of processing certain input sizes (standard version of AlexNet for example, input image with

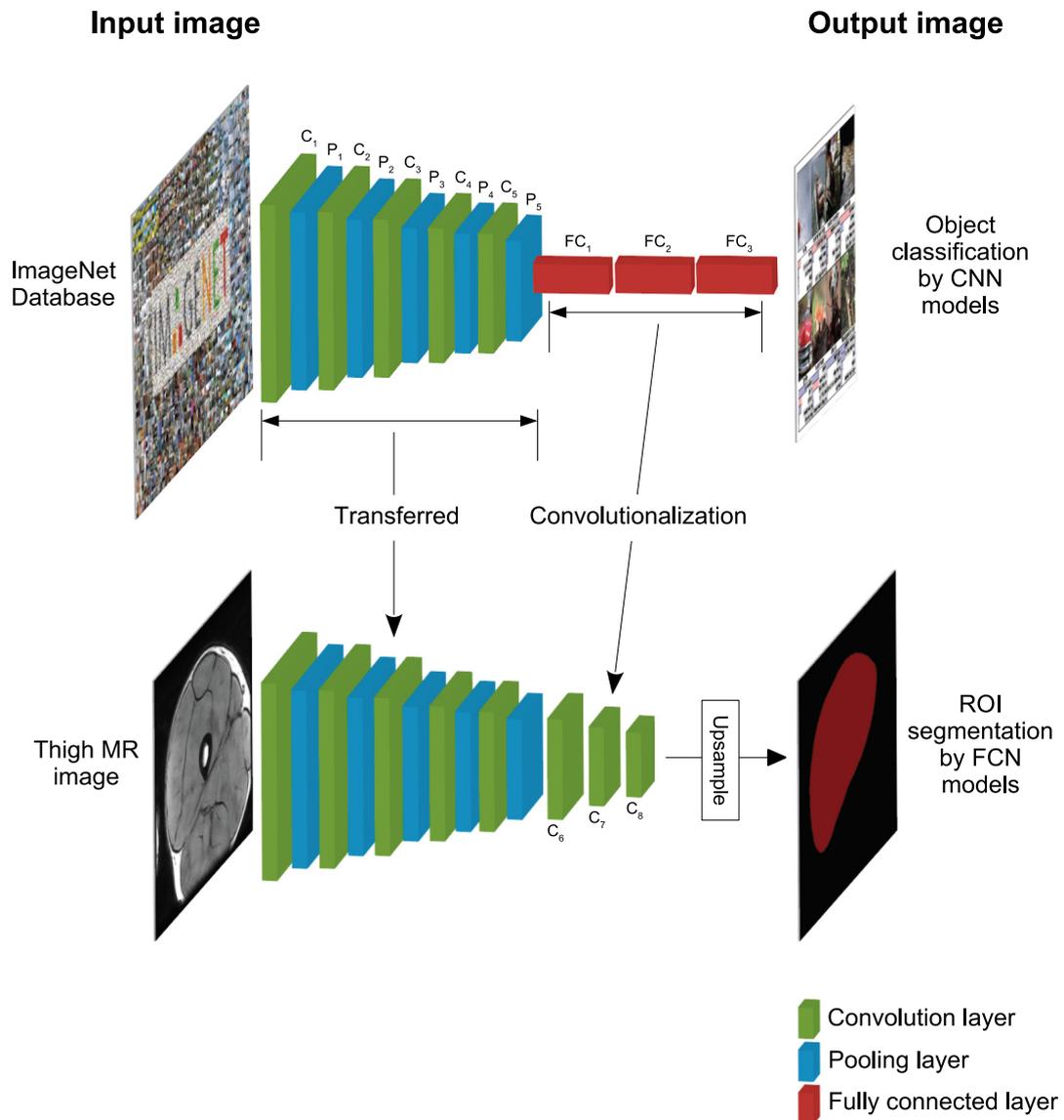


FIGURE 7.6: Transfer learning procedure of deep CNNs to obtain optimized weights initializations. Three fully connected layers of CNN were removed and replaced by three convolutional layers, making the pre-trained model fully convolutional.

a fixed size of 224×224 will be processed regardless of the original size of the input image) and compute a general nonlinear “function”. However, fully convolutional networks (FCNs) compute a nonlinear “filter” and are therefore able to work on any size of input images with equal size of output images.

FCN-AlexNet. *FCN-AlexNet* is a customized FCN model based on AlexNet architecture. Since AlexNet specializes in object classifications (due to being trained on ImageNet dataset - with 15 million of images and with 1000 of different object classes),

a few adjustments are made at its fully-connected classification layers. Earlier network's convolutional, ReLU and pooling layers are maintained (for lower- and high-level features extraction), while the fully connected layers are removed and replaced with equivalent convolutional layers.

Due to the nature of the fully connected layer that normalises the vectors of output from previous layers, these often resulted in extracted features with relative coordinates and thus, produced non-spatial predictions for the final output of the domain task. The sizes of the kernels are adjusted according to the input size of these additional convolutional layers. Deconvolutional layers with similar stride sizes and scaling factors to the previous convolutional layers are then introduced to generate a pixel-wise prediction (based on object classes) for every pixel of the input image.

FCN-VGGNet. VGG-16 architecture is customized to three models, *FCN-32s*, *FCN-16s* and *FCN-8s*, where each model magnifies the output with different upsampling layers. For FCN-32s, the fully connected layers are convolutionized and a 32-pixel stride size is applied in deconvolution layers. This works well for object classification, localization and detection tasks as precise pixel-wise predictions are not the main priority. FCN-16s and FCN-8s perform better for object segmentation tasks, as both extract extra low-level features from the input image to produce a more precise output result by 16 x 16 and 8 x 8 pixel blocks, respectively.

In FCN-16s, the final output is a product of upsampling of two layers, which are upsamples of the fourth pooling layer and (upsampling of the seventh convolutional layer) x 2. For FCN-8s, the final output of the model is a product of upsampling of the third pooling layer, (upsampling of the fourth pooling layer) x 2 and (upsampling of the seventh convolutional layer) x 4.

PSPNet. In this chapter, the implementation of PSPNet is a straightforward process, without any customization to its “prototxt” or model layers.

7.3.3 Configurations

Machine 1. (1) Hardware: CPU - Intel i7-6700 @ 4.00Ghz, GPU - NVIDIA 1080 TITAN X 11Gb, RAM - 32Gb DDR5 (2) Framework: Caffe [193] (3) Training and

validation: All pre-trained models are applied to mid-scan MRI dataset (4) Training epochs: 30 (5) Solver type - Stochastic Gradient Descent (SGD) and Adam. (6) Base learning rate: 0.0001.

Machine 2. (1) Hardware: CPU - Intel i7-6700 @ 4.00Ghz, GPU - 2 x NVIDIA GeForce GTX 980 Ti 6Gb, RAM - 32Gb DDR5 (2) Framework: Pytorch [257] (3) Training and validation: All pre-trained models are applied to whole-scan MRI dataset (4) Training epochs: 50 (5) Solver type - Stochastic Gradient Descent (SGD). (6) Base learning rate: 0.0001.

7.4 Results

Two datasets of 1110 MR images were manually segmented and labelled. To implement the semantic segmentation domain, these segmented MR images are then trained, validated and tested by the pre-trained models of FCN-AlexNet, FCN-VGG-16 (consists of FCN-32s, FCN-16s and FCN-8s) and PSPNet. Both mid-scan and whole-scan MRI datasets were split into the configuration of 70%, 10% and 20% for training, validation and testing, respectively (mid-scan MRI dataset: 77, 11, 22 images and whole-scan MRI dataset: 700, 100, 200 images), with 5-fold cross-validation procedure (5 processing batches to analyse all MR images).

To evaluate the performance of these FCN models, all of the semantically segmented (tested) images are compared and measured with the corresponding labelled (manually segmented ROI) images. In addition to *Jaccard Similarity Index (JSI)* used in previous chapters, *Sensitivity*, *Specificity*, *Matthews Correlation Coefficient (MCC)* and *Dice Similarity Coefficient (DSC)* indexes are also employed as the performance metrics for the evaluation of segmented ROI. The formulations of each metric are described as follows:

$$Sensitivity = \frac{TP}{TP + FN} , \quad (7.3)$$

where $TP = |MP \wedge GT|$ and $FN = |GT| - TP$.

$$Specificity = \frac{TN}{FP + TN} , \quad (7.4)$$

where $TN = \forall - |MP \vee GT|$ and $FP = |MP| - TP$. MP , GT , TP , TN , FP and FN , as they stand for are: Model Prediction image, Ground Truth image, True Positives, True Negatives, False Positives and False Negatives, respectively. *Matthews Correlation Coefficient (MCC)* [258] is given as:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} , \quad (7.5)$$

and *Dice Similarity Coefficient (DSC)* (also shown in Section 3.3.):

$$DSC = \frac{2 \times TP}{(2 \times TP + FP + FN)} . \quad (7.6)$$

Sensitivity (also called the *true positive rate* or the *recall*) measures the proportion of positives that are correctly identified or predicted as an ROI. It measures the proportion of excellent ROIs who were correctly identified to the total number of excellent ROIs. *Specificity* (also called the *true negative rate*) measures the proportion of negatives that are correctly identified as such, measures the proportion of poor ROIs who were correctly rejected to the total number of poor ROIs. Both *Sensitivity* and *Specificity* are most commonly associated with a binary classification test and statistically determine the performance of the test. A good binary classification test always results with high metric values (or percentage). Generally, if *Sensitivity* is high and *Specificity* is low, there is no need to concern about the excellent candidates (of a dataset) but the poor candidates must be re-assessed to eliminate false positives (poor candidates that mistakenly selected). However, excellent candidates must be re-assessed to eliminate false negatives (excellent candidates that mistakenly rejected), when *Sensitivity* value is low and *Specificity* is high. An average binary classification test always results with average values which are almost similar for both *Sensitivity* and *Specificity*. *Matthews Correlation Coefficient (MCC)* often used in machine learning as a means of the quality of binary classifications (of two classes), with formulation that based on values of true and false positives and negatives. *MCC* is considered as a balanced evaluation, which can be used even if the classes are different in size [259] and regarded as a correlation coefficient between the observed and predicted binary classifications. It returns a value between -1 and +1, where coefficient of +1 represents a perfect prediction and -1 indicates total disagreement between prediction and observation. Table 7.1 summarizes the mean index of semantic segmentation results of all the FCNs for the testing of 1110 MR images.

TABLE 7.1: Mean performance index for all models on mid-scan (MD) and whole-scan (WD) datasets.

Network	JSI		DSC		Sensitivity		Specificity		MCC	
	MD	WD	MD	WD	MD	WD	MD	WD	MD	WD
PSPNet	0.8263	0.9194	0.9026	0.9565	0.9029	0.9583	0.9252	0.9611	0.828	0.9209
FCN-AlexNet	0.8487	0.8923	0.9169	0.9332	0.901	0.9446	0.9804	0.9769	0.8925	0.9153
FCN-32s	0.8198	0.9211	0.899	0.9581	0.8796	0.9768	0.9776	0.9467	0.8712	0.9232
FCN-16s	0.8669	0.9199	0.9266	0.9574	0.9399	0.9782	0.9743	0.9442	0.9059	0.922
FCN-8s	0.8973	0.9316	0.9442	0.9639	0.9708	0.9822	0.9748	0.9514	0.9282	0.9336

Overall, FCN-8s performed best in regards to the semantic segmentation task of the ROI of thigh MR images, with an average accuracy of 0.8973 (89.73%) for mid-scan dataset (MD) and 0.9316 (93.16%) for whole-scan dataset (WD) by *JSI*. The average *DSC* performance index for the corresponding datasets was also notably higher, i.e. 0.9442 (94.42%) and 0.9639 (96.39%), for MD and WD, respectively. For the uniformity of the thesis as a whole and the less-forgiving nature of its measurement, *JSI* will be used comprehensively from here after. *DSC* was included partially because it was universally employed as a means of performance evaluation in machine learning. Although both *JSI* and *DSC* adapted a slightly different formulation from one another, both indexes are good for performance assessment purpose.

FCN-8s also excelled at *Sensitivity* and *MCC* indexes and ranked third overall in the *Specificity* performance division. The difference however, was not substantial from the first-ranked FCN-AlexNet model, in this regard (0.0056 and 0.0255 of *Specificity* differences for MD and WD, respectively).

7.4.1 FCN-8s model with different solver types

Stochastic Gradient Descent (SGD) [190] is arguably predominantly used in the methodology of training a deep learning model, mainly due to its strength of simplicity in implementation and fast processing, even for problems (or datasets) with many training patterns. Semantic segmentation results in Table 7.1 also generated by the application of SGD as its solver type in all pre-trained models.

Table 7.2 demonstrates the performance impact on the application of different solver methods used for semantic segmentation in thigh MR datasets by FCN-8s model.

From Table 7.2, the various applications of solver types on FCN-8s during the training and validation suggested that Adam solver generally performs better for the semantic

TABLE 7.2: Mean *JSI* performance index for different solver types used in FCN-8s on mid-scan (MD) and whole-scan (WD) datasets.

	MD					WD	
	AdaGrad [192]	Adam [191]	NAG [195]	RMSprop [196]	SGD	Adam	SGD
JSI	0.8778	0.927	0.8972	0.8829	0.8973	0.9469	0.9316

segmentation of thigh MRI, with both MD and WD datasets having a *JSI* performance increment of 0.0297 (2.97%) and 0.0153 (1.53%) in accuracy, respectively compared to FCN-8s with SGD solver. Unfortunately, none of the results can be generated here with the application of AdaDelta solver [194] on FCN-8s due to the lower percentage ($\approx 70\%$) of learning accuracy and higher loss during the training and validation, even after the number of epoch was increased.

7.4.2 Segmentation of MRI Whole-scan Dataset Trained Models on MRI Mid-scan Dataset

The main intention of distributing the dataset into two (MD and WD) is to observe for any significance in *JSI* performance index for the semantic segmentation task of the ROI of the thigh MRI. Should the ROI segmentation of a whole-scan dataset succeed with greater level of accuracy compared to the results in Chapter 6, then adapting the FCN-8s model (as it proves to be the best model to this domain thus far) on whole-scan dataset can be further explored, replacing the previous proposed framework, as in Chapter 6, for a fully-automated ROI segmentation model.

However, should better semantic segmentation results of the mid-scan dataset be obtained compared to results in Chapter 4, then the application of FCN-8s model on mid-scan dataset will be used, replacing the benchmark of semi-automatic frameworks in Chapter 4. This in turn can be employed to the framework as proposed in Chapter 6 for further investigations to enhance the overall segmentation framework.

To further observe the behaviour of the pre-trained models, we combined the MD and WD datasets, and trained this newly refreshed “all-scan” dataset (AD) with all the pre-trained models proposed. For this purpose, the new training, validation and testing configurations are segregated as: 900, 100 and 110 images, respectively, with 1000 images in WD being used for training and validation and 110 images in MD primarily for testing. The hypothesis was that the value of weight initialization will be much more optimized

since there are more images involved for training and validation procedures, and therefore the testing performance of each pre-trained model should be higher than that of MD and WD datasets. Table 7.3 shows the results of this experiment, with results for MD and WD are also included, for comparisons.

TABLE 7.3: Mean *JSI* performance index of all models on mid-scan (MD), whole-scan (WD) and all-scan (AD) datasets.

		Model						
		FCN-AlexNet		FCN-8s		FCN-16s	FCN-32s	PSPNet
		Adam	SGD	Adam	SGD	SGD	SGD	SGD
Dataset	MD	0.897	0.8487	0.927	0.8973	0.8669	0.8198	0.8263
	WD	0.9253	0.8923	0.9469	0.9316	0.9199	0.9211	0.9194
	AD	0.9228	0.8988	0.9388	0.9391	0.9378	0.9412	0.9081

As expected, overall FCN-8s (with Adam solver) performed better compared to the other pre-trained models, throughout all the datasets (MD, WD or AD), with FCN-8s for WD dataset (94.69%) remains the best. However, surprisingly for AD, the incorporation of SGD solver on FCN-32s performed better (compared to the other models within AD), with average *JSI* performance of 0.9412 (94.12%), an increment of 0.201 (or 2.01%) by *JSI* compared to FCN-32s for WD. Also note that the mean *JSI* performances for FCN-8s (with both Adam and SGD solver methods), FCN-16s and FCN-32s, in AD dataset are also close to one another, with only a small performance difference between them (± 0.0034 or 0.34%). This suggests that an AD dataset trained, validated and tested on VGG-16 customized models (FCN-8s, -16s and -32s) has the potential to perform well in the domain of semantic segmentation and confirms the concept of having sufficient input (and labelled) ground truth data will improve the output domain of the deep learning model. Table 7.4 depicts the average processing (semantic segmentation testing) time (in second) for a single thigh MR image for all proposed models.

TABLE 7.4: Mean processing (testing) time (in second) for all models to semantically segment an ROI of thigh MR image.

		Model											
		FCN-AlexNet		FCN-8s				FCN-16s		FCN-32s		PSPNet	
		Adam	SGD	AdaGrad	Adam	NAG	RMSprop	SGD	Adam	SGD	Adam	SGD	SGD
Dataset	MD	0.34	0.24	0.46	0.64	0.45	0.37	0.61	0.58	0.39	0.61	0.55	0.05
	WD	0.18	0.15	-	0.1	-	-	0.03	-	0.03	-	0.03	0.05
	AD	0.1	0.09	-	0.17	-	-	0.03	-	0.04	-	0.03	0.05

FCN-8s with SGD solver (in AD and WD) was the best in terms of computation time for the semantic segmentation task of thigh MRI, with 30 ms (0.03 sec) average processing time per image. That was 100 times faster than the proposed method in Chapter 6

(3 sec), 567 times faster than the proposed method in Chapter 4 (17 sec) and 733 times faster the proposed method in Chapter 5 (22 sec). In correlation to the best performer in terms of accuracy (FCN-8s with Adam solver in WD; average processing time of 0.1 sec per image), the processing time was 30, 170 and 220 times faster than the frameworks in Chapter 6, 4 and 5, respectively. This indicates that in general, the FCN-8s model is ideal for the domain of semantic segmentation as it is not only accurate, but also considerably fast. Overall, the application of any pre-trained model for testing (semantic segmentation) the ROI of thigh MRI is reliable in terms of processing time, with FCN-8s (with Adam solver in MD) taking the longest time - 0.64 sec. And that still is 4.7 times faster than the proposed method in Chapter 6.

