

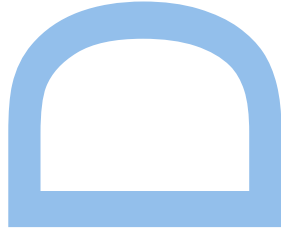
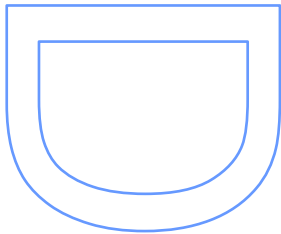
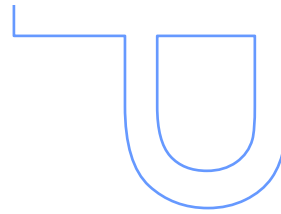
Development of algorithms for detection and quantification of rheumatic diseases in musculoskeletal ultrasound

Nelson Costa Martins

Doctoral Programme in Computer Science – MAPi

Department of Computer Science - FCUP

2019





Development of algorithms for detection and quantification of rheumatic diseases in musculoskeletal ultrasound

Nelson Costa Martins

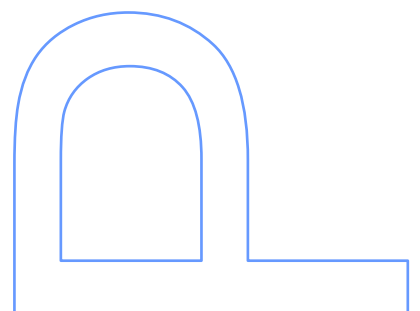
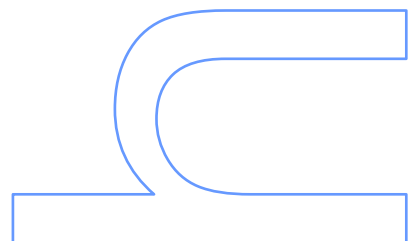
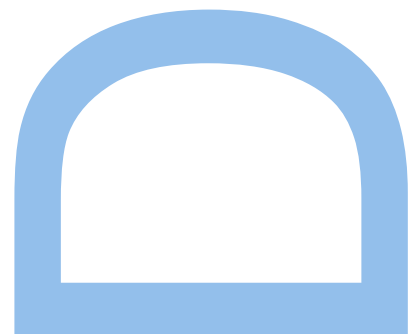
Doctoral Programme in Computer Science - MAPi
Department of Computer Science, Faculty of Sciences
University of Porto
2019

Supervisor

Miguel Tavares Coimbra,
Assistant Professor, Department of Computer Science Faculty of
Sciences of the University of Porto, Portugal

Co-Supervisor

Manuel João Oliveira Ferreira
Head of Research, Development and Innovation
Neadvance, Machine Vision, SA, Braga, Portugal



UNIVERSITY OF PORTO

PHD THESIS

Development of Algorithms for Detection and Quantification of Rheumatic Diseases in Musculoskeletal Ultrasound

Author:

Nelson MARTINS

Supervisors:

Doutor Miguel COIMBRA

Doutor Manuel FERREIRA

*Submitted in part fulfillment of the requirements for the degree of
Doctor of Philosophy in Computer Sciences, for the MAPi doctoral program.*

Faculty of Sciences of the University of Porto

&

Neadvance, Machine Vision, SA

November 2019

Acknowledgements

I, Nelson MARTINS, would like to thank Doctor Manuel Ferreira and Doctor Miguel Coimbra as my main supporters. I'm also very thankful to the Neadvance - Machine Vision SA, Faculty of Sciences of the University of Porto and Telecommunications Institute - Porto teams for their help as friends and work partners. To Dra Mónica Bogas and Dra Filipa Teixeira many thanks, their medical experience was very important along this work. And finally, to my family and friends for their unconditional support.

UNIVERSIDADE DO PORTO

Resumo

Programa Doutoral

MAPI

Doutoramento em Informática

Desenvolvimento de Algoritmos de Detecção e Quantificação de Doenças Reumáticas para Ecografia Músculo-Esquelética

por Nelson MARTINS

As doenças reumáticas são a maior causa de dor e perda de mobilidade nos países desenvolvidos, fazendo delas um grave problema social, económico e de saúde pública. Devido às suas vantagens, a ecografia tem vindo a ser introduzida na prática clínica dos reumatologistas de forma a facilitar o diagnóstico e o seguimento dos pacientes. Ao contrário da radiografia, que é atualmente a modalidade imagiológica *standard*, a ecografia não utiliza radiação ionizante e permite uma detecção mais precoce e melhores seguimentos de algumas doenças. No entanto, a dificuldade na interpretação e aquisição das imagens ecográficas reduz a sua aceitação e, por esse motivo, são necessárias soluções inovadoras que potenciem a sua utilização em contexto clínico.

Neste trabalho utilizam-se imagens da segunda articulação metacarpofalângica pela sua importância para o diagnóstico de doenças reumáticas tal como a Artrite Reumatóide. Relativamente a estas imagens, foram identificadas problemas em aberto na literatura, os quais se procuraram responder com as abordagens descritas ao longo deste trabalho.

Segmentação do metacarpo e da falange utilizando contornos ativos locais. Esta técnica permitiu identificar 80% das imagens com uma distância média de Hausdorff inferior a 3 pixels;

A identificação do tendão extensor foi alcançada aplicando uma variante do algoritmo de contornos ativos abertos, utilizando simetria de fase como pré-processamento e dados de forma do tendão para reforçar o modelo. Por fim, o tendão foi obtido pela optimização dos valores das energias recorrendo a algoritmos genéticos. Os resultados

alcançados apontam para uma distância média de Hausdorff inferior a 10 pixels (ou 0.5 mm) em 95% das imagens;

Para a identificação da cápsula articular foram testadas duas abordagens distintas, uma baseada no algoritmo SLIC, e outra baseada em redes neuronais convolucionais, mais propriamente o modelo *UNet*. Os resultados indicam uma clara superioridade do modelo *UNet* que tem ainda a vantagem de não depender de outros métodos de segmentação (como é o caso do SLIC) no processo de inicialização. Os resultados demonstraram que 90% das imagens foram identificadas com um coeficiente de Dice superior a 0.6.

Por fim, foi testada uma abordagem preliminar de extração de características e classificação para o problema específico da sinovite. Esta abordagem partiu das segmentações do metacarpo, falange e tendão extensor para criar um conjunto de máscaras para caracterizar localmente a imagem e identificar as alterações criadas na mesma devido a sinovite. Tendo sido obtido um index de Youden's de 0.84 para a identificação de sinovite e de 0.94 para a classificação.

Palavras-chave: Doenças Reumáticas; Artrite Reumatóide; Sinovite; Articulação Metacarpofalângica; Ecografia; Contornos ativos; Simetria de Fase; SLIC; Redes Neuronais Convolucionais; Máquinas de Vectores de Suporte; Extração de Características.

UNIVERSITY OF PORTO

Abstract

MAPI

Doctoral Program

Doctor of Philosophy in Computer Sciences

Development of Algorithms for Detection and Quantification of Rheumatic Diseases in Musculoskeletal Ultrasound

by Nelson MARTINS

Rheumatic diseases are the main cause of impairment and pain in developed countries, which makes them a critical social, health and economic problem. Due to their main advantages, ultrasounds are now being used in rheumatology to diagnosis and evaluate rheumatic diseases in their early stages. Unlike radiography, which is the current standard, ultrasounds are less expensive, do not use ionizing radiation and can lead to better early diagnosis and follow-up outcomes. The difficulties in the interpretation and acquisition of this type of image reduce its acceptance, and thus, new and innovative solutions are needed to help doctors in the diagnostic process.

In this study, images of the second metacarpophalangeal joint will be used because they are of great importance in the diagnosis of rheumatic diseases such as Rheumatoid Arthritis. Regarding these images, there are image processing problems that are still unsolved, and this work aims to solve them.

The segmentation of the metacarpus and phalangeal bone was achieved using Localizing Active Contours. This approach allowed the identification of 80% of the images with a Modified Hausdorff Distance below 3 pixels.

The extensor tendon was identified with the proposal of an Open Ended Active Contours method using Phase Symmetry pre-processing, prior structure knowledge and Genetic Algorithm based optimization. The results show that the segmentation was achieved with a confidence of 95% for a Modified Hausdorff Distance below 10 pixels;

For the segmentation of the joint capsule, two distinct approaches were tested: one using the Simple Linear Iterative Clustering algorithm followed by a special shape constrained merge strategy, and the other using convolutional neural networks, more precisely the *UNet* model. The results show that the *UNet* model outperforms the clustering method, without the necessity of other segmentation methods to limit the joint capsule search zone, as is the case for the Simple Linear Iterative Clustering method. The segmentation was achieved with a Dice Coefficient higher than 0.6 in 90% of the images.