This also confirms: 1) the general influence of utilizing the simplicity of update value formulation of the weight W_{t+1} as per Equation 3.74 for SGD, evidently results in faster computation, compared to the complexity of Adams's weight update $(W_{t+1})_i$ as of Equation 3.77; and 2) vice-versa for the accuracy department, as Adam generally performed better throughout all datasets, compared to SGD.

7.4.3 Post-processing

Notwithstanding the promising results set out above, there is still a need to investigate the possibility of improving the end results of semantic segmentation by the FCN-8s model. Here, the inclusion of a post-processing stage by image processing techniques will be explored, with the goal of enhancing the ROI segmentation result. This FCN-8s model dedicated experimentations will be applied to both Adam and SGD solvers on all datasets (MD, WD and AD). The process involves image filtering and morphological procedures to the segmented ROI. Figure 7.7 demonstrates the significance of the inclusion of this stage and Table 7.5 shows the mean *JSI* performance index with the implementation of post-processing stage by image processing techniques described earlier.

The results above demonstrated that by incorporating the proposed post-processing stage, the semantic segmentation output could be enhanced. The overall accuracy performances are increased throughout all of the datasets, with FCN-8s (Adam solver in WD) achieved the best result. However, the performance increment was minimal, due to the fact that only $< 1\%$ (or 0.86% to be exact across output in Table 7.5) of semantically

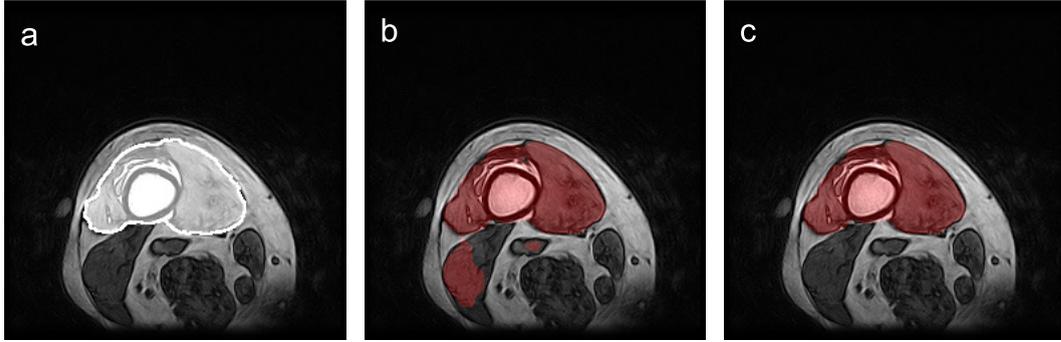


FIGURE 7.7: Image filtering and morphological to the semantic segmentation image. (a) Ground truth label (in white) with original MR image. (b) Semantic segmentation result by FCN8-s model (for this sample: with Adam solver in WD dataset). (c) Post-processing result of (b).

TABLE 7.5: Mean *JSI* performance index of FCN-8s model on mid-scan (MD), whole-scan (WD) and all-scan (AD) datasets, without and with post-processing, respectively.

		FCN-8s			
		Without post-processing		With post-processing	
		Adam	SGD	Adam	SGD
Dataset	MD	0.927	0.8973	0.927	0.9006
	WD	0.9469	0.9316	0.9502	0.9325
	AD	0.9388	0.9391	0.9398	0.9392

segmented images are affected by this “oversegmentation” condition as shown in Figure 7.7(b). With additional processing, the overall mean processing time to process a thigh MR image also increased. It is measured that an extra 17 ms (0.017 sec) is needed for this purpose (for FCN-8s with Adam solver in WD), increasing the overall processing time to $0.1 + 0.017 = 0.117$ sec, which was still 26, 145 and 188 times faster than the frameworks in Chapters 6, 4 and 5, respectively.

This also indicates that it has out-performed the mean *JSI* benchmark of 0.95 (or 95%) set in Chapter 4 for the semi-automatic segmentation framework, and therefore warrants the in depth analysis before further application. Table 7.6 shows the average time taken (in minute) per image for training and validation procedures for all models used. Red and green boxes represent procedures that was performed by either Machine 1 or Machine 2 configurations, respectively.

TABLE 7.6: Average time taken for training and validation procedures for all models used (in minute) for an image. Red and green boxes depict model that was trained in Machine 1 and 2, respectively.

		Model											
		FCN-AlexNet		FCN-8s					FCN-16s		FCN-32s		PSPNet
		Adam	SGD	AdaGrad	Adam	NAG	RMSprop	SGD	Adam	SGD	Adam	SGD	SGD
Dataset	MD	0.37	0.26	0.73	1.15	1.28	1.26	0.79	1.26	0.73	1.58	1.03	0.15
	WD	0.22	0.23	-	0.67	-	-	0.35	-	0.35	-	0.29	0.6
	AD	0.55	0.48	-	1.09	-	-	0.99	-	0.91	-	0.91	1.32

7.5 Discussion

Five pre-trained models of FCN-AlexNet, FCN-32s, FCN-16s, FCN-8s and PSPNet, from three existing deep learning architectures (AlexNet, VGG-16 and PSPNet) have been explored for the semantic segmentation task of the thigh MRI. All deep models (apart from PSPNet) were customized by removing the fully connected layers and replacing these layers with additional convolutional layers. These pre-trained FCN models are then transferred for the task of ROI semantic segmentation. Some of the benefits of this conversion and transfer learning process are: 1) capability to process input image of various sizes; and 2) the number of parameters to update are significantly reduced because of the weight distributions in a convolutional layer, making the learning process faster, allow robust pixel-wise prediction from the extracted feature maps and better overall performance.

FCN-8s emerges as the best all-around deep learning model for the task, and Adam solver type works the best overall for the thigh MR images, although evidently Adam solver also proved as one of the slowest, due to the complexity in its parameter updates of the weight during forward and backward pass. By default, this FCN-8s with Adam solver, trained, validated and tested on mid-scan (MD), whole-scan (WD) and all-scan (AD) datasets, produced a mean semantic segmentation accuracy by *JSI* of 0.927 (92.7%), 0.9469 (94.69%) and 0.9388 (93.88%), within an average processing time per image of 0.64 sec, 0.1 sec and 0.17 sec, respectively for the corresponding datasets.

With the implementation of post-processing stage of image filtering and morphology to the ROI segmented output, the results are enhanced to 0.927 (92.7%), 0.9502 (95.02%) and 0.9398 (93.98%) for MD, WD and AD, respectively. The best *JSI* performance result of 0.9502 (95.02%) suggests that FCN-8s model with Adam solver in WD has out-performed the mean *JSI* benchmark of 0.95 (or 95%) set in Chapter 4 for the semi-automatic segmentation framework. And this was achieved by segmenting images

throughout the whole dataset (one MRI dataset consist of either 13 or 26 images) not just some random MR image or mid-scan MR image like the analysis in Chapter 4 and 5. Overall, the performance of the combination framework of customized deep CNN with image processing technique is exceptional, in terms of accuracy (0.9502 or 95.02%), speed (0.117 sec testing/segmentation time per image) and reliability (< 1% of images are affected by issue related to oversegmentation).

However, in machine learning, having a sole end-to-end network for your domain task is preferable and considered as the best solution for an automated domain system. And since by default (without the inclusion of post-processing stage), the output performance of FCN-8s is comparable to the new benchmark achieved, in the future it is: (1) best to explore other potential deep CNNs that are specified in the semantic segmentation domain (such as U-Net [260], SegNet [261] and FCN-DenseNet [262]) and investigate, evaluate and compare their performances with the current benchmark method; (2) making adjustments to FCN-8s parameters (type of layer used, additional dropout loss if necessary or enhance batch normalization) and hyperparameters (learning rate, momentum or regularization coefficient) for better optimization in terms of training and testing; and (3) the prospect of developing our own deep nets that specifies and optimizes to the ROI segmentation for significant scientific outcome.

Hypothetically, in our endeavour to engineer an end-to-end solution for ROI segmentation by using deep learning algorithm, proposing our own network can possibly be accomplished by: (1) training the model with more MRI samples; (2) re-evaluation on derivation of the normalized initialization for better regularization; and (3) configuring the best architecture model for weight optimization and distribution by referring to the analysis gathered from exploring other semantic segmentation specialized models as above.

7.6 Conclusion

In general, automatic segmentation of ROI of MRI of the thigh by classic image processing technique, computer vision and/or statistical analysis algorithms is prone to susceptibility due to their reliance on specific rules with certain assumptions. Here, we have successfully demonstrated the application of several ANNs in the form of FCNs that

may progressively expand their application towards the automated analysis on medical imaging. End-to-end performances of accuracy (by *JSI*) and processing time by FCN-8s (a VGG-16 based CNN) with Adam solver type are 0.9469 (94.69%) and 0.1 sec per image, respectively. Combined with additional post-processing stage, the segmentation accuracy increased to 0.9502 (95.02%), whilst the processing time increased to 0.117 sec per image. Overall, the automatic segmentation framework by FCN-8s (with Adam solver) is fast, accurate and robust and has out-performed the mean *JSI* benchmark of 0.95 (or 95%) set in Chapter 4.

Chapter 8

Conclusion

This chapter summarizes the main research findings and interprets them in respect of the existing literature and other solutions to similar challenges for ROI segmentation of skeletal muscles as well as highlighting possible future directions.

8.1 Introduction

For the past thirty years, medical imaging has developed significantly in terms of design and engineering; image or data acquisition; reliability and sustainability; and output performances. For the output performance, medical image analysis has gradually transformed from a fully supervised technique to a more flexible unsupervised method, whether it is for classification, detection, localization or segmentation.

The work set out in this thesis focused on medical imaging analysis of human skeletal muscles obtained by MRI. This focus was chosen because human muscles play important roles in providing support for posture and making movement possible, in thermoregulation and in paracrine and endocrine functions and in any situation where muscle mass and structure are compromised it has the potential to restrict physical function and cause disability. During ageing, physical activity levels decrease and the consequent disuse alongside other biological processes such as hormonal and/or other endocrine changes contribute to muscle atrophy or muscle wasting and in turn muscle weakness. Several large-scale research projects to investigate the causes of atrophy and weakness

of leg muscle have been performed. In one such study at Manchester Metropolitan University (MMU), MR images were collected from hundreds of adults aged between 18 and 90 years. The MR images can be analyzed for muscle, fat and connective tissue content by examining pixel areas and pixel intensity, and quantification of the results reveals a lower muscle mass and more fat infiltration in older people compared to young people (the effects of ageing or disease).

The current “gold standard” manual analysis for specified regions of interest (ROI) segmentation for the thigh MRI is well established but laborious, time consuming, operator dependent and prone to intra- and inter-operator variability. Therefore, the main goal of the work presented in this thesis was to automate the current image analysis of thigh MRI without compromising output performance. An automated segmentation framework would offer substantial time saving, especially for scientists that are studying the effects of ageing or other conditions on “muscle quality” using MR images of the thigh skeletal muscles.

8.2 Research Findings

The five major objectives of the projects are: 1) establishing the best algorithms for the ROI segmentation task; 2) designing a reliable, fast and accurate ROI segmentation and quantification framework for thigh MRI; 3) developing an automated segmentation and quantification framework for MRI of the thigh. The framework should provide acceptable results without any supervision; 4) proposing an ROI segmentation and quantification framework for serial MR images within the same dataset (subject); and 5) designing a semantic segmentation framework for MR image of the thigh based on current machine learning approach of deep learning algorithm. The ROI (for all segmentation frameworks) consists of quadriceps muscle, femur and bone marrow.

In the first step (Chapter 4), image processing techniques such as thresholding, followed by convex hull pixel identification for unwanted background removal and analysis of cross-sectional area for the ROI were combined with the human input to delineate the quadriceps from the hamstring muscles. The mean accuracy as determined by *JSI* and processing time per image for this semi-automated segmentation framework were 0.95 and 17 sec, respectively. The proposed algorithms in this chapter were developed to

improve speed (a time saving of 87% compared to manual segmentation), efficiency and increase the ability to distinguish between different tissue types and generate quantitative data. Although this foundation segmentation framework yields accurate results, reliable and fast, it was semi-supervised (partially operator-dependent) and analysis was done on MR images within the middle region of a dataset, where images are tend to be less affected by bias field.

Next, the research continued with the development of a fully automated process by using a combination of statistical analysis and image processing algorithms without the need of user interactions (Chapter 5). The analysis was done exclusively on a single transverse-plane section of the mid-thigh MR images. The overall performance for this automatic segmentation framework was 0.8526 and 22 sec, for mean segmentation accuracy by *JSI* and processing time per image, respectively. Three major limitations of this proposed framework will be explored in future investigations: 1) The inclusion of more mid-scan MR images in the training dataset to re-construct the shape model (by PCA) for a more robust representation of the initial contour before the application of deformable model; 2) The re-configuration of optimized values of the region-based active contour dependent constants k_a and k_b to enhance the computations of average derivation of inside and outside of shape model's placement; and 3) In theory, a more optimized and robust ROI segmentation output performance can be accomplished by engaging a more dynamic and adaptive computation of the transformation properties (scaling factor s , orientation Δ , and rotation angle ϕ) of the shape model.

In the next step, the development of combination frameworks that leads to an automatic MRI dataset segmentation method was proposed (Chapter 6). To acquire the best results, the semi-automated framework of Chapter 4 was re-applied and this is done by only one simple manual interaction to segment the mid-scan MR image. Serial images in the same MRI dataset are then automatically segmented by the customized framework of snake active contour curve transformation, based on template geometric information of the earlier semi-automatically segmented mid-scan image as the shape regulator. Overall, the method needs minimal supervision and offers exceptional performance in terms of segmentation accuracy (mean *JSI* of 0.9334) and processing time (mean of 3 sec per image). As part of the segmentation process was originated from the foundation framework of algorithms in Chapter 4, this proposed framework also limited by its

dependency on human interaction. However, this restriction was progressively refined and the results were reported in Chapter 7.

With the aim of replacing the semi-automatic segmentation method in Chapter 4 and Chapter 6 to a fully automatic thigh MR image segmentation method, the next step was to apply an artificial neural network, a concept in machine learning that is inspired by the nervous system. Five pre-trained models of FCN-AlexNet, FCN-32s, FCN-16s, FCN-8s and PSPNet, from three existing deep learning architectures (AlexNet, VGG-16 and PSPNet) have been explored for applicability to the semantic segmentation task of thigh MRI. All deep models (apart from PSPNet) were customized by removing the fully connected layers and replacing these layers with additional convolutional layers. These pre-trained FCN models were then transferred for the task of ROI semantic segmentation. This improved capability of the models to deal with various sizes of input images and most importantly, results in a significant reduction in training and testing time due to the weight optimization. With the implementation of the post-processing stage of image filtering and morphology to the ROI segmented output by FCN-8s model (with Adam solver) on whole-scan (WD) dataset, the results have out-performed the benchmark set in Chapter 4. However, further investigations are desired for the end-to-end solution by machine learning algorithm and to do so, requirements have been defined that may improve the image segmentation further such as: 1) training the model with more MRI samples; 2) re-evaluation on derivation of the normalized initialization for better regularization; and 3) configuring the best architecture model for weight optimization and distribution by referring to the analysis gathered from exploring other semantic segmentation specialized neural networks. Conclusively, this is one potential avenue for future work.

The research has overall progressed from a manual segmentation performed by specialized technicians prior to any research set out in this thesis and moved through novel frameworks first requiring manual and supervised interactions, towards a fully automated framework able to accurately quantify skeletal muscle tissue and bone of the human thigh. The mid-scan is universally the preferred MR image to process as it is here that muscle has the largest cross-sectional area and images have fewer artefacts, echo and other substantial noise such as intensity inhomogeneity, which is a consequence of gradient pulse eddy current effects and inherent effects of pulsed radiation on MRI radio frequency (RF) coils from the MRI machine.

Table 8.1 outlines the results of all contributory chapters (Chapter 4-7) of this research and illustrates the progress of the research over time, in general.

TABLE 8.1: Mean performance index of research findings from Chapter 4, 5, 6 and 7, including manual “ground truth” procedure.

	Ground Truth	Chapter 4	Chapter 5	Chapter 6	Chapter 7
Segmentation Framework	Manual	Semi-automatic MR Image	Automatic MR Image	Automatic MRI Dataset	Semantic (Automatic)
Mean Accuracy (<i>JSI</i>)	-	0.95	0.8526	0.9334	0.9502
Standard Deviation	-	0.0091	0.0915	0.0352	0.0184
Mean Processing Time (per image)	132 sec	17 sec	22 sec	3 sec	0.117 sec

The MRI dataset used in this project is currently unavailable for public use in any form or media but can be requested and released upon consensus and mutual agreement from both the potential researcher(s) and research group of The Institute for Biomedical Research into Human Movement (IRM) and School of Healthcare Science, Manchester Metropolitan University (MMU). Benchmarking the proposed frameworks in the contributory chapters of this thesis with other MRI datasets used in previous works of thigh MRI segmentation for immediate, explicit and thorough comparisons was also not feasible at the moment, due to the similar inaccessible reason (and vice-versa scenario for our MRI dataset implementation on previous related works, i.e. the codes were not available to public and due to the framework’s complexity, it was laborious to replicate the previous related works).

The database of our MR images was derived from a 0.25-Tesla MRI scanner (Esaote G-scan, Genoa, Italy), which generates a lower intensity of magnetic field and more limited scanning area compared with more common 1-Tesla, 1.5-Tesla or 3-Tesla MRI scanners generally used from previous related works. Even with such limitation, our analysis has proven of its robustness and can be used to analyse thigh images derived from any MRI scanner from 0.25-Tesla or higher. Technically, with such quality, the proposed frameworks also can be applied to other MR images with virtually similar transverse plane configurations of body segments such as lower leg or upper arm MRIs, that are acquired also from any MRI scanner from 0.25-Tesla or higher.

8.3 The Challenges

Figure 8.1 demonstrates the MR images of the thigh of three different subjects, from proximal, mid-scan and distal scans.