Finally, a preliminary feature extraction and classification study was presented, specifically addressing the synovitis. The proposed approach started from the segmentation of the metacarpus, phalange and extensor tendon to create a set of masks. These masks were used to locally characterize the images and detect the anatomical changes provoked by the synovitis.

Keywords: Rheumatic Diseases; Rheumatoid Arthritis; Synovitis; Metacarpophalangeal Joint; Ultrasounds; Active Contours; Phase Symmetry; Simple Linear Iterative Clustering; Convolutional Neural Networks; Support Vector Machines; Feature Extraction.

Contents

Acknowledgements	iii
Resumo	v
Abstract	vii
Contents	ix
List of Figures	xiii
List of Tables	xvii
Abbreviations	xix
1 Introduction	1
1.1 Motivation	1
1.2 Research Questions	4
1.3 Scientific Contributions	4
1.4 Thesis Outline	7
2 Ultrasound Imaging	9
2.1 Introduction	9
2.2 Acquisition Equipment	10
2.3 Image Characteristics	12
2.4 Doppler Mode	13
3 Rheumatology Background	17
3.1 Introduction	17
3.1.1 Causes	18
3.1.2 Epidemiology	19
3.2 Rheumatoid Arthritis	20
3.2.1 Diagnosis	21
3.2.2 Treatment	24
3.2.3 Costs	25
3.3 Ultrasound for Rheumatoid Arthritis	25
3.3.1 Metacarpophalangeal Joint	26
3.3.2 Selected Acquisition Protocol	28

4	Image Processing and Analysis Background	31
4.1	Ultrasound Imaging Existing Work	32
4.1.1	Pre-Processing	32
4.1.2	Segmentation	33
4.1.3	Feature Extraction and Classification	34
4.2	Active Contours Model	35
4.2.1	Internal Energy	36
4.2.2	External Energy	36
4.2.3	Optimization	37
4.3	Localizing Active Contours	37
4.4	Phase Symmetry	39
4.5	Simple Linear Iterative Clustering	41
4.6	Convolutional Neural Networks	42
4.7	Performance Metrics	45
4.7.1	Modified Hausdorff Distance	45
4.7.2	Dice Similarity Coefficient	47
5	Metacarpus and Phalange Identification	49
5.1	Introduction	49
5.2	Proposed Work	51
5.2.1	Pre-processing	52
5.2.2	Contour Initialization	52
5.2.3	Contour Refinement - LAC	52
5.2.4	Upper Line Extraction	53
5.3	Results and Discussion	54
5.4	Complementary Results	55
5.5	Conclusions	57
6	Extensor Tendon Identification	59
6.1	Introduction	59
6.2	Segmentation Framework	61
6.2.1	Internal Energy	62
6.2.2	External Energy	63
6.2.3	Area Constraints	65
6.2.4	Additional Constraints	66
6.2.5	Implementation	67
6.3	Results and Discussion	68
6.3.1	Dataset and Metrics	68
6.3.2	Number of Points Optimization	69
6.3.3	Weights Optimization	70
6.3.4	Visual Results	72
6.4	Complementary Results	73
6.4.1	New Data Results	73
6.4.2	Fully Automatic Segmentation	74
6.5	Conclusions	77
7	Joint Capsule Identification	81

7.1	Introduction	81
7.2	Proposed Work	84
7.2.1	Split and Merge	84
7.2.2	Convolutional Neural Networks	88
7.3	Results and Discussion	92
7.3.1	Metrics	93
7.3.2	Results - <i>SM</i>	93
7.3.3	Results - CNN	96
7.3.4	Comparative Results and Discussion	97
7.4	Conclusions	100
8	Synovitis Detection and Grading	103
8.1	Introduction	103
8.2	Proposed Work	105
8.2.1	Feature Extraction	106
8.2.2	Feature Selection	108
8.3	Results and Discussion	109
8.4	Conclusions	112
9	Conclusions and Future work	115
A	Database Information and Study	133
A.1	Database Characteristics	133
A.2	Clinical Visual Features	134
A.2.1	Distances Between Bones, Extensor Tendon and Skin Line	135
A.2.2	Joint Capsule Area	136
A.2.3	Intensity Based	138

List of Figures

1.1	Illustration of the bone anatomy of the hand and respective nomenclature.	3
2.1	Typical acquisition system used in medical ultrasound imaging.	10
2.2	Example of ultrasound transducers.	11
2.3	Example of image visualization enhancement by speckle noise reduction.	13
2.4	Examples of artifacts found in ultrasound images.	14
2.5	Examples of a B-Mode, Doppler and Power Doppler images of the kidney from different patients.	15
3.1	Example of the progression of Rheumatoid Arthritis on the hands.	20
3.2	Illustration of a healthy metacarpophalangeal joint.	27
3.3	Illustration of a normal joint and a joint with synovitis.	28
3.4	Representation of the second metacarpophalangeal joint ultrasound image acquisition.	29
4.1	Representation of the Localizing Active Contour algorithm	38
4.2	Example of a symmetric and an asymmetric object.	39
4.3	Representation of the odd and even orthogonal filters.	40
4.4	Difference between a fully connected layer and a convolutional layer.	43
4.5	Example of receptive field growth in deeper layers.	44
4.6	Example of the low level and high-level features learn in shallow and deep layers (from the bottom to the top).	44
4.7	Examples of the Hausdorff distance between two lines	46
4.8	Visual representation of the Dice Coefficient in three different cases.	47
5.1	Examples of ultrasound images of the dorsal view of the second metacarpophalangeal joint.	50
5.2	Proposed pipeline for the identification of the metacarpus and phalange.	51
5.3	Visual results of the metacarpus and phalange segmentation using the Localizing Active Contours algorithm.	54

5.4	Segmentation error for the metacarpus and phalange, using the vertical distance in pixels.	55
5.5	Segmentation error for the metacarpus and phalange using the vertical distance in pixels, in all dataset.	56
5.6	Segmentation error for the metacarpus and phalange using the Modified Hausdorff Distance in millimeters, in all dataset.	57
6.1	Example of ultrasound images of the dorsal view of the second metacarpophalangeal joint.	60
6.2	Proposed open ended active contour model framework.	62
6.3	Example of the phase symmetry pre-processing results.	64
6.4	Box plot of the area measurements.	66
6.5	Approximation error using different number of points.	69
6.6	Results obtained for different parameter configurations of the proposed method.	71
6.7	Visual results of the extensor tendon segmentation.	72
6.8	Results obtained for the segmentation of the extensor tendon, in all database.	74
6.9	Diagram of the fully automatic extensor tendon segmentation method.	75
6.10	Results obtained in the test data for the different automatic and manual segmentation settings.	75
6.11	Visual results of the fully automatic segmentation of the extensor tendon for different configurations.	77
7.1	Example of ultrasound images of the second metacarpophalangeal joint with different joint capsules aspects and synovitis grades.	82
7.2	Joint capsule segmentation pipeline using a Split-Merge-Refine approach.	84
7.3	Results of the Simple Linear Iterative Clustering algorithm in the metacarpophalangeal joint images.	86
7.4	The <i>UNet</i> model.	89
7.5	Result of adding skip connections.	90
7.6	Proposed changes to the <i>UNet</i> model.	91
7.7	Boxplot of the <i>DICE</i> values obtained for the different configurations.	94
7.8	Segmentation results obtained for the different configurations.	95
7.9	Visual results obtained for the segmentation of the joint capsule for different configurations.	95
7.10	Visual results obtained for the segmentation of the joint capsule for different <i>UNet</i> parameterizations.	97
7.11	Boxplot of the <i>DICE</i> values obtained for the test images.	98

7.12 Results obtained for the segmentation of the joint capsule using three different methods.	99
7.13 Visual results obtained for the segmentation of the joint capsule using three different methods.	100
8.1 Representation of the synovitis grades convention.	104
8.2 Example of images with different synovitis grades.	105
8.3 Manual annotations used in the feature extraction.	106
8.4 Proposed masks for the synovitis identification and grading.	107
8.5 Selected features for the different synovitis detection and grading tests.	113
A.1 Manual annotations used to validate the proposed algorithms.	134
A.2 Box plot of the mean distances between the bones and the extensor tendon.	135
A.3 Box plot of the mean distances between the extensor tendon and the skin line.	136
A.4 Box plot of the joint capsule area extracted from the manual annotations for each synovitis grade.	137
A.5 Plot of the mean intensity extracted from inside and outside the joint capsule manual annotations.	139
A.6 Plot of the intensity standard deviation extracted from inside and outside the joint capsule manual annotations.	139