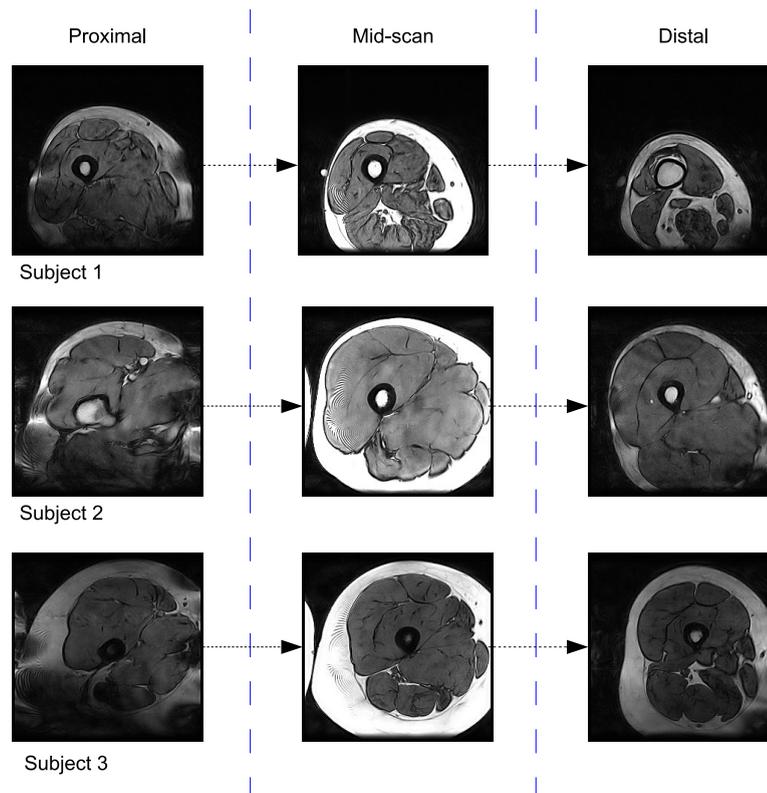


FIGURE 8.1: Configuration of MR images from three different subjects, from proximal, mid-scan and distal scans.

Processing and analysing MRI scans with different muscles configurations, even within the same subject, were some of the major challenges for this research as the shape of the individual muscle changes significantly along the length of the muscle (as illustrated in Figure 8.1). Other substantial noise such as intensity inhomogeneity, image artefact and echo can also be observed from this figure and such complications are always associated with MR images in proximal and distal sections and may considerably affecting the overall performance of the segmentation framework.

In terms of ROI analysis, the manual segmentation is prone to intra- and inter-operator variability [8, 11]. Even if the manual annotation is done by the same person on the same MR image, it is difficult to get 100% accuracy.

Moreover, in OsiriX, the output resolution of the ground truth image is natively processed in 1526 x 1526 resolution. This raster (pixel-based) image is then downsized to a

256 x 256 dimension for performance measurement, as the analysis of input and output of proposed segmentation frameworks are all done within the 256 x 256 resolution. This downsizing reflected upon the proposed algorithm performances, especially when processing MR images with substantial edges/corners of the ROI (usually near the exterior region of the Rectus Femoris muscle) and could suggest another strain towards achieving the 100% accuracy.

Ideally, having a great amount of data is paramount to efficiently train a deep neural network. However, in the real world it is not often possible and for this research, additional labelled images for training means extra time has to be reserved and spent for the manual annotation procedure and this was another challenge that associates to the research.

8.4 The Trials

After the proposed semi-automatic segmentation framework, an approach towards automatic segmentation method by combining the framework of atlas construction (based on segmented mid-scan MR image) and image registration methods for MRI thigh segmentation were explored. These methods propagate the desired ROI between atlas image and the targeted MR image for quadriceps muscles, femur cortical layer and bone marrow segmentations. Among all image registration methods that can be put to use in achieving model-based segmentation, one popular approach is based on maximizing mutual information [263, 264]. Good results with this technique have been reported for rigid and non-rigid registration [162, 265, 266]. Atlas-based segmentation was successfully achieved using the fluid-flow model introduced in [267] and overall was computationally efficient. Authors in [268, 269] presented a joint registration-segmentation scheme, but could only accommodate rigid motion and an improved method was later developed by using the Bayesian method [270] that simultaneously estimates linear registration. Methods in [271–273] worked with non-rigid registration. However, these methods only work well with images that are matched/paired from the same imaging modality or with similar intensity profile. Our experimentations for this atlas-based registration approach suffered in all performance departments (accuracy, processing time and reliability), especially the analysis on sequential images apart from a few middle-scanned images. Figure 8.2 depicts our trial with this method.

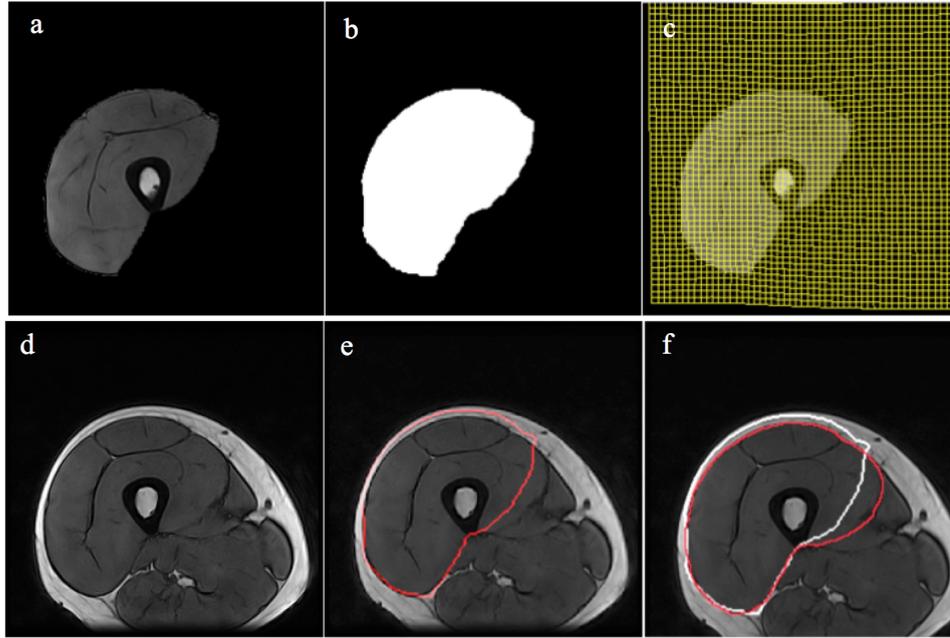


FIGURE 8.2: Atlas-registration method. (a) Atlas (reference) image. (b) Atlas region. (c) Atlas registration. (d) Moving (targeted) image. (e) Overlay of atlas boundary on moving images. (f) Ideal atlas-registration process.

Another trial includes the computation of the “suspected area” of muscle border, which is an area that statistically lies within ± 15 pixels (in x -axis) of centroid location of mid-scan image, with the orientation of a shape model (that was formed by PCA, in Chapter 5). Hessian and Canny Edge Detection methods are employed to segregate any edges within this “suspected area”, combined with thresholding to filter any unwanted (or weak) edges. The “suspected area” with its content (of useful objects associated with muscle border) is extracted. Any objects unequal to orientation of the shape model and with a size that is substantially small (≤ 5 pixels), are then eliminated. Remaining objects are thresholded by morphological application of a skeleton. This gives us information such as branch-points and end-points of the remaining objects, which are then utilized to predict the muscle border, by using polynomial coefficients. The measured spline curve is used as part of the regulator for the mid-scan image before the process of convergence by active contour takes place. This method was evaluated as a means to regulate the spline curve estimation area that represents the muscle border between the quadriceps and the rest of the thigh muscles (such as hamstring, etc). In the ideal scenario (MR images with prominent image features that represent muscle borders and little artefact), the required information can be extracted and processed accurately with this method. However, the reality of the images used throughout this

thesis meant that the analysis was badly affected by the generic estimation results and it was therefore deemed unsuitable for the ROI segmentation of thigh MRI. Figure 8.3 shows our trial with this method. For this reason, it was necessary to develop the novel solutions set out in this thesis.

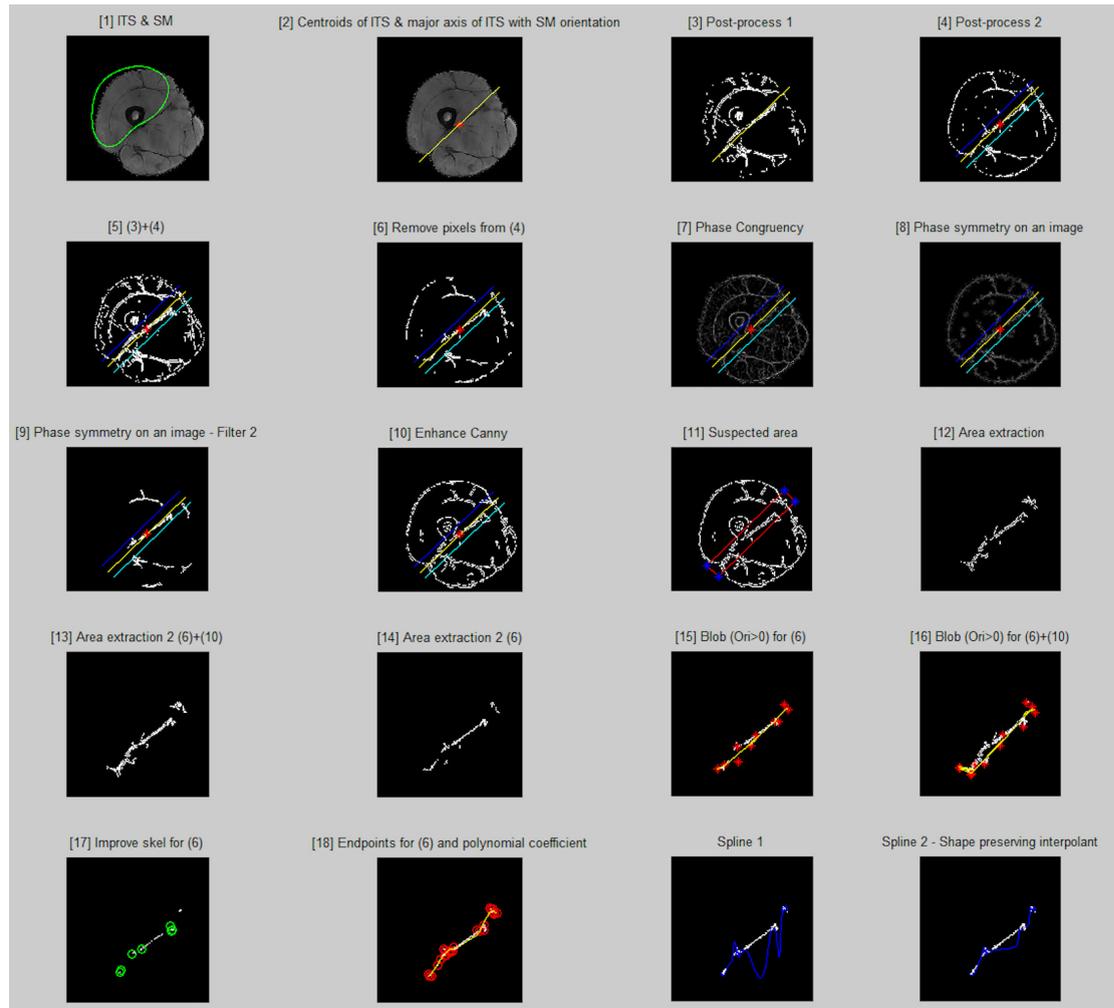


FIGURE 8.3: Procedure of isolating the “suspected area” and spline curve estimation.

8.5 Future Works

Although the overall performance of the proposed segmentation framework by fully convolutional network model can be regarded as the best solution, there are still certain areas that can be improved. These future works or further recommendations are outlined as follows:

1. The possibilities of investigating and studying with raw DICOM (.dcm) image resolution (rather than just a 256 x 256 image resolution that we are currently working with) and analysing the core differences. In theory, output analysis and performance could be enhanced, due to the additional feature maps (such as edges and curves) that can be extracted by the algorithm rather than working within the restrictions of a confined and compressed image.
2. The re-configuration of optimized values of the region-based active contour dependent constants k_a and k_b of the framework in Chapter 5 to enhance the computations of average derivation of inside and outside of shape model's placement.
3. With the replacement of a fixed value of scaling factor s , together with the application of more dynamic and adaptive computation of orientation Δ and rotation angle ϕ of the shape model, the proposed segmentation method may be optimized further to provide better overall performance.
4. As mentioned in the conclusion section of Chapter 7, having a sole end-to-end network for the segmentation task is preferable and considered as the best solution for an automated domain system in machine learning. By default (without the inclusion of post-processing techniques), the output performance of FCN-8s is still credible to the new benchmark achieved (mean segmentation accuracy by *JSI* of 0.9469 compared to 0.9502). In the future it is best to explore other potential deep CNNs that are specified in the semantic segmentation domain (such as U-Net [260], SegNet [261], FCN-DenseNet [262] and many more) and investigate, evaluate and compare their performances with the current benchmark method.
5. Formulating prospective adjustments to FCN-8s parameters, hyperparameters or overall architecture for better regularization and optimization in terms of training and testing.
6. Developing and assembling our own deep convolutional network that specify and optimise to semantic segmentation of the ROI.
7. To further optimise the weight initialization of the deep network model, more MR images and datasets for training and validation procedure will be performed. This may also include MR images acquired from other MR machines and protocols, to develop a universally applicable segmentation solution.

8. Future developments should find solutions to provide clinicians with the individual details of each of the quadriceps muscles (Vastus Lateralis, Vastus Medialis, Vastus Intermedius and Rectus Femoris) for a more robust and detailed framework.

8.6 Final Remarks

This thesis has presented works designed to establish and improve the current image analysis of thigh MRI. The proposed automatic segmentation framework is developed and analyzed from the combination of image processing and the novelty and emerging machine learning techniques and yields exceptional ROI segmentation output performances (excellent precision, fast processing time and robust) and is hoped to be of benefit and utilized for the medical community.

Appendix A

OsiriX

A.1 3-D Viewer for Volume and Surface Rendering

Figure A.1 shows a concept on how the initiation of ROI volume reconstructions would be. The red-ish or brown-ish region represents segmented muscles fibers and yellow-ish regions represent femur bone marrow. 3-D reconstruction would require the simulation of fiber generation to fill in the gaps for one thigh MRI dataset/scans (muscles and marrow gaps as per figure A.1).

A.2 Area Conversion

Figure A.2 demonstrates optimum dimension measurements of MATLAB and OsiriX applications used in the development of semi- and fully-automatic segmentation framework. The framework's code structure was designed and processed using MATLAB and since MATLAB is strictly matrix dependent, the processed output (for image) data will always displayed in pixels. Pixel to cm^2 area conversions is required as a quantitative measurement and comparison for physiologists and physicists. Red arrows correspond to exact output image dimension view size in MATLAB environment while blue arrows associate with OsiriX. Therefore the conversion of pixel to cm^2 :

$$D_{cm} = \frac{D_{pix} \times 18}{256}, \quad (\text{A.1})$$

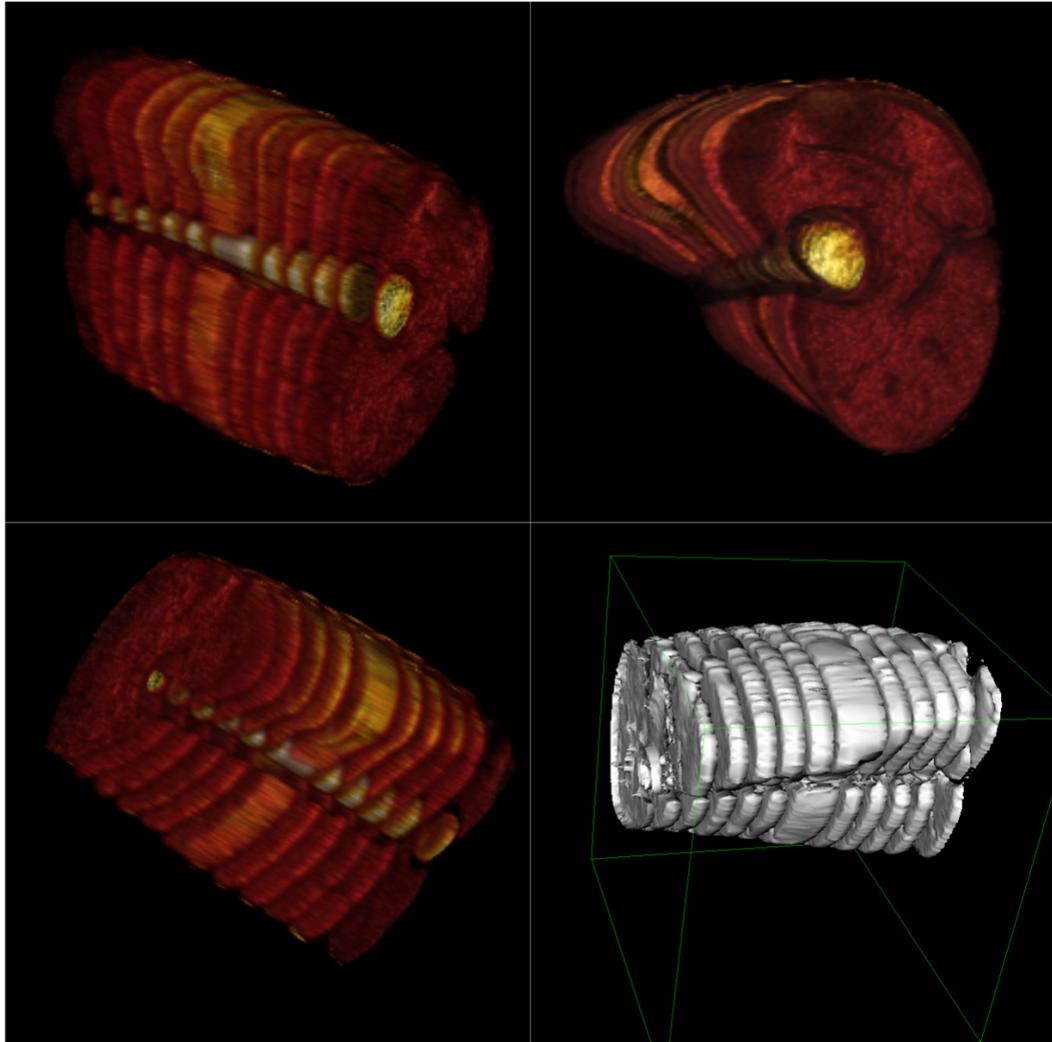


FIGURE A.1: 3-D volume rendering by OsiriX for one thigh MRI dataset.

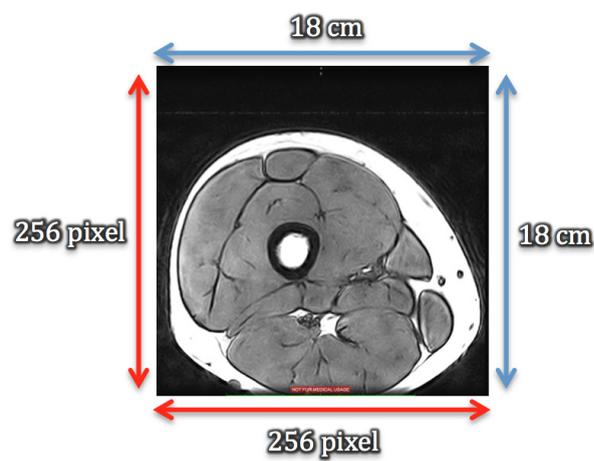


FIGURE A.2: Corresponding dimensions of pixel (in red arrows from MATLAB) and cm (in blue arrows from OsiriX).

where D_{pix} = ROI equivalent diameter (in pixel); D_{cm} = measurement of ROI equivalent diameter in cm. Area (in cm^2) can then be measured by circle surface equation of:

$$Area(cm^2) = \frac{\pi(D_{cm})^2}{4} . \quad (A.2)$$

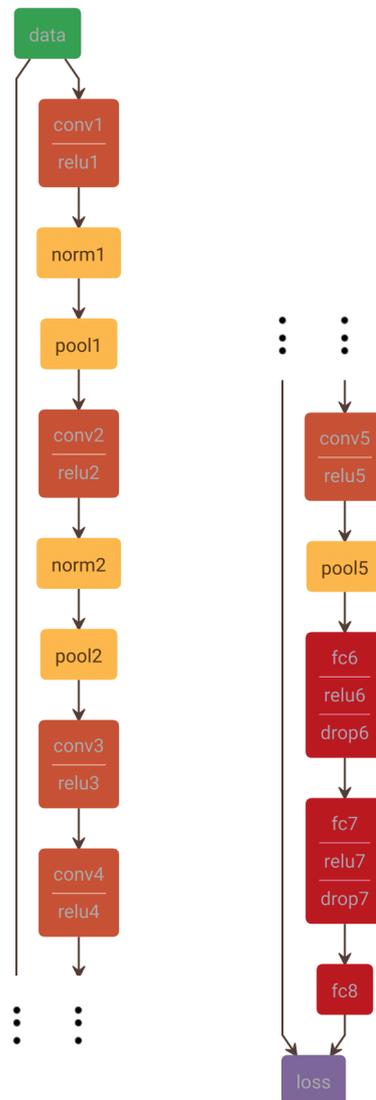
Appendix B

Architecture Visualization for Convolutional Neural Network

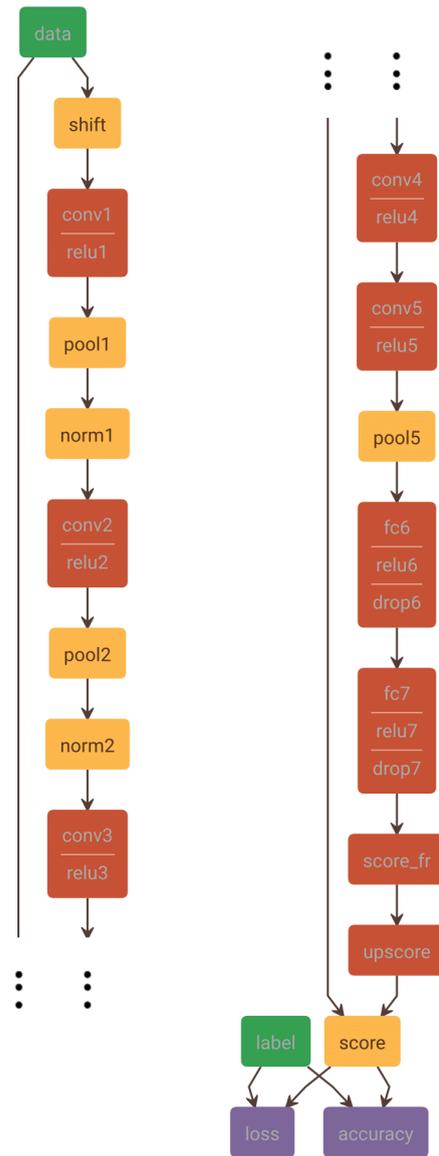
This appendix demonstrates the architecture visualization of Convolutional Neural Networks employed in Chapter 7.

B.1 AlexNet Based Models

B.1.1 AlexNet

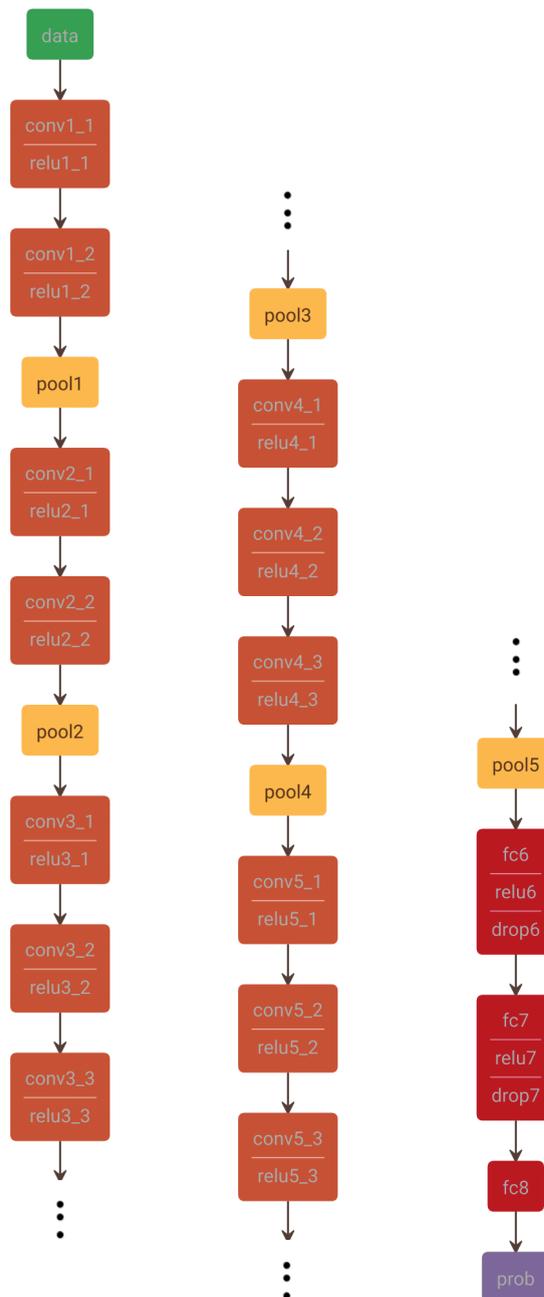


B.1.2 FCN-AlexNet

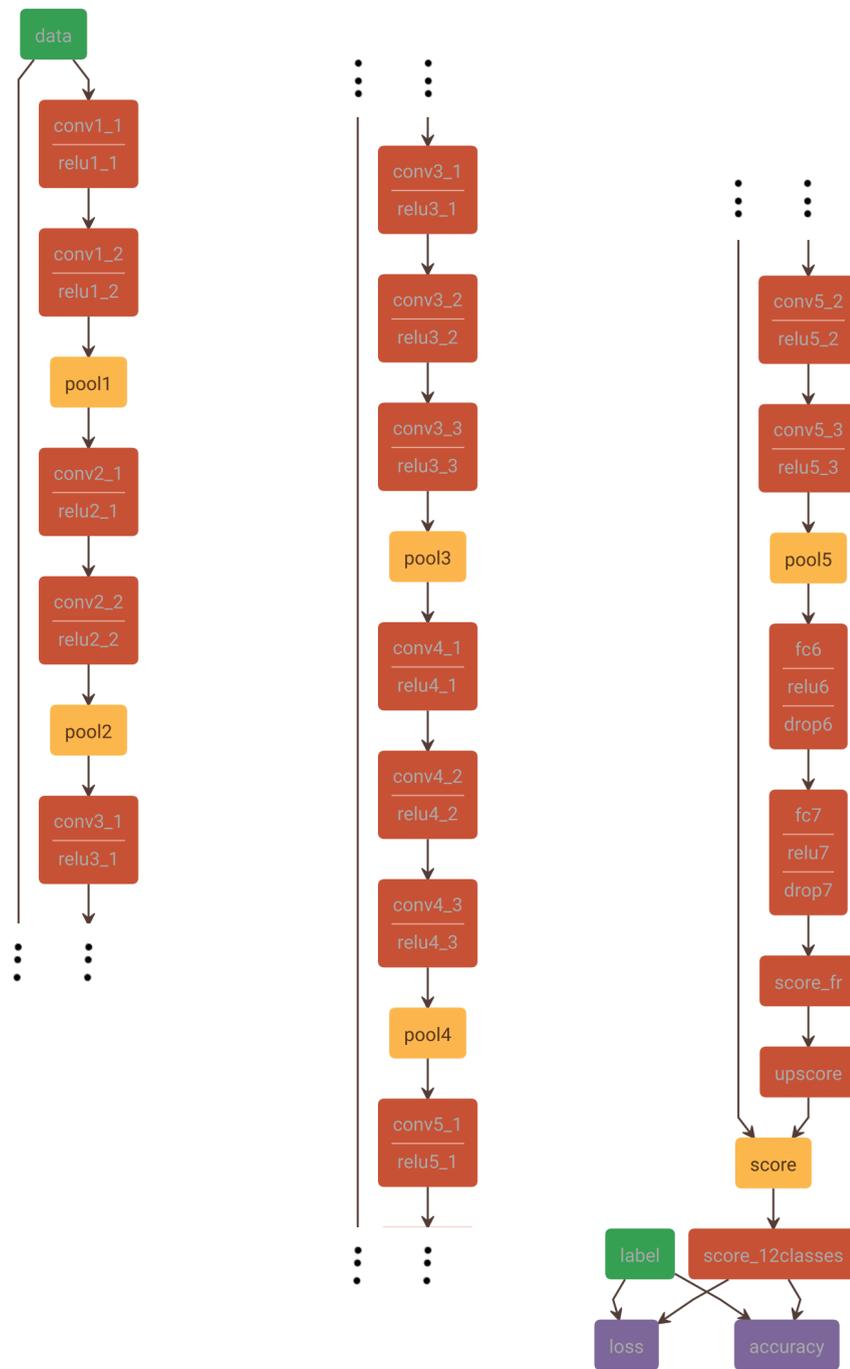


B.2 VGG-16 Based Models

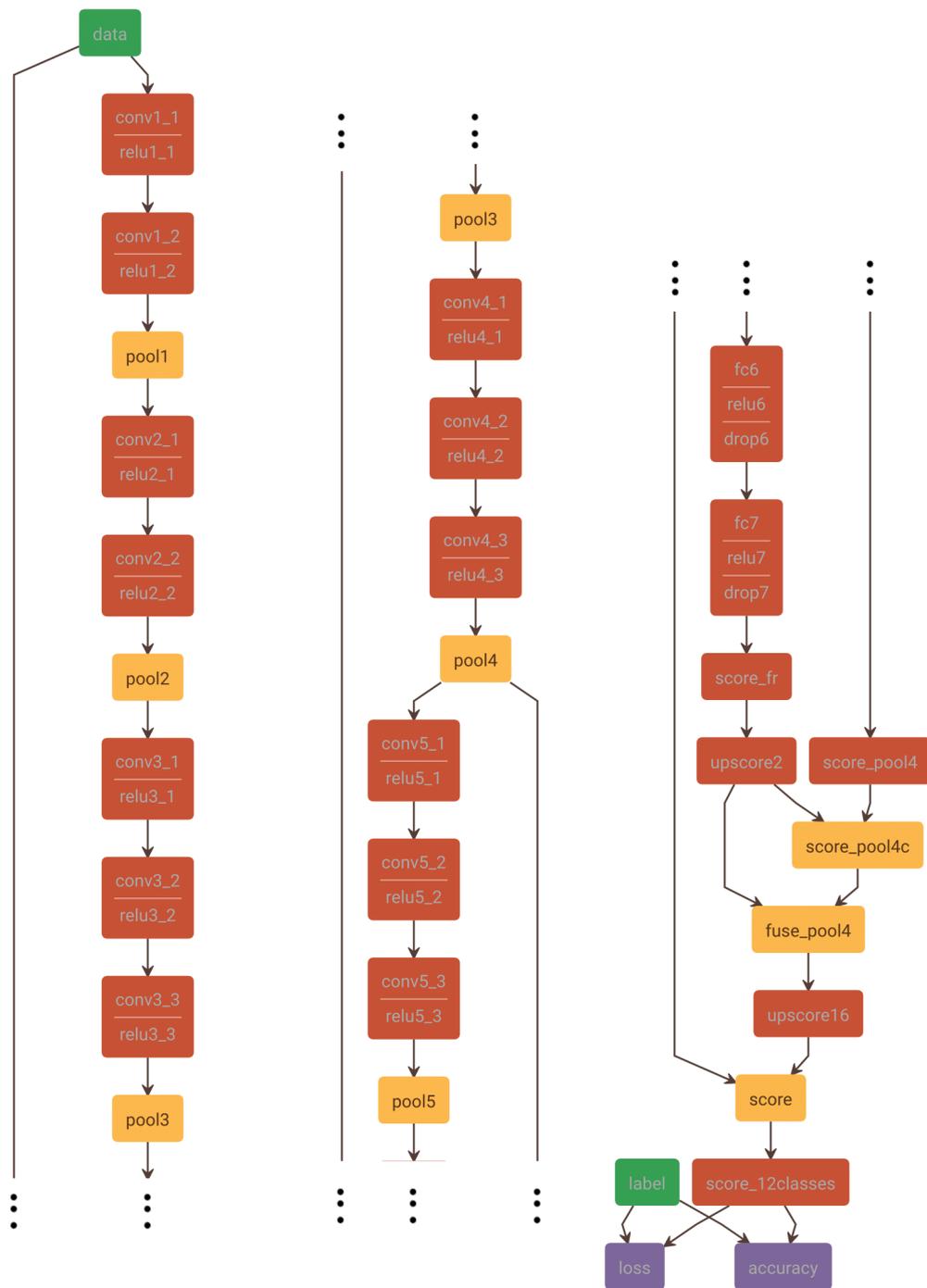
B.2.1 VGG-16



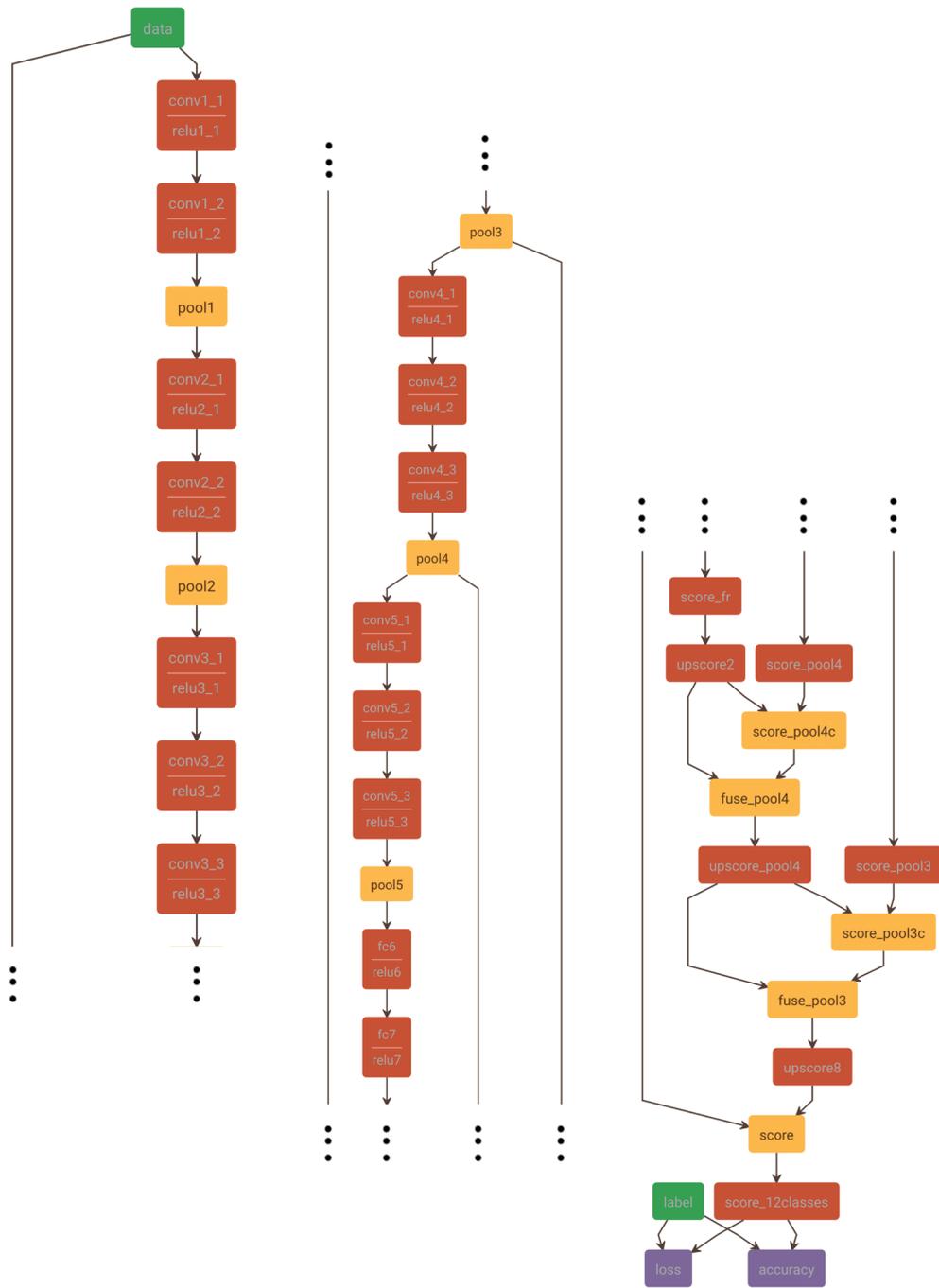
B.2.2 FCN-32s



B.2.3 FCN-16s



B.2.4 FCN-8s



Appendix C

Publications

All (front pages) of the related publications during the research progress are attached to this appendix.