List of Tables

3.1	Epidemiology of major rheumatic diseases.	20
3.2	2010 ACR/EULAR classification criteria for rheumatoid arthritis.	22
6.1	<i>Confidence</i> obtained for the different configurations and for three MHD. . .	76
7.1	p -values resulting from the <i>Welch's t – test</i> of different pairs of methods. .	94
7.2	Mean <i>DICE</i> results in the test set.	96
7.3	p -values resulting from the <i>Welch's t – test</i> of different pairs of methods. .	98
8.1	Database organization and number of images used in the Monte Carlo cross-validation train and validation (Train/Validation) attending to the synovitis grade.	110
8.2	Feature selection results in the four tested configurations.	111
8.3	Summary of the results obtained before and after the feature selection phase for each configuration.	112
A.1	Final database size according to the synovitis grade.	133
A.2	Statistical significance (p -values) between the mean distance measurement and the synovitis degree.	136
A.3	Statistical significance of the difference in the joint capsule area in different synovitis grades.	137
A.4	Correlation between the joint capsule area and the mean distance between the bones and the extensor tendon for each synovitis grade.	138

Abbreviations

ACR	A merican C ollege of R heumatology
AUC	A rea U nder the C urve
bsDMARD	b iosimilar D isease M odifying A nti R heumatic D rugs
CAD	C omputer A ided D iagnosis
CAT	C omputed A xial T omography
DMARD	D isease M odifying A nti R heumatic D rugs
DNN	D eep N eural of N etwork
EM	E xpectation M aximization
EULAR	E Uropean L eaue A gainst R heumatism
IDC	I nverted D ice C oefficient
LAC	L ocalizing A ctive C ontours
MCPJ	M eta C arpo P halangeal J oint
MHD	M odified H ausdorff D istance
MRF	M arkov R andom F ield
MRI	M agnetic R esonance I maging
MTPJ	M eta T arso P halangeal J oint
PIP	P roximal I nter P halangeal
PCA	P incipal C omponent A nalysis
RA	R heumatoid A rthritis
SLIC	S imple L inear I terative C lustering
SVM	S upport V ector M achine
ULSAM	U nidade L ocal S áude A lto M inho
US	U ltra S ound

Chapter 1

Introduction

Contents

1.1 Motivation	1
1.2 Research Questions	4
1.3 Scientific Contributions	4
1.4 Thesis Outline	7

1.1 Motivation

Rheumatic diseases are pathologies that affect bones, joints and connective tissues. They are the main cause of impairment and pain in developed countries, making them a serious health problem with high social and economic implications [1]. Due to the increase in longevity and negative lifestyles, the number of rheumatic problems is likely to grow in the coming years. This creates pressure for the diagnosis of these diseases because, if not diagnosed and treated properly and in time, they progress faster, causing severe and irreversible physical deformations and, leading to incapacity [2]. An early diagnosis is then fundamental to give the patient a better quality of life and prevent faster deterioration of his/her condition.

The number of rheumatic diseases is very large and covers a broad range of different types of pathologies, but their assessment is usually based on physical exams, imaging techniques and sample based analysis [3]. A physical exam tests the mobility of the joints, pain, and other movement characteristics, such as morning stiffness and lack of

strength. Sample based analysis tests the presence and concentration of specific antibodies and other substances in the body [4]. The imaging techniques allow visualization of the internal body structures. In rheumatology, the most common imaging techniques are magnetic resonance imaging (MRI), radiography and ultrasound imaging. Due to their characteristics, ultrasounds have seen their use increase in the last few years. These characteristics include their low cost, ease of operation, ease of acceptance by patients, possibility of acquiring several images in different positions, and comparable results to the other imaging techniques [5, 6]. One limitation of the imaging techniques is the subjectivity of the diagnosis and the difficulty of image interpretation. This is especially true in ultrasound images, because the images tend to be more noisy and their acquisition more operator dependent. Moved by this problem and its potential impact, the research, development and testing of new advanced image processing algorithms are proposed in this work. Due to the previously mentioned advantages, ultrasound imaging was selected for use in this study. The aim is to contribute to the creation of a computer aided diagnosis (CAD) system to automatically extract meaningful information from the images that complements the information that the rheumatologists already extract from them empirically. In this manner, it is also expected to help improve the diagnosis and follow up for rheumatic diseases.