C.1 Published Article(s)

C.1.1 SPIE 2014

Atlas-registration based image segmentation of MRI human thigh muscles in 3-D space

Ezak Ahmad^a, Moi Hoon Yap^a, Hans Degens^b and Jamie S. McPhee^b

^a*School of Computing and Mathematics, Manchester Metropolitan University, Manchester, United Kingdom*

^b*School of Healthcare Science, Manchester Metropolitan University, Manchester, United Kingdom*
{e.ahmad, m.yap, h.degens, j.s.mcphee}@mmu.ac.uk

ABSTRACT

Automatic segmentation of anatomic structures of magnetic resonance thigh scans can be a challenging task due to the potential lack of precisely defined muscle boundaries and issues related to intensity inhomogeneity or bias field across an image. In this paper, we demonstrate a combination framework of atlas construction and image registration methods to propagate the desired region of interest (ROI) between atlas image and the targeted MRI thigh scans for quadriceps muscles, femur cortical layer and bone marrow segmentations. The proposed system employs a semi-automatic segmentation method on an initial image in one dataset (from a series of images). The segmented initial image is then used as an atlas image to automate the segmentation of other images in the MRI scans (3-D space). The processes include: ROI labeling, atlas construction and registration, and morphological transform correspondence pixels (in terms of feature and intensity value) between the atlas (template) image and the targeted image based on the prior atlas information and non-rigid image registration methods.

Keywords: Image segmentation, image registration, magnetic resonance imaging (MRI)

1. INTRODUCTION

Human thigh muscles play an important role in locomotion and other interactions with the environment. Thigh muscles consist of two main compartments - the quadriceps muscle group on the anterior and the hamstrings on the posterior. Physical activity levels evidently decrease during aging [1] and the consequent disuse can impact on muscle size and function. The effect of ageing is not only limited to disuse, factors such as hormonal and other endocrine changes also contribute to muscle degeneration [2]. Figure 1, shows a transverse-plane cross section of the mid-thigh in a young (Figure 1(a)) and older (Figure 1 (b)) man and it is clear that the muscles of the older man were smaller and the whole-thigh had more fatty adipose tissue.

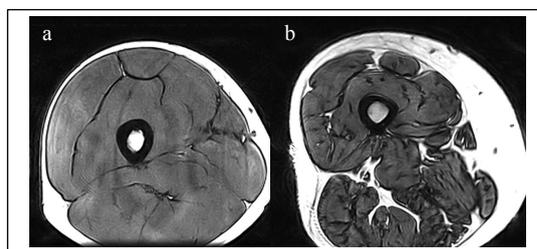


Figure 1: MRI scans from (a) Young and (b) Elderly.

Medical Imaging 2014: Image Perception, Observer Performance, and Technology Assessment, edited by Claudia R. Mello-Thoms, Matthew A. Kupinski, Proc. of SPIE Vol. 9037, 90371L
© 2014 SPIE · CCC code: 1605-7422/14/\$18 · doi: 10.1117/12.2043606

Proc. of SPIE Vol. 9037 90371L-1

Downloaded From: <http://proceedings.spiedigitallibrary.org/> on 03/24/2014 Terms of Use: <http://spiedigitallibrary.org/terms>

C.1.2 ICCSCE 2014

Enhancement of MRI Human Thigh Muscle Segmentation by Template-based Framework

Ezak Ahmad*, Moi Hoon Yap*, Hans Degens[†] and Jamie McPhee[†]

*School of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, Manchester, M1 5GD, UK

[†]School of Healthcare Science, Manchester Metropolitan University, Manchester, M1 5GD, UK

Abstract—Image segmentation of anatomic structures is often an essential step in medical image analysis. A variety of segmentation methods have been proposed, but none provides automatic segmentation of the thigh. In magnetic resonance images of the thigh, the segmentation is complicated by factors, such as artifacts (e.g. intensity inhomogeneity and echo) and inconsistency of soft and hard tissue compositions, especially in muscle from older people, where accumulation of intermuscular fat is greater than in young muscles. In this paper, the combination framework that leads to a segmentation enhancement method for region of interest segmentation are demonstrated. Appropriate methods of image pre-processing, thresholding, manual interaction of muscle border, template conversion and deformable contours in combination with image filters are applied. Prior geometrical information in an initial template image is used to automatically derive the muscle outlines by application of snake active contours, in serial images within a single MRI dataset. Our approach has an average segmented output accuracy of 93.34% by Jaccard Similarity Index, and reduced the processing time by 97.73% per image compared to manual segmentation.

Index Terms—Magnetic resonance imaging, image segmentation, active contour

I. INTRODUCTION

Medical imaging provides a vast amount of physiological and functional information, which improves diagnosis and patient treatment, especially when supported by modern quantitative image analysis methods. However, the analysis of images to provide information to clinicians and the patient presents many problems and requires substantial technical and medical expertise. By definition, segmentation in computing is a process that allows the partitioning of an image (in 2- or 3-dimensions) into contiguous regions with mutual and cohesive properties (such as intensity or texture).

Manual segmentation of specified regions of interest (ROI) in MRI is well established, but laborious, time consuming and operator dependent [1], [2]. Methods for performing segmentation by computer vision algorithms vary widely depending on the specific application, imaging modality and targeted anatomical area. A reliable computerized (semi-automatic or fully-automatic) segmentation system will help tremendously in large-scale research to increase efficiency and reduce analysis time compared to manual segmentation. However, difficulties can arise due to image artifact, noise, echo, overlapping of pixel/voxel intensities and non-uniform 2-D pixel intensity on its output.

In recent years, several segmentation algorithms have been developed and applied to a small number of images, but the issue of compatibility of these algorithms to one another has not been properly documented. A popular technique for image segmentation proposed in [3], derived from k-means method, is the fuzzy c-mean (FCM) algorithm. It performs fuzzy partitioning through iterative optimization of the objective function with update to cluster centers. However, the algorithm does not incorporate any information about the spatial context, which makes it susceptible to noise and artifacts [4]. To overcome this issue, various modifications that integrate local spatial information as constraints in the objective function were proposed, such as by Li and Shen [4]. Yang *et al.* [5] and Li *et al.* [6] described a comprehensive version of the robust fuzzy clustering method (RFCM) for noisy image segmentation. While Xue-xi *et al.* [7] suggested a hybrid intelligent color segmentation method, which combines the region growing and clustering methods.

The application of active contour models or snakes for ROI segmentation or extraction is a common approach in computer vision and medical image analysis and to date, these techniques have been progressively developed and carefully adapted (according to datasets to be processed). Generally, this framework attempts to minimize an energy associated to the initial contour and object contour as a sum of an external or internal energy. Conventionally, active contour models can be divided into several groups, with two of the major ones being; edge-based [8], [9] and region-based [10], [11] models. Meanwhile, over- or under-segmentation commonly occurs in snakes due to its energy function sensitivity, weak or without edges and gradient variations in an image. Researchers [12]–[17] have incorporated active contour when segmenting various kinds of images, including medical images. Therefore, to enhance the snake's deformable model robustness or border convergence, practical customizations with other algorithms should be considered.

In a recent publication, Ahmad *et al.* [18] used Atlas-Registration for automated segmentation of the quadriceps muscles of the human thigh. An atlas image was constructed from the output of a semi-automatic segmented region. Intensity, energy gradient, geometric formation and feature information of the atlas were registered and these details were used as references for automatic segmentation of serial MRI scans within the same dataset by using image registration method.

C.1.3 ICBET 2018

Automatic Segmentation of MRI Human Thigh Muscles: Combination of Reliable and Fast Framework Methods for Quadriceps, Femur and Marrow Segmentation

Ezak Ahmad^a, Jamie S. McPhee^b, Hans Degens^b, Moi Hoon Yap^a

^aSchool of Computing, Mathematics and Digital Technology

^bSchool of Healthcare Sciences

Manchester Metropolitan University, Manchester, UK

ezak@dr.com

ABSTRACT

This paper outlines the procedure for automated segmentation of human thigh magnetic resonance images. Datasets that included sequential transverse-plane images (approximately 20 sequential images) were used to develop the framework and the images of mid-thigh were selected to establish the segmentation algorithm because these images contain essential parametric and geometric information that represents the subsequent images of the same dataset. Hence, if the mid-thigh can be decomposed accurately, it would significantly improve the prospects of decomposing all remaining transverse images within the dataset. The segmentation framework was designed based on the integration of statistical signal processing and medical imaging algorithms. The results showed good segmentation accuracy and fast processing time.

CCS CONCEPTS

• **Computing methodologies** → **Image processing**;

Keywords

Magnetic resonance imaging (MRI), Image processing, Automatic segmentation, Thigh muscles, Quadriceps.

1. INTRODUCTION

Skeletal muscle strength is a main component of functional capacity and when strength of leg muscles is low it can impact adversely on a person's mobility levels and quality of life. For these reasons, strength is commonly assessed by research studies into physical health [1, 2, 3, 4, 5]. The main interest is to understand the association between muscle strength and functional deficiencies, what causes such conditions and its mechanisms. To date, it is still not clear what causes the age associated reductions in muscle mass and strength, but they are ascribed to an inherent ageing process [6], that is often modified by dietary deficiencies [7, 8] and/or disease [9].

ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.

ICBET '18, April 23–25, 2018, Bali, Indonesia
 © 2018 Association for Computing Machinery.
 ACM ISBN 978-1-4503-6369-3/18/04...\$15.00
<https://doi.org/10.1145/3208955.3208958>

Muscle strength is often presented as the maximum amount of force that someone can exert, regardless of body size or weight and naturally it is higher in those with larger muscles, but there is also variation between people in the force that can be produced per unit muscle mass and this is sometimes referred to as muscle "quality". The muscle quality can be estimated from a simple division of maximum force into muscle cross sectional area. Measurements of skeletal muscle force are straightforward and require the patient simply to apply as much force as possible against an immovable object. The object includes a strain gauge that measures forces. However, measuring skeletal muscle mass is more difficult [10].

In medical imaging, various image modalities has been used to screen different human anatomy [11, 12]. For thigh, there are various techniques commonly used, ranging from limb-circumference which is fraught with error, through the modern clinical x-ray or magnetic resonance imaging (MRI) techniques, to computer vision technique using depth sensor [13, 14]. For any measurement of the muscle quality, it is clearly vital to include an accurate measurement of the skeletal muscle size and in this respect the MRI is the gold-standard. The aim of this paper specifically is to develop a method to automatically and accurately determine muscle size from thigh MRI mid-scans. MRI has the additional benefits that it can help in diagnosing muscular pathologies. It is non-invasive, non-ionizing and can also identify fat infiltrations and bone size. Automatic segmentation techniques have been successfully developed and published [15, 16, 17, 18, 19, 20, 21, 22], with some generating excellent segmentation accuracies in a short period of time.

Indeed, establishing automated system for individual muscle segmentations within skeletal muscles or within the same muscle group of the same body segment is a key challenge in medical imaging. It is technically difficult to achieve due to poor contrast between connective and muscle tissues, and sometimes none or little information of muscle borders to assist the segmentation, even within the same muscle group. With this in mind, the most commonly used approach is to manually segment the regions of interest (ROI) and this requires a very good knowledge of human anatomy. This process is operator-dependent, tedious, time-consuming and can have questionable reproducibility by novice users. The issue holding back development of automatic segmentation of individual muscles from MR images is that individual muscles share similar MR properties and can appear indistinguish-

C.2 Other(s)

C.2.1 Journal of Software: Practice and Experience

Page 1 of 25

Software: Practice and Experience

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

SOFTWARE—PRACTICE AND EXPERIENCE
Softw. Pract. Exper. 2010; **00**:1–25
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/spe

Formulating Efficient Software Solution for Digital Image Processing System

Thomas Sherwood, Ezak Ahmad, Moi Hoon Yap*

School of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, Chester Street, Manchester, M1 5GD, UK

SUMMARY

Digital Image Processing Systems are complex, being usually composed of different computer vision libraries. Algorithm implementations cannot be directly used in conjunction with other algorithms developed using other computer vision libraries. This paper formulate a software solution by proposing a processor with the capability of handling different types of image processing algorithms, which allow the end-users to install new image processing algorithms from any library. This approach has other functionalities like capability to process one or more images; manage multiple processing jobs simultaneously; and maintain the manner in which an image was processed for later use. It is a computational efficient and promising technique to handle variety image processing algorithms. To promote the reusability and adaptation of the package for new types of analysis, a feature of sustainability is established. The system past the testing procedures by using unit testing, integration testing and usability testing. Future work involves introducing the capability to connect to another instance of processing service with better performance. Copyright © 2010 John Wiley & Sons, Ltd.

Received ...

KEY WORDS: image processing, software solution, sustainability, plugin, OpenCV, Matlab

1. INTRODUCTION

Digital image processing is complex and inconsistency due to various programming languages and variation computer vision libraries. The domain of image processing has increased vastly in recent years [1], spanning across a range of applications such as photography, forensics and medical imaging [2]. The term simply relates to the process (or set of processes) applied to the detector and dataset of a radiograph exposure [3]. Motivations for processing an image stem from not only the amount of information perceived as image form, but also for autonomous machine control [4].

A mechanism for implementing the algorithms is required, in order to provide a means to perform the transformations. Larkins et. al. [5] discuss an existing high-level toolbox known as Matlab, providing a plethora of existing algorithms and components for re-use in building more complex algorithms. They highlight how Matlab is easy enough for novice users to grasp while still providing powerful processing and data crunching capabilities. However algorithm implementations cannot be directly used in conjunction with other algorithms developed using other technologies, for instance C++ processing classes. Culjak et. al.[6] discusses an alternative to Matlab known as OpenCV, which provides a suite of processing algorithms and assistant classes written in C. They discuss how the library is also widely used, providing heavily optimised solutions to particular algorithms. A C++ wrapper is available for OpenCV, allowing for easy integration into higher-level languages.

*Correspondence to: School of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, Chester Street, Manchester, M1 5GD, UK.

C.2.2 SPIE 2018 - Paper 1

Capillary Detection in Transverse Muscle Sections

Ahmad Nadim Baharum^a, Moi Hoon Yap^a, Glenn Ferris^b, Ezak Ahmad Fadzrin^a, and Hans Degens^b

^aSchool of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, Chester Street, Manchester, UK

^bSchool of Healthcare Sciences, Manchester Metropolitan University, Chester Street, Manchester, UK

ABSTRACT

Manual identification of capillaries in transverse muscle sections is laborious and time consuming. Although the process of classifying a structure as a capillary is facilitated by (immuno)histochemical staining methods, human judgement is still required in a significant number of cases. This is mainly due to the fact that not all capillaries stain as strongly: they may have an elongated appearance and/or there may be staining artefacts that would lead to a false identification of a capillary. Here we propose two automated methods of capillary detection: a novel image processing approach and an existing machine learning approach that has been previously used to detect nuclei-shaped objects. The robustness of the proposed methods was tested on two sets of differently stained muscle sections. On average, the image processing approach scored a True Positive Rate of 0.817 and a harmonic mean (F1 measure) of 0.804 whilst the machine learning approach scored a True Positive Rate 0.843 and F1 measure of 0.846. Both proposed methods are thus able to mimic most of the manual capillary detection, but further improvements are required for practical applications.

Keywords: histology, muscles, capillary detection, machine learning, image processing

1. INTRODUCTION

Muscles are indispensable for locomotion and maintenance of posture in most animals. The muscle performs its function by the sliding of myosin and actin filaments along each other thereby shortening and producing force at the same time, resulting in motion or maintenance of body posture. This process requires adenosine triphosphate (ATP) that can be generated both anaerobically and aerobically. The latter process generates more ATP per glucose moiety than anaerobic metabolism and also can be used to generate ATP from fatty acids. The aerobic energy metabolism requires an adequate supply of oxygen and energy substrates that is derived from the capillary blood and it is indeed a common observation that the capillary network is denser in muscle with a large aerobic potential than those with a low aerobic potential [1–4].

The exploration of the human body has a long history. One of the first recorded instances of human dissection was performed by the ancient Egyptians to unravel the anatomy and understand how the human body works [6]. With the advancements in screening technology, many techniques have been developed [7] that opened up ways to study anatomy even at the microscopical level. One such area that has been studied with microscopy is the micro-circulation and initially capillaries were identified by injecting Indian ink into the circulation [1]. Recent work analysing the capillary network in muscle tissues used staining protocols to reveal capillaries [8–10], where the subsequent identification of capillaries and muscle fibres is performed manually [2, 3]. More recent work by Ballak et al. [11] has shown that at least some of the analysis can be automated. However, the identification of muscle fibres and capillaries was still done manually.

It is highly desirable to automate the capillary identification and therefore, this paper evaluates the efficacy of two methods of automated capillary detection on two image datasets; a novel image processing method and a machine learning method [12].