In medicine, and rheumatology in particular, the use of quantitative information is essential to obtain an evidence based diagnosis. Consequently, a CAD system can be an added value for doctors and their patients. In fact, such systems have been used to help visualize and extract information from medical information. The introduction of digital images and videos created the opportunity to use digital manipulation. From simple contrast enhancement, zooming and panning to more complex tools, such as statistical analysis and even artificial intelligence, they all help handle and enhance the information present in the images.

To achieve the proposed CAD system, work on automatic detection of physiological structures will be conducted, as well as quantification and classification of relevant physiological parameters. In this manner, it is expected that the assessment of rheumatic diseases will be improved, making it less subjective and reducing the variability intra and inter observer. Moreover, with more precise measures, the follow-up for these diseases could be more accurate and the treatment adjustment more assertive,

allowing a direct comparative analysis of the treatment responses for each patient. Ideally, this could also be used to create gold standards, improving the diagnosis even further.

Given that the number of possible acquisition protocols is very broad, it was decided to narrow the scope and focus the work on a single acquisition. The longitudinal-dorsal view of the metacarpophalangeal joint (MCPJ) was selected because it provides clues regarding the hand small joint, which is often the first structure to be affected by rheumatic diseases, such as rheumatoid arthritis. In Fig. 1.1 an illustration of the hand bone structure can be seen.

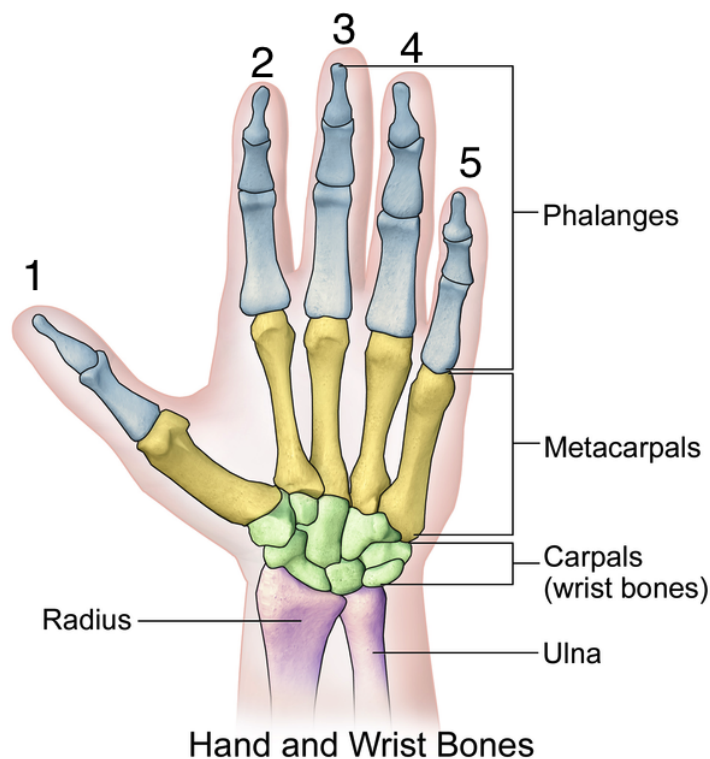


FIGURE 1.1: Illustration of the bone anatomy of the hand and respective nomenclature.¹

The metacarpals are the bones in the palmar region of the hand; they are directly connected to the phalanges and carpal bones, which are, respectively, the fingers and wrist bones. The numbers above each finger in Fig. 1.1 indicates the convention in the naming of each finger, which is used to facilitate communication between medical staff. The rule is that the thumb is the first finger, the index finger is the second, and so on, until the pinky, which is the fifth finger. This is used for both hands and feet.

¹Adapted from: Blausen.com staff. Medical gallery of blausen medical 2014. *WikiJournal of Medicine* 1 (2), 2014

The selection of the longitudinal-dorsal view was based on a literature review (explained in more detail in Chapter 3) and the experience of the rheumatologist that contributed in the creation of this work. Nevertheless, the proposed algorithms and metrics might be useful for other ultrasound images, such as metatarsophalangeal joints and interphalangeal joints, because the image characteristics are similar.

1.2 Research Questions

Three main research questions will be addressed during this PhD:

- Is it possible to, accurately and automatically, identify the structures present in ultrasound images of the MCPJ (bones, tendon and joint capsule)?
- Using feature extraction techniques, is it possible to classify an image regarding the presence or absence of pathology?
- Can this approach improve diagnosis and/or follow up for rheumatic diseases?

Regarding the last question we assume that, given its complexity and necessary resources, we may not fully answer it. It involves several patients and Rheumatologists and a long period of time to track the evolution of the disease. Because of that we will aim for a preliminary study and move later to a larger study.

1.3 Scientific Contributions

Throughout the PhD work several contributions were proposed; next, a summary of all contributions is presented:

Journal publications

- N. Martins, S. Sultan, D. Veiga, M. Ferreira, F. Teixeira, and M. Coimbra. A new active contours approach for finger extensor tendon segmentation in ultrasound images using prior knowledge and phase symmetry. *IEEE Journal of Biomedical and Health Informatics*, pages 1–1, 2018. ISSN 2168-2194. doi: 10.1109/JBHI.2017.2723819

Conference publications

- N. Martins, M. S. Sultan, D. Veiga, M. Ferreira, and M. Coimbra. Segmentation of the metacarpus and phalange in musculoskeletal ultrasound images using local active contours. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4097–4100, Aug 2016. doi: 10.1109/EMBC.2016.7591627
- N. Martins, M. S. Sultan, D. Veiga, M. Ferreira, and M. Coimbra. Joint capsule segmentation in ultrasound images of the metacarpophalangeal joint using a split and merge approach. In *2018 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, pages 243–246, March 2018. doi: 10.1109/BHI.2018.8333414
- Martins, N., et al. "Fully Automatic Finger Extensor Tendon Segmentation in Ultrasound Images of the Metacarpophalangeal Joint." *Engineering in Medicine and Biology Society (EMBC)*, 2018 IEEE 40th Annual International Conference of the. IEEE, 2018.

Other Contributions

- M. S. Sultan, N. Martins, M. J. Ferreira, and M. T. Coimbra. Segmentation of bones mcp joint region of the hand from ultrasound images. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3001–3004, Aug 2015. doi: 10.1109/EMBC.2015.7319023
- Malik Saad Sultan, Nelson Martins, Diana Veiga, Manuel Ferreira, and Miguel Coimbra. Automatic segmentation of extensor tendon of the mcp joint in ultrasound images. In *Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies*, BIOSTEC 2016, pages 71–76, Portugal, 2016. SCITEPRESS - Science and Technology Publications, Lda. ISBN 978-989-758-170-0. doi: 10.5220/0005692500710076
- M. S. Sultan, N. Martins, D. Veiga, M. J. Ferreira, and M. T. Coimbra. Tracking of the anterior mitral leaflet in echocardiographic sequences using active contours. In *2016 38th Annual International Conference of the IEEE Engineering*

in Medicine and Biology Society (EMBC), pages 1074–1077, Aug 2016. doi: 10.1109/EMBC.2016.7590889

- Malik Saad Sultan, Nelson Martins, Eva Costa, Diana Veiga, Manuel João Ferreira, Sandra da Silva Mattos, and Miguel Tavares Coimbra. Real-time anterior mitral leaflet tracking using morphological operators and active contours. In *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2017) - Volume 2: BIOIMAGING, Porto, Portugal, February 21-23, 2017.*, pages 39–46, 2017. doi: 10.5220/0006244700390046
- Eva Costa, Nelson Martins, Malik Saad Sultan, Diana Veiga, Manuel João Ferreira, Sandra da Silva Mattos, and Miguel Tavares Coimbra. A preliminary approach for the segmentation of mitral valve regurgitation jet in doppler ecocardiography images. In *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2017) - Volume 2: BIOIMAGING, Porto, Portugal, February 21-23, 2017.*, pages 47–54, 2017. doi: 10.5220/0006248900470054
- M. S. Sultan, N. Martins, E. Costa, D. Veiga, M. J. Ferreira, S. Mattos, and M. T. Coimbra. Tracking large anterior mitral leaflet displacements by incorporating optical flow in an active contours framework. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3244–3247, July 2017. doi: 10.1109/EMBC.2017.8037548
- M. S. Sultan, N. C. Martins, E. Costa, D. Veiga, M. J. Ferreira, S. Mattos, and M. T. Coimbra. Virtual m-mode for echocardiography: A new approach for the segmentation of the anterior mitral leaflet