Further author information: (Send correspondence to M.H. Yap)
M.H. Yap.: E-mail: m.yap@mmu.ac.uk, Telephone: +44 161 247 1503

C.2.3 SPIE 2018 - Paper 2

End-to-End Breast Ultrasound Lesions Recognition with a Deep Learning Approach

Moi Hoon Yap^a, Manu Goyal^a, Fatima Osman^b, Ezak Ahmad^a, Reyer Zwiggelaar^c, Robert Martí^d, Erika Denton^e, and Arne Juetten^e

^aManchester Metropolitan University, Chester Street, Manchester, UK

^bSudan University of Science and Technology, Khartoum, Sudan

^cAberystwyth University, Aberystwyth, UK

^dUniversity of Girona, Girona, Spain

^eNorwich and Norfolk University Hospital, Norwich, UK

ABSTRACT

The existing methods for breast ultrasound lesions detection and recognition are based on multi-stage processing, such as preprocessing, filtering/denoising, segmentation and classification. The performance of these processes dependent on the prior stages. To improve the current state of the art, we proposed an end-to-end breast ultrasound lesions detection and recognition using a deep learning approach. We implemented a popular semantic segmentation framework, i.e. Fully Convolutional Network (FCN-AlexNet) for our experiment. To overcome data deficiency, we used the pre-trained model from ImageNet for the transfer learning FCN-AlexNet. We validated our results on two datasets, which consists of a total of 113 malignant lesions and 356 benign lesions. We assessed the performance of the model using the following split: 70% for training data, 10% for validation data, and 20% testing data. The results shown that our proposed method performed better on benign lesions, with a *Dice* score of 0.6879, when compared to the malignant lesions with a *Dice* score of 0.5525. When considering the number of images with *Dice* score > 0.5 , 79.45% of the benign lesions were successfully segmented and correctly recognised, but only 65% of the malignant lesions were successfully segmented and correctly recognised. This paper provides the first end-to-end solution for breast ultrasound lesions recognition. The future challenges for this proposal is to obtain more datasets and customize the deep learning framework to improve the results.

Keywords: breast ultrasound lesions, breast cancer detection, fully convolutional network, AlexNet

1. INTRODUCTION

According to Breast Cancer Care [1], breast cancer is the most common cancer in the UK. One in eight women will be diagnosed with breast cancer in their lifetime and one person is diagnosed every 10 minutes [1]. Over the past years, many research into using different image modalities [2] and technical methods were developed [3, 4] to aid early diagnosis of the disease. These efforts has led to further research challenge and demand for robust computerised method.

Mammography is known as the gold standard for breast cancer diagnosis [2]. However, ultrasound is an important complementary modality to increase the accuracy of the diagnosis and for those who are vulnerable even to a tiny radiation exposure. Other alternatives are tomography or magnetic resonance. However, ultrasound is the cheapest option and widely used in clinical practice [5].

Conventional computerised methods on breast ultrasound cancer diagnosis composed of multi-stages, including pre-processing, detection of region of interest (ROI), segmentation and classification [6–8]. These processes rely on hand-crafted features including descriptions in spatial domain (texture information, shape and edge descriptors) and frequency domain. With the advancement in deep learning methods, we can detect and recognise objects without hand-crafted features. This paper presents the limitation of the state of the art and conducts a feasibility study on the use of a deep learning approach as an end-to-end solution for fully automated breast lesions recognition.

Further author information: (Send correspondence to M.H. Yap)
M.H.Yap.: E-mail: m.yap@mmu.ac.uk, Telephone: +44 161 247 1503

Bibliography

- [1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [2] Alberto Garcia-Garcia, Sergio Orts-Escolano, Sergiu Oprea, Victor Villena-Martinez, and Jose Garcia-Rodriguez. A review on deep learning techniques applied to semantic segmentation. *arXiv preprint arXiv:1704.06857*, 2017.
- [3] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network. *arXiv preprint arXiv:1612.01105*, 2016.
- [4] Paul Blakey and Twin Eagles Educational. *The muscle book*. Bibliotek Books, 1992.
- [5] JAMES F Sallis. Age-related decline in physical activity: a synthesis of human and animal studies. *Medicine and science in sports and exercise*, 32(9):1598–1600, 2000.
- [6] H Degens. The role of systemic inflammation in age-related muscle weakness and wasting. *Scandinavian journal of medicine & science in sports*, 20(1):28–38, 2010.
- [7] Lorena Urricelqui, Armando Malanda, and Arantxa Villanueva. Automatic segmentation of thigh magnetic resonance images. *World Academy of Science, Engineering and Technology*, 34, 2009.
- [8] Vincenzo Positano, Tore Christiansen, Maria Filomena Santarelli, Steffen Ringgaard, Luigi Landini, and Amalia Gastaldelli. Accurate segmentation of subcutaneous and intermuscular adipose tissue from mr images of the thigh. *Journal of Magnetic Resonance Imaging*, 29(3):677–684, 2009.

- [9] Jeffrey W Prescott, Thomas M Best, Mark S Swanson, Furqan Haq, Rebecca D Jackson, and Metin N Gurcan. Anatomically anchored template-based level set segmentation: application to quadriceps muscles in mr images from the osteoarthritis initiative. *Journal of digital imaging*, 24(1):28–43, 2011.
- [10] F Picard and L Guarente. Molecular links between aging and adipose tissue. *International Journal of Obesity*, 29:S36–S39, 2005.
- [11] Shawn Andrews, Ghassan Hamarneh, Azadeh Yazdanpanah, Bahareh HajGhanbari, and W Darlene Reid. Probabilistic multi-shape segmentation of knee extensor and flexor muscles. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*, pages 651–658. Springer, 2011.
- [12] Matthew Morrison and Y Attikiouzel. A probabilistic neural network based image segmentation network for magnetic resonance images. In *Neural Networks, 1992. IJCNN., International Joint Conference on*, volume 3, pages 60–65. IEEE, 1992.
- [13] P-Y Baudin, Noura Azzabou, Pierre G Carlier, and Nikos Paragios. Automatic skeletal muscle segmentation through random walks and graph-based seed placement. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 1036–1039. IEEE, 2012.
- [14] E Jeffrey Metter, Laura A Talbot, Matthew Schrager, and Robin Conwit. Skeletal muscle strength as a predictor of all-cause mortality in healthy men. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 57(10):B359–B365, 2002.
- [15] Walter R Frontera, Kieran F Reid, Edward M Phillips, Lisa S Krivickas, Virginia A Hughes, Ronenn Roubenoff, and Roger A Fielding. Muscle fiber size and function in elderly humans: a longitudinal study. *Journal of Applied Physiology*, 105(2):637–642, 2008.
- [16] Ryota Akagi, Yohei Takai, Megumi Ohta, Hiroaki Kanehisa, Yasuo Kawakami, and Tetsuo Fukunaga. Muscle volume compared to cross-sectional area is more appropriate for evaluating muscle strength in young and elderly individuals. *Age and ageing*, 38(5):564–569, 2009.

- [17] S Barbat-Artigas, Y Rolland, M Zamboni, and Mylène Aubertin-Leheudre. How to assess functional status: a new muscle quality index. *The journal of nutrition, health & aging*, 16(1):67–77, 2012.
- [18] Lei Chen, David R Nelson, Yang Zhao, Zhanglin Cui, and Joseph A Johnston. Relationship between muscle mass and muscle strength, and the impact of comorbidities: a population-based, cross-sectional study of older adults in the united states. *BMC geriatrics*, 13(1):1, 2013.
- [19] Joseph C Dunn. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. 1973.
- [20] Yanling Li and Yi Shen. Fuzzy c-means clustering based on spatial neighborhood information for image segmentation. *Systems Engineering and Electronics, Journal of*, 21(2):323–328, 2010.
- [21] Zhang Yang, Fu-Lai Chung, and Wang Shitong. Robust fuzzy clustering-based image segmentation. *Applied Soft Computing*, 9(1):80–84, 2009.
- [22] Yanling Li and Yi Shen. Robust image segmentation algorithm using fuzzy clustering based on kernel-induced distance measure. In *Computer Science and Software Engineering, 2008 International Conference on*, volume 1, pages 1065–1068. IEEE, 2008.
- [23] Zhang Xue-xi and Yang Yi-min. Hybrid intelligent algorithms for color image segmentation. In *Control and Decision Conference, 2008. CCDC 2008. Chinese*, pages 264–268. IEEE, 2008.
- [24] Morgan McGuire. An image registration technique for recovering rotation, scale and translation parameters. *NEC Res. Inst. Tech. Rep., TR*, pages 98–018, 1998.
- [25] James P Boyle, Theodore J Thompson, Edward W Gregg, Lawrence E Barker, and David F Williamson. Projection of the year 2050 burden of diabetes in the us adult population: dynamic modeling of incidence, mortality, and prediabetes prevalence. *Population health metrics*, 8(1):29, 2010.
- [26] Bernd Rechel, Emily Grundy, Jean-Marie Robine, Jonathan Cylus, Johan P Mackenbach, Cecile Knai, and Martin McKee. Ageing in the european union. *The Lancet*, 381(9874):1312–1322, 2013.

- [27] Piotr Juchno Silvia Andueza Robustillo, Veronica Corsini and Monica Marcu. Short analytical web note demography report 2015 edition, July 2015. URL <http://ec.europa.eu/eurostat/en/web/products-statistical-books/-/KE-BM-15-003>.
- [28] Steven Dunstan. General lifestyle survey overview: a report on the 2010 general lifestyle survey. *Newport: Office for National Statistics*, 2012.
- [29] Valiollah Hajhashemi, Alireza Ghannadi, and Hadi Jafarabadi. Black cumin seed essential oil, as a potent analgesic and antiinflammatory drug. *Phytotherapy Research*, 18(3):195–199, 2004.
- [30] Stuart Ira Fox. Human physiology 9th editon. 1996.
- [31] RH Edwards, DK Hill, and DA Jones. Heat production and chemical changes during isometric contractions of the human quadriceps muscle. *The Journal of physiology*, 251(2):303–315, 1975.
- [32] Elaine N Marieb, Katja Hoehn, Patricia Brady Wilhelm, and Nina Zanetti. Human anatomy & physiology: International edition with human anatomy and physiology atlas, 7/e, 2006.
- [33] Gf Burnstock. Structure of smooth muscle and its innervation. *Smooth muscle*, pages 1–69, 1970.
- [34] Kevin T Patton and Gary A Thibodeau. *Anatomy & physiology*. Elsevier Health Sciences, 2014.
- [35] Rebecca L Johnson. *The Muscular System*. Lerner Publications, 2004.
- [36] Keith L Moore, Arthur F Dalley, and Anne MR Agur. *Clinically oriented anatomy*. Lippincott Williams & Wilkins, 2013.
- [37] Donna Lind Infeld. *Disciplinary Approaches to Aging: Biology of aging*. Taylor & Francis, 2002.
- [38] KL Minaker. Common clinical sequelae of aging. *Goldman's Cecil Medicine. 24th ed. Philadelphia, PA: Elsevier Saunders*, 2011.
- [39] Carrie A Laughton, Mary Slavin, Kunal Katdare, Lee Nolan, Jonathan F Bean, D Casey Kerrigan, Edward Phillips, Lewis A Lipsitz, and James J Collins. Aging,

- muscle activity, and balance control: physiologic changes associated with balance impairment. *Gait & posture*, 18(2):101–108, 2003.
- [40] Thomas BL Kirkwood. Understanding the odd science of aging. *Cell*, 120(4):437–447, 2005.
- [41] Randye G Jacobson and Franklin P Flowers. Skin changes with aging and disease. *Wound Repair and Regeneration*, 4(3):311–315, 1996.
- [42] Alastair JJ Wood and Richard Eastell. Treatment of postmenopausal osteoporosis. *New England journal of medicine*, 338(11):736–746, 1998.
- [43] Howard W Stoudt. The anthropometry of the elderly. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 23(1):29–37, 1981.
- [44] Yvonne D’Arcy. Is low back pain getting on your nerves? *The Nurse Practitioner*, 34(5):10–17, 2009.
- [45] Ian Janssen, Steven B Heymsfield, ZiMian Wang, and Robert Ross. Skeletal muscle mass and distribution in 468 men and women aged 18–88 yr. *Journal of applied physiology*, 89(1):81–88, 2000.
- [46] Walter R Frontera, VIRGINIA A Hughes, Karyn J Lutz, and William J Evans. A cross-sectional study of muscle strength and mass in 45-to 78-yr-old men and women. *Journal of applied physiology*, 71(2):644–650, 1991.
- [47] K Sreekumaran Nair. Aging muscle. *The American journal of clinical nutrition*, 81(5):953–963, 2005.
- [48] Hirofumi Tanaka and Douglas R Seals. Endurance exercise performance in masters athletes: age-associated changes and underlying physiological mechanisms. *The Journal of physiology*, 586(1):55–63, 2008.
- [49] David Westaway, Stephen J DeArmond, Juliana Cayetano-Canlas, Darlene Groth, Dallas Foster, Shu-Lian Yang, Marilyn Torchia, George A Carlson, and Stanley B Prusiner. Degeneration of skeletal muscle, peripheral nerves, and the central nervous system in transgenic mice overexpressing wild-type prion proteins. *Cell*, 76(1):117–129, 1994.

- [50] Jerrold T Bushberg and John M Boone. *The essential physics of medical imaging*. Lippincott Williams & Wilkins, 2011.
- [51] KMM Rao and VDP Rao. Medical image processing. In *proc. of workshop on Medical Image Processing and Applications*, volume 8, 2006.
- [52] Charles J Pavlin, Kasia Harasiewicz, Michael D Sherar, and F Stuart Foster. Clinical use of ultrasound biomicroscopy. *Ophthalmology*, 98(3):287–295, 1991.
- [53] Timothy G Leighton. What is ultrasound? *Progress in biophysics and molecular biology*, 93(1):3–83, 2007.
- [54] Marek J Maryanski, John C Gore, Richard P Kennan, and Robert J Schulz. Nmr relaxation enhancement in gels polymerized and cross-linked by ionizing radiation: a new approach to 3d dosimetry by mri. *Magnetic resonance imaging*, 11(2):253–258, 1993.
- [55] Katsuhisa Omori and Masaya Yamashita. Rf coil positioning device for mri, November 19 1991. US Patent 5,066,915.
- [56] WM Faulkner. Basic principles of mri. *Outsource Inc*, 1996.
- [57] Barry LW Chapman. Gradients: The heart of the mri machine. *Current Medical Imaging Reviews*, 2(1):131–138, 2006.
- [58] Michel M Ter-Pogossian. Positron emission tomography (pet). In *Diagnostic Imaging in Medicine*, pages 273–277. Springer, 1983.
- [59] Anja Wagner, Heiko Mahrholdt, Thomas A Holly, Michael D Elliott, Matthias Regenfus, Michele Parker, Francis J Klocke, Robert O Bonow, Raymond J Kim, and Robert M Judd. Contrast-enhanced mri and routine single photon emission computed tomography (spect) perfusion imaging for detection of subendocardial myocardial infarcts: an imaging study. *The Lancet*, 361(9355):374–379, 2003.
- [60] National cancer institute. URL <http://www.cancer.gov/types/breast/mammograms-fact-sheet>.
- [61] Claude R Brice and Claude L Fennema. Scene analysis using regions. *Artificial intelligence*, 1(3):205–226, 1970.

- [62] Theodosios Pavlidis. Segmentation of pictures and maps through functional approximation. *Computer Graphics and Image Processing*, 1(4):360–372, 1972.
- [63] A Rosenfeld and AC Kak. Enhancement. *Digital Picture Processing*. New York, NY, Academic, pages 173–175, 1976.
- [64] Rafael C Gonzalez and Paul Wintz. Image enhancement. *Digital image processing*, pages 161–201, 1977.
- [65] H William Strauss, Barry L Zaret, Peter J Hurley, TK Natarajan, and Bertram Pitt. A scintiphotographic method for measuring left ventricular ejection fraction in man without cardiac catheterization. *The American journal of cardiology*, 28(5):575–580, 1971.
- [66] D Ballard and Jack Sklansky. Tumor detection in radiographs. *Computers and Biomedical Research*, 6(4):299–321, 1973.
- [67] James S Duncan and Nicholas Ayache. Medical image analysis: Progress over two decades and the challenges ahead. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(1):85–106, 2000.
- [68] Vincent Barra and Jean-Yves Boire. Segmentation of fat and muscle from mr images of the thigh by a possibilistic clustering algorithm. *Computer methods and programs in biomedicine*, 68(3):185–193, 2002.
- [69] JP Mattei, Y Le Fur, N Cuge, S Guis, PJ Cozzone, and D Bendahan. Segmentation of fascias, fat and muscle from magnetic resonance images in humans: the dispimag software. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 19(5): 275–279, 2006.
- [70] Han Kang, Antonio Pinti, Laurent Vermeiren, Abdelmalik Taleb-Ahmed, and Xiyani Zeng. An automatic fcm-based method for tissue classification application to mri of thigh. In *Bioinformatics and Biomedical Engineering, 2007. ICBBE 2007. The 1st International Conference on*, pages 510–514. IEEE, 2007.
- [71] Hyoungku Kang, Antonio Pinti, L Vermeiren, A Taleb-Ahmed, and X Zeng. Tissue classification for mri of thigh using a modified fcm method. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 5579–5584. IEEE, 2007.

- [72] Changming Wang, Xiaojuan Guo, Li Yao, Ke Li, and Zhen Jin. A practical method for muscles extraction and automatic segmentation of leg magnetic resonance images. In *Complex Medical Engineering, 2007. CME 2007. IEEE/ICME International Conference on*, pages 885–890. IEEE, 2007.
- [73] Peijie Jiang, Chunlan Yang, Shuicai Wu, and Xiangyan Ruan. The research of tissue segmentation method on thigh mri. In *Multimedia Technology (ICMT), 2010 International Conference on*, pages 1–3. IEEE, 2010.
- [74] Sokratis Makrogiannis, Suraj Serai, Kenneth W Fishbein, Catherine Schreiber, Luigi Ferrucci, and Richard G Spencer. Automated quantification of muscle and fat in the thigh from water-, fat-, and nonsuppressed mr images. *Journal of Magnetic Resonance Imaging*, 35(5):1152–1161, 2012.
- [75] Senthil Purushwalkam, Baihua Li, Qinggang Meng, and Jamie McPhee. Automatic segmentation of adipose tissue from thigh magnetic resonance images. In *Image Analysis and Recognition*, pages 451–458. Springer, 2013.
- [76] Alexander Valentinitich, Dimitrios C Karampinos, Hamza Alizai, Karupppasamy Subburaj, Deepak Kumar, Thomas M Link, and Sharmila Majumdar. Automated unsupervised multi-parametric classification of adipose tissue depots in skeletal muscle. *Journal of Magnetic Resonance Imaging*, 37(4):917–927, 2013.
- [77] Sara Orgiu, Claudio L Lafortuna, Fabio Rastelli, Marcello Cadioli, Andrea Falini, and Giovanna Rizzo. Automatic muscle and fat segmentation in the thigh from t1-weighted mri. *Journal of Magnetic Resonance Imaging*, 2015.
- [78] Pierre-Yves Baudin, Noura Azzabou, Pierre G Carlier, Nikos Paragios, et al. Manifold-enhanced segmentation through random walks on linear subspace priors. In *Proceedings of the British Machine Vision Conference*, 2012.
- [79] P-Y Baudin, Noura Azzabou, Pierre G Carlier, and Nikos Paragios. Prior knowledge, random walks and human skeletal muscle segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pages 569–576. Springer, 2012.
- [80] Simon Andrews and Ghassan Hamarneh. The generalized log-ratio transformation: Learning shape and adjacency priors for simultaneous thigh muscle segmentation. 2015.

- [81] Pim Moeskops, Max A Viergever, Adriënne M Mendrik, Linda S de Vries, Manon JNL Benders, and Ivana Išgum. Automatic segmentation of mr brain images with a convolutional neural network. *IEEE transactions on medical imaging*, 35(5):1252–1261, 2016.
- [82] Sérgio Pereira, Adriano Pinto, Victor Alves, and Carlos A Silva. Brain tumor segmentation using convolutional neural networks in mri images. *IEEE transactions on medical imaging*, 35(5):1240–1251, 2016.
- [83] Phi Vu Tran. A fully convolutional neural network for cardiac segmentation in short-axis mri. *arXiv preprint arXiv:1604.00494*, 2016.
- [84] Raghav Mehta and Jayanthi Sivaswamy. M-net: A convolutional neural network for deep brain structure segmentation. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*, pages 437–440. IEEE, 2017.
- [85] Moi Hoon Yap, Gerard Pons, Joan Martí, Sergi Ganau, Melcior Sentís, Reyer Zwiggelaar, Adrian K Davison, and Robert Martí. Automated breast ultrasound lesions detection using convolutional neural networks. *IEEE Journal of Biomedical and Health Informatics*, 2017.
- [86] Shu Liao, Yaozong Gao, Aytakin Oto, and Dinggang Shen. Representation learning: a unified deep learning framework for automatic prostate mr segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 254–261. Springer, 2013.
- [87] Michiel Kallenberg, Kersten Petersen, Mads Nielsen, Andrew Y Ng, Pengfei Diao, Christian Igel, Celine M Vachon, Katharina Holland, Rikke Rass Winkel, Nico Karssemeijer, et al. Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring. *IEEE transactions on medical imaging*, 35(5):1322–1331, 2016.
- [88] Christian F Baumgartner, Lisa M Koch, Marc Pollefeys, and Ender Konukoglu. An exploration of 2d and 3d deep learning techniques for cardiac mr image segmentation. *arXiv preprint arXiv:1709.04496*, 2017.
- [89] Stefano Trebeschi, Joost JM van Griethuysen, Doenja MJ Lambregts, Max J Lahaye, Chintan Parmer, Frans CH Bakers, Nicky HGM Peters, Regina GH Beets-Tan, and Hugo JWL Aerts. Deep learning for fully-automated localization and

- segmentation of rectal cancer on multiparametric mr. *Scientific Reports*, 7(1):5301, 2017.
- [90] MR Avendi, Arash Kheradvar, and Hamid Jafarkhani. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac mri. *Medical image analysis*, 30:108–119, 2016.
- [91] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331, 1988.
- [92] Wei-Ying Ma and Bangalore S Manjunath. Edgeflow: a technique for boundary detection and image segmentation. *Image Processing, IEEE Transactions on*, 9(8):1375–1388, 2000.
- [93] Satyanad Kichenassamy, Arun Kumar, Peter Olver, Allen Tannenbaum, and Anthony Yezzi. Gradient flows and geometric active contour models. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pages 810–815. IEEE, 1995.
- [94] Vicent Caselles, Ron Kimmel, and Guillermo Sapiro. Geodesic active contours. *International journal of computer vision*, 22(1):61–79, 1997.
- [95] Stanley Osher and James A Sethian. Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. *Journal of computational physics*, 79(1):12–49, 1988.
- [96] James Albert Sethian. *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, volume 3. Cambridge university press, 1999.
- [97] Nikos Paragios and Rachid Deriche. Geodesic active regions for supervised texture segmentation. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 926–932. IEEE, 1999.
- [98] Liana M Lorigo, Olivier Faugeras, W Eric L Grimson, Renaud Keriven, Ron Kikinis, and Carl-Fredrik Westin. Co-dimension 2 geodesic active contours for mra segmentation. In *Information Processing in Medical Imaging*, pages 126–139. Springer, 1999.

- [99] Yanlin Guo and Baba C Vemuri. Hybrid geometric active models for shape recovery in medical images. In *Information Processing in Medical Imaging*, pages 112–125. Springer, 1999.
- [100] Anthony Yezzi Jr, Andy Tsai, and Alan Willsky. A statistical approach to snakes for bimodal and trimodal imagery. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 898–903. IEEE, 1999.
- [101] Timothy F Cootes, C Beeston, Gareth J Edwards, and Christopher J Taylor. A unified framework for atlas matching using active appearance models. In *Information Processing in Medical Imaging*, pages 322–333. Springer, 1999.
- [102] Yongmei Wang and Lawrence H Staib. Boundary finding with correspondence using statistical shape models. In *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on*, pages 338–345. IEEE, 1998.
- [103] Lawrence H Staib and James S Duncan. Boundary finding with parametrically deformable models. *IEEE transactions on pattern analysis and machine intelligence*, 14(11):1061–1075, 1992.
- [104] Benjamin Gilles and Dinesh K Pai. Fast musculoskeletal registration based on shape matching. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*, pages 822–829. Springer, 2008.
- [105] Prasanna K Sahoo, SAKC Soltani, and Andrew KC Wong. A survey of thresholding techniques. *Computer vision, graphics, and image processing*, 41(2):233–260, 1988.
- [106] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.
- [107] Larry S Davis. A survey of edge detection techniques. *Computer graphics and image processing*, 4(3):248–270, 1975.
- [108] David Marr and Ellen Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London B: Biological Sciences*, 207(1167):187–217, 1980.

- [109] John Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):679–698, 1986.
- [110] Jianping Fan, Walid G Aref, Mohand-Said Hacid, and Ahmed K Elmagarmid. An improved automatic isotropic color edge detection technique. *Pattern Recognition Letters*, 22(13):1419–1429, 2001.
- [111] Tamar Peli and David Malah. A study of edge detection algorithms. *Computer graphics and image processing*, 20(1):1–21, 1982.
- [112] Irwin Sobel and Gary Feldman. A 3x3 isotropic gradient operator for image processing. 1968.
- [113] Xin Wang. Laplacian operator-based edge detectors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(5):886–890, 2007.
- [114] GT Shrivakshan and C Chandrasekar. A comparison of various edge detection techniques used in image processing. *IJCSI International Journal of Computer Science Issues*, 9(5):272–276, 2012.
- [115] Rishi R Rakesh, Probal Chaudhuri, and CA Murthy. Thresholding in edge detection: a statistical approach. *Image Processing, IEEE Transactions on*, 13(7):927–936, 2004.
- [116] Rolf Adams and Leanne Bischof. Seeded region growing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(6):641–647, 1994.
- [117] P Gibbs, DL Buckley, SJ Blackband, and A Horsman. Tumour volume determination from mr images by morphological segmentation. *Physics in medicine and biology*, 41(11):2437, 1996.
- [118] Arnau Oliver, Xavier Munoz, Joan Batlle, Lluís Pacheco, and Jordi Freixenet. Improving clustering algorithms for image segmentation using contour and region information. In *Automation, Quality and Testing, Robotics, 2006 IEEE International Conference on*, volume 2, pages 315–320. IEEE, 2006.
- [119] Guy B Coleman and Harry C Andrews. Image segmentation by clustering. *Proceedings of the IEEE*, 67(5):773–785, 1979.

- [120] James C Bezdek, LO Hall, and L-P Clarke. Review of mr image segmentation techniques using pattern recognition. *Medical physics*, 20(4):1033–1048, 1992.
- [121] Stephen C Johnson. Hierarchical clustering schemes. *Psychometrika*, 32(3):241–254, 1967.
- [122] Bing Jian and Baba C Vemuri. A robust algorithm for point set registration using mixture of gaussians. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1246–1251. IEEE, 2005.
- [123] H Digabel and Christian Lantuéjoul. Iterative algorithms. In *Proc. 2nd European Symp. Quantitative Analysis of Microstructures in Material Science, Biology and Medicine*, volume 19, page 8. Stuttgart, West Germany: Riederer Verlag, 1978.
- [124] Christian Lantuejoul. *La squelettisation et son application aux mesures topologiques des mosaïques polycristallines*. PhD thesis, 1978.
- [125] Serge Beucher and Christian Lantuéjoul. Use of watersheds in contour detection. In *International workshop on image processing, real-time edge and motion detection*, 1979.
- [126] Alina N Moga and Moncef Gabbouj. Parallel image component labelling with watershed transformation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(5):441–450, 1997.
- [127] AEL Allaoui and Mbarek Nasri. Medical image segmentation by marker-controlled watershed and mathematical morphology. *The Int. J. Multimedia & Its Applications*, 4(3):2–3, 2012.
- [128] S Ravi and AM Khan. Bio-medical image segmentation using marker controlled watershed algorithm: A case study.
- [129] Chen Jia-Xin and Liu Sen. A medical image segmentation method based on watershed transform. In *Computer and Information Technology, 2005. CIT 2005. The Fifth International Conference on*, pages 634–638. IEEE, 2005.
- [130] Vicente Grau, AUJ Mewes, M Alcaniz, Ron Kikinis, and Simon K Warfield. Improved watershed transform for medical image segmentation using prior information. *Medical Imaging, IEEE Transactions on*, 23(4):447–458, 2004.

- [131] Tian Yun, Zhou Ming-quan, Wu Zhong-ke, and Wang Xing-ce. A region-based active contour model for image segmentation. In *Computational Intelligence and Security, 2009. CIS'09. International Conference on*, volume 1, pages 376–380. IEEE, 2009.
- [132] Tim McInerney and Demetri Terzopoulos. Deformable models in medical image analysis: a survey. *Medical image analysis*, 1(2):91–108, 1996.
- [133] Andrew Blake and Andrew Zisserman. *Visual reconstruction*, volume 2. MIT press Cambridge, 1987.
- [134] David Mumford and Jayant Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*, 42(5):577–685, 1989.
- [135] Song Chun Zhu and Alan Yuille. Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(9):884–900, 1996.
- [136] Daniel Cremers, Florian Tischhuser, Joachim Weickert, and Christoph Schnorr. Diffusion snakes: Introducing statistical shape knowledge into the mumford-shah functional. *International journal of computer vision*, 50(3):295–313, 2002.
- [137] Tony F Chan and Luminita A Vese. Active contours without edges. *Image processing, IEEE transactions on*, 10(2):266–277, 2001.
- [138] Shawn Lankton and Allen Tannenbaum. Localizing region-based active contours. *Image Processing, IEEE Transactions on*, 17(11):2029–2039, 2008.
- [139] Alain Dervieux and Franois Thomasset. A finite element method for the simulation of a rayleigh-taylor instability. In *Approximation methods for Navier-Stokes problems*, pages 145–158. Springer, 1980.
- [140] John A Bogovic, Jerry L Prince, and Pierre-Louis Bazin. A multiple object geometric deformable model for image segmentation. *Computer Vision and Image Understanding*, 117(2):145–157, 2013.
- [141] Nikos Paragios. Curve propagation, level set methods and grouping. In *Handbook of Mathematical Models in Computer Vision*, pages 145–159. Springer, 2006.

- [142] Daniel Cremers, Mikael Rousson, and Rachid Deriche. A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. *International journal of computer vision*, 72(2):195–215, 2007.
- [143] Nikos K Paragios and Rachid Deriche. A pde-based level-set approach for detection and tracking of moving objects. In *Computer Vision, 1998. Sixth International Conference on*, pages 1139–1145. IEEE, 1998.
- [144] Hong-Kai Zhao, Tony Chan, Barry Merriman, and Stanley Osher. A variational level set approach to multiphase motion. *Journal of computational physics*, 127(1):179–195, 1996.
- [145] Junmo Kim, John W Fisher III, Anthony Yezzi, Mujdat Cetin, and Alan S Wilksy. Nonparametric methods for image segmentation using information theory and curve evolution. In *Image Processing. 2002. Proceedings. 2002 International Conference on*, volume 3, pages 797–800. IEEE, 2002.
- [146] Ian L Dryden and Kanti V Mardia. *Statistical shape analysis*, volume 4. Wiley Chichester, 1998.
- [147] Christopher G Small. *The statistical theory of shape*. Springer Science & Business Media, 2012.
- [148] Timothy F Cootes, Christopher J Taylor, David H Cooper, and Jim Graham. Active shape models-their training and application. *Computer vision and image understanding*, 61(1):38–59, 1995.
- [149] John G Csernansky, Sarang Joshi, Lei Wang, John W Haller, Mokhtar Gado, J Philip Miller, Ulf Grenander, and Michael I Miller. Hippocampal morphometry in schizophrenia by high dimensional brain mapping. *Proceedings of the National Academy of Sciences*, 95(19):11406–11411, 1998.
- [150] Liang Wang, Tieniu Tan, Weiming Hu, and Huazhong Ning. Automatic gait recognition based on statistical shape analysis. *Image Processing, IEEE Transactions on*, 12(9):1120–1131, 2003.
- [151] Tobias Heimann and Hans-Peter Meinzer. Statistical shape models for 3d medical image segmentation: a review. *Medical image analysis*, 13(4):543–563, 2009.

- [152] Timothy F Cootes, Gareth J Edwards, and Christopher J Taylor. Active appearance models. In *Computer Vision ECCV98*, pages 484–498. Springer, 1998.
- [153] Lassi Hyvärinen. Principal component analysis. In *Mathematical Modeling for Industrial Processes*, pages 82–104. Springer, 1970.
- [154] Bruce C Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *Automatic Control, IEEE Transactions on*, 26(1):17–32, 1981.
- [155] M Mudrova and A Prochazka. Principal component analysis in image processing. In *Proceedings of the MATLAB Technical Computing Conference, Prague*, 2005.
- [156] Mark Richardson. Principal component analysis. URL: <http://people.maths.ox.ac.uk/richardsonm/SignalProcPCA.pdf> (last access: 3.5. 2013). Aleš Hladnik Dr., Ass. Prof., Chair of Information and Graphic Arts Technology, Faculty of Natural Sciences and Engineering, University of Ljubljana, Slovenia ales.hladnik@ntf.uni-lj.si, 2009.
- [157] Jonathon Shlens. A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*, 2014.
- [158] Bruce D Lucas, Takeo Kanade, et al. An iterative image registration technique with an application to stereo vision. In *IJCAI*, volume 81, pages 674–679, 1981.
- [159] Frederik Maes, Andre Collignon, Dirk Vandermeulen, Guy Marchal, and Paul Suetens. Multimodality image registration by maximization of mutual information. *Medical Imaging, IEEE Transactions on*, 16(2):187–198, 1997.
- [160] JB Antoine Maintz and Max A Viergever. A survey of medical image registration. *Medical image analysis*, 2(1):1–36, 1998.
- [161] Hrvoje Kalinic. Atlas-based image segmentation: A survey. *Department of Electronic Systems and Information Processing, University of Zagreb*, 2008.
- [162] Daniel Rueckert, Luke I Sonoda, Carmel Hayes, Derek LG Hill, Martin O Leach, and David J Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *Medical Imaging, IEEE Transactions on*, 18(8):712–721, 1999.

- [163] Alexis Roche, Grégoire Malandain, Xavier Pennec, and Nicholas Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *Medical Image Computing and Computer-Assisted Intervention MICCAI98*, pages 1115–1124. Springer, 1998.
- [164] Josien PW Pluim, JB Antoine Maintz, Max Viergever, et al. Mutual-information-based registration of medical images: a survey. *Medical Imaging, IEEE Transactions on*, 22(8):986–1004, 2003.
- [165] Padhraic Smyth. Belief networks, hidden markov models, and markov random fields: a unifying view. *Pattern recognition letters*, 18(11):1261–1268, 1997.
- [166] Karsten Held, Elena Rota Kops, Bernd J Krause, William M Wells III, Ron Kikinis, and Hans-Wilhelm Muller-Gartner. Markov random field segmentation of brain mr images. *Medical Imaging, IEEE Transactions on*, 16(6):878–886, 1997.
- [167] Jean Talairach and Pierre Tournoux. Co-planar stereotaxic atlas of the human brain. 3-dimensional proportional system: an approach to cerebral imaging. 1988.
- [168] JL Lancaster, LH Rainey, JL Summerlin, CS Freitas, PT Fox, AC Evans, AW Toga, and JC Mazziotta. Automated labeling of the human brain: a preliminary report on the development and evaluation of a forward-transform method. *Human brain mapping*, 5(4):238, 1997.
- [169] Nancy C Andreasen, R Rajarethinam, Theodore Cizadlo, Stephan Arndt, Victor W Swayze, Laura A Flashman, Daniel S O’Leary, James C Ehrhardt, William TC Yuh, et al. Automatic atlas-based volume estimation of human brain regions from mr images. *Journal of computer assisted tomography*, 20(1):98–106, 1996.
- [170] D Louis Collins, Colin J Holmes, Terrence M Peters, and Alan C Evans. Automatic 3-d model-based neuroanatomical segmentation. *Human brain mapping*, 3(3):190–208, 1995.
- [171] Christos Davatzikos. Spatial normalization of 3d brain images using deformable models. *Journal of computer assisted tomography*, 20(4):656–665, 1996.

- [172] Gary E Christensen, Sarang C Joshi, Michael Miller, et al. Volumetric transformation of brain anatomy. *Medical Imaging, IEEE Transactions on*, 16(6):864–877, 1997.
- [173] Georges B Aboutanos and Benoit M Dawant. Automatic brain segmentation and validation: image-based versus atlas-based deformable models. In *Medical Imaging 1997*, pages 299–310. International Society for Optics and Photonics, 1997.
- [174] Fei Wang, Baba C Vemuri, and Stephan J Eisenschenk. Joint registration and segmentation of neuroanatomic structures from brain mri. *Academic radiology*, 13(9):1104–1111, 2006.
- [175] Lilla Zöllei, Martha Shenton, William Wells, and Kilian Pohl. The impact of atlas formation methods on atlas-guided brain segmentation. In *Proceedings of Medical image computing and computer-assisted intervention: MICCAI International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 39–46, 2007.
- [176] Paul M Thompson and Arthur W Toga. Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations. *Medical image analysis*, 1(4):271–294, 1997.
- [177] Michael E Leventon, W Eric L Grimson, and Olivier Faugeras. Statistical shape influence in geodesic active contours. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 1, pages 316–323. IEEE, 2000.
- [178] Yunmei Chen, Hemant D Tagare, Sheshadri Thiruvankadam, Feng Huang, David Wilson, Kaundinya S Gopinath, Richard W Briggs, and Edward A Geiser. Using prior shapes in geometric active contours in a variational framework. *International Journal of Computer Vision*, 50(3):315–328, 2002.
- [179] Andy Tsai, Anthony Yezzi Jr, William Wells III, Clare Tempany, Dewey Tucker, Ayres Fan, W Eric Grimson, and Alan Willsky. Model-based curve evolution technique for image segmentation. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–463. IEEE, 2001.
- [180] Andy Tsai, Anthony Yezzi Jr, William Wells, Clare Tempany, Dewey Tucker, Ayres Fan, W Eric Grimson, and Alan Willsky. A shape-based approach to the

- segmentation of medical imagery using level sets. *Medical Imaging, IEEE Transactions on*, 22(2):137–154, 2003.
- [181] Xavier Bresson, Pierre Vandergheynst, and Jean-Philippe Thiran. A variational model for object segmentation using boundary information and shape prior driven by the mumford-shah functional. *International Journal of Computer Vision*, 68(2):145–162, 2006.
- [182] Luminita A Vese and Tony F Chan. A multiphase level set framework for image segmentation using the mumford and shah model. *International journal of computer vision*, 50(3):271–293, 2002.
- [183] Alan L Yuille. Generalized deformable models, statistical physics, and matching problems. *Neural Computation*, 2(1):1–24, 1990.
- [184] Eric Bardinet, Laurent D Cohen, and Nicholas Ayache. Tracking and motion analysis of the left ventricle with deformable superquadrics. *Medical image analysis*, 1(2):129–149, 1996.
- [185] John J Hopfield. Artificial neural networks. *IEEE Circuits and Devices Magazine*, 4(5):3–10, 1988.
- [186] Simon S Haykin, Simon S Haykin, Simon S Haykin, and Simon S Haykin. *Neural networks and learning machines*, volume 3. Pearson Upper Saddle River, NJ, USA., 2009.
- [187] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [188] Nikhil Buduma and Nicholas Locascio. *Fundamentals of Deep Learning: Designing Next-generation Machine Intelligence Algorithms.* ” O’Reilly Media, Inc.”, 2017.
- [189] David E Rumelhart, Geoffrey E Hinton, Ronald J Williams, et al. Learning representations by back-propagating errors. *Cognitive modeling*, 5(3):1, 1988.
- [190] Léon Bottou. Stochastic gradient descent tricks. In *Neural networks: Tricks of the trade*, pages 421–436. Springer, 2012.

- [191] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [192] John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for on-line learning and stochastic optimization. *Journal of Machine Learning Research*, 12(Jul):2121–2159, 2011.
- [193] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM international conference on Multimedia*, pages 675–678. ACM, 2014.
- [194] Matthew D Zeiler. Adadelata: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*, 2012.
- [195] Yurii Nesterov. A method of solving a convex programming problem with convergence rate $o(1/k^2)$. In *Soviet Mathematics Doklady*, volume 27, pages 372–376, 1983.
- [196] G Hinton, N Srivastava, and K Swersky. Rmsprop: Divide the gradient by a running average of its recent magnitude. *Neural networks for machine learning, Coursera lecture 6e*, 2012.
- [197] Alan H Cheetham and Joseph E Hazel. Binary (presence-absence) similarity coefficients. *Journal of Paleontology*, pages 1130–1136, 1969.
- [198] Paul Jaccard. The distribution of the flora in the alpine zone. *New phytologist*, 11(2):37–50, 1912.
- [199] Lee R Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.
- [200] Alex P Zijdenbos, Benoit M Dawant, Richard Margolin, Andrew C Palmer, et al. Morphometric analysis of white matter lesions in mr images: method and validation. *Medical Imaging, IEEE Transactions on*, 13(4):716–724, 1994.
- [201] Helena Chmura Kraemer. Extension of the kappa coefficient. *Biometrics*, pages 207–216, 1980.

- [202] Torsten Rohlfing, Robert Brandt, Randolph Menzel, and Calvin R Maurer Jr. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *NeuroImage*, 21(4):1428–1442, 2004.
- [203] Osirix imaging software, advance open source pacs workstation dicom viewer. URL www.osirix-viewer.com.
- [204] John G Sled, Alex P Zijdenbos, and Alan C Evans. A nonparametric method for automatic correction of intensity nonuniformity in mri data. *Medical Imaging, IEEE Transactions on*, 17(1):87–97, 1998.
- [205] Medical image processing, analysis, and visualization. URL <http://mipav.cit.nih.gov>.
- [206] Antero Koivula, Jyrki Alakuijala, and Osmo Tervonen. Image feature based automatic correction of low-frequency spatial intensity variations in mr images. *Magnetic resonance imaging*, 15(10):1167–1175, 1997.
- [207] Jagath C Rajapakse and Frithjof Kruggel. Segmentation of mr images with intensity inhomogeneities. *Image and Vision Computing*, 16(3):165–180, 1998.
- [208] Bradley Efron. The convex hull of a random set of points. *Biometrika*, 52(3-4):331–343, 1965.
- [209] Ezak Ahmad, Moi Hoon Yap, Hans Degens, and Jamie S McPhee. Atlas-registration based image segmentation of mri human thigh muscles in 3d space. In *SPIE Medical Imaging*, pages 90371L–90371L. International Society for Optics and Photonics, 2014.
- [210] Kelly Rehm, Kirt Schaper, Jon Anderson, Roger Woods, Sarah Stoltzner, and David Rottenberg. Putting our heads together: a consensus approach to brain/non-brain segmentation in t1-weighted mr volumes. *NeuroImage*, 22(3):1262–1270, 2004.
- [211] David W Shattuck, Gautam Prasad, Mubeena Mirza, Katherine L Narr, and Arthur W Toga. Online resource for validation of brain segmentation methods. *NeuroImage*, 45(2):431–439, 2009.

- [212] Simon K Warfield, Kelly H Zou, and William M Wells. Validation of image segmentation and expert quality with an expectation-maximization algorithm. In *Medical Image Computing and Computer-Assisted Intervention MICCAI 2002*, pages 298–306. Springer, 2002.
- [213] J Richard Landis and Gary G Koch. The measurement of observer agreement for categorical data. *biometrics*, pages 159–174, 1977.
- [214] Mehul P Sampat, Zhou Wang, Shalini Gupta, Alan Conrad Bovik, and Mia K Markey. Complex wavelet structural similarity: A new image similarity index. *Image Processing, IEEE Transactions on*, 18(11):2385–2401, 2009.
- [215] Audrey F Manley. Physical activity and health: A report of the surgeon general. Technical report, DTIC Document, 1997.
- [216] Marjolein Visser, Dorly JH Deeg, and Paul Lips. Low vitamin d and high parathyroid hormone levels as determinants of loss of muscle strength and muscle mass (sarcopenia): the longitudinal aging study amsterdam. *The Journal of Clinical Endocrinology & Metabolism*, 88(12):5766–5772, 2003.
- [217] KA1 Stockton, Kerrie Mengersen, Jenny Davida Paratz, D Kandiah, and KL Bennell. Effect of vitamin d supplementation on muscle strength: a systematic review and meta-analysis. *Osteoporosis international*, 22(3):859–871, 2011.
- [218] Rita Rastogi Kalyani, Mark Corriere, and Luigi Ferrucci. Age-related and disease-related muscle loss: the effect of diabetes, obesity, and other diseases. *The Lancet Diabetes & Endocrinology*, 2(10):819–829, 2014.
- [219] Marybeth Brown and Eileen M Hasser. Complexity of age-related change in skeletal muscle. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 51(2):B117–B123, 1996.
- [220] Ezak Ahmad, Moi Hoon Yap, Hans Degens, and Jamie McPhee. Enhancement of mri human thigh muscle segmentation by template-based framework. In *Control System, Computing and Engineering (ICCSCE), 2014 IEEE International Conference on*, pages 405–410. IEEE, 2014.
- [221] M Greenacre and R Primicerio. Measures of distance between samples: Euclidean. *Fundacion BBVA Publication (December 2013). ISBN*, pages 978–84, 2008.

- [222] Peter Kovési. Image features from phase congruency. *Videre: Journal of computer vision research*, 1(3):1–26, 1999.
- [223] Ke Wang, Pengfeng Xiao, Xuezhi Feng, and Guiping Wu. Image feature detection from phase congruency based on two-dimensional hilbert transform. *Pattern Recognition Letters*, 32(15):2015–2024, 2011.
- [224] Peter Kovési. Phase congruency: A low-level image invariant. *Psychological research*, 64(2):136–148, 2000.
- [225] Nhat Vu and BS Manjunath. Shape prior segmentation of multiple objects with graph cuts. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.
- [226] Tony Chan and Wei Zhu. Level set based shape prior segmentation. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 1164–1170. IEEE, 2005.
- [227] Ghassan Hamarneh and Xiaoxing Li. Watershed segmentation using prior shape and appearance knowledge. *Image and Vision Computing*, 27(1):59–68, 2009.
- [228] Daniel Freedman and Tao Zhang. Interactive graph cut based segmentation with shape priors. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 755–762. IEEE, 2005.
- [229] Mohamed N Ahmed, Sameh M Yamany, Nevin Mohamed, Aly A Farag, and Thomas Moriarty. A modified fuzzy c-means algorithm for bias field estimation and segmentation of mri data. *Medical Imaging, IEEE Transactions on*, 21(3):193–199, 2002.
- [230] RT Jagoe and MPKJ Engelen. Muscle wasting and changes in muscle protein metabolism in chronic obstructive pulmonary disease. *European Respiratory Journal*, 22(46 suppl):52s–63s, 2003.
- [231] Martin Thoma. A survey of semantic segmentation. *arXiv preprint arXiv:1602.06541*, 2016.
- [232] Evan Shelhamer, Jonathan Long, and Trevor Darrell. Fully convolutional networks for semantic segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(4):640–651, 2017.

- [233] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.
- [234] Hyeonwoo Noh, Seunghoon Hong, and Bohyung Han. Learning deconvolution network for semantic segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1520–1528, 2015.
- [235] Son Lam Phung, Abdesselam Bouzerdoum, and Douglas Chai. Skin segmentation using color pixel classification: analysis and comparison. *IEEE transactions on pattern analysis and machine intelligence*, 27(1):148–154, 2005.
- [236] Hongsheng Li, Rui Zhao, and Xiaogang Wang. Highly efficient forward and backward propagation of convolutional neural networks for pixelwise classification. *arXiv preprint arXiv:1412.4526*, 2014.
- [237] Itamar Arel, Derek C Rose, and Thomas P Karnowski. Deep machine learning—a new frontier in artificial intelligence research [research frontier]. *IEEE computational intelligence magazine*, 5(4):13–18, 2010.
- [238] Xue-Wen Chen and Xiaotong Lin. Big data deep learning: challenges and perspectives. *IEEE access*, 2:514–525, 2014.
- [239] David Belanger and Andrew McCallum. Structured prediction energy networks. In *International Conference on Machine Learning*, pages 983–992, 2016.
- [240] Alessandro Lusci, Gianluca Pollastri, and Pierre Baldi. Deep architectures and deep learning in chemoinformatics: the prediction of aqueous solubility for drug-like molecules. *Journal of chemical information and modeling*, 53(7):1563–1575, 2013.
- [241] Yi Sun, Xiaogang Wang, and Xiaoou Tang. Deep learning face representation from predicting 10,000 classes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1891–1898, 2014.
- [242] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Domain adaptation for large-scale sentiment classification: A deep learning approach. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 513–520, 2011.

- [243] Andy Zeng, Kuan-Ting Yu, Shuran Song, Daniel Suo, Ed Walker, Alberto Rodriguez, and Jianxiong Xiao. Multi-view self-supervised deep learning for 6d pose estimation in the amazon picking challenge. In *Robotics and Automation (ICRA), 2017 IEEE International Conference on*, pages 1386–1383. IEEE, 2017.
- [244] James Bergstra, Olivier Breuleux, Pascal Lamblin, Razvan Pascanu, Olivier Delalleau, Guillaume Desjardins, Ian Goodfellow, Arnaud Bergeron, Yoshua Bengio, and Pack Kaelbling. Theano: Deep learning on gpus with python. 2011.
- [245] Znaonui Liang, Gang Zhang, Jimmy Xiangji Huang, and Qmming Vivian Hu. Deep learning for healthcare decision making with emrs. In *Bioinformatics and Biomedicine (BIBM), 2014 IEEE International Conference on*, pages 556–559. IEEE, 2014.
- [246] Scott R Granter, Andrew H Beck, and David J Papke Jr. Straw men, deep learning, and the future of the human microscopist: Response to artificial intelligence and the pathologist: Future frenemies?. *Archives of Pathology & Laboratory Medicine*, 141(5):624–624, 2017.
- [247] Manu Goyal, Neil D Reeves, Adrian K Davison, Satyan Rajbhandari, Jennifer Spragg, and Moi Hoon Yap. Dfunet: Convolutional neural networks for diabetic foot ulcer classification. *arXiv preprint arXiv:1711.10448*, 2017.
- [248] Manu Goyal, Neil D Reeves, Satyan Rajbhandari, Jennifer Spragg, and Moi Hoon Yap. Fully convolutional networks for diabetic foot ulcer segmentation. *arXiv preprint arXiv:1708.01928*, 2017.
- [249] Jhan S Alarifi, Manu Goyal, Adrian K Davison, Darren Dancey, Rabia Khan, and Moi Hoon Yap. Facial skin classification using convolutional neural networks. In *International Conference Image Analysis and Recognition*, pages 479–485. Springer, 2017.
- [250] Manu Goyal and Moi Hoon Yap. Multi-class semantic segmentation of skin lesions via fully convolutional networks. *arXiv preprint arXiv:1711.10449*, 2017.
- [251] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

- [252] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [253] Song Han, Jeff Pool, John Tran, and William Dally. Learning both weights and connections for efficient neural network. In *Advances in neural information processing systems*, pages 1135–1143, 2015.
- [254] Hoo-Chang Shin, Holger R Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M Summers. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE transactions on medical imaging*, 35(5):1285–1298, 2016.
- [255] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, 2015.
- [256] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3431–3440, 2015.
- [257] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. 2017.
- [258] Brian W Matthews. Comparison of the predicted and observed secondary structure of t4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure*, 405(2):442–451, 1975.
- [259] Sabri Boughorbel, Fethi Jarray, and Mohammed El-Anbari. Optimal classifier for imbalanced data using matthews correlation coefficient metric. *PloS one*, 12(6): e0177678, 2017.
- [260] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241. Springer, 2015.

- [261] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv:1511.00561*, 2015.
- [262] Simon Jégou, Michal Drozdal, David Vazquez, Adriana Romero, and Yoshua Bengio. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. *arXiv preprint arXiv:1611.09326*, 2016.
- [263] Paul Viola and William M Wells III. Alignment by maximization of mutual information. *International journal of computer vision*, 24(2):137–154, 1997.
- [264] André Collignon, Frederik Maes, Dominique Delaere, Dirk Vandermeulen, Paul Suetens, and Guy Marchal. Automated multi-modality image registration based on information theory. In *Information processing in medical imaging*, volume 3, pages 263–274, 1995.
- [265] Charles R Meyer, Jennifer L Boes, Boklye Kim, Peyton H Bland, Kenneth R Zasadny, Paul V Kison, Kenneth Koral, Kirk A Frey, and Richard L Wahl. Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations. *Medical image analysis*, 1(3):195–206, 1997.
- [266] Philippe Thévenaz and Michael Unser. Optimization of mutual information for multiresolution image registration. *Image Processing, IEEE Transactions on*, 9(12):2083–2099, 2000.
- [267] Gary E Christensen, Michael I Miller, Michael W Vannier, and Ulf Grenander. Individualizing neuro-anatomical atlases using a massively parallel computer. *Computer*, 29(1):32–38, 1996.
- [268] Anthony Yezzi, Lilla Zollei, and Tina Kapur. A variational framework for joint segmentation and registration. In *Mathematical Methods in Biomedical Image Analysis, 2001. MMBIA 2001. IEEE Workshop on*, pages 44–51. IEEE, 2001.
- [269] Paul P Wyatt and J Alison Noble. Map mrf joint segmentation and registration of medical images. *Medical Image Analysis*, 7(4):539–552, 2003.
- [270] Bruce Fischl, David H Salat, Evelina Busa, Marilyn Albert, Megan Dieterich, Christian Haselgrove, Andre Van Der Kouwe, Ron Killiany, David Kennedy, Shuna

- Klaveness, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355, 2002.
- [271] Stefano Soatto and Anthony J Yezzi. Deformation deforming motion, shape average and the joint registration and segmentation of images. In *Computer VisionECCV 2002*, pages 32–47. Springer, 2002.
- [272] Frédéric JP Richard and Laurent D Cohen. A new image registration technique with free boundary constraints: application to mammography. *Computer Vision and Image Understanding*, 89(2):166–196, 2003.
- [273] Baba C Vemuri, Yunmei Chen, and Zhizhou Wang. Registration assisted image smoothing and segmentation. In *Computer VisionECCV 2002*, pages 546–559. Springer, 2002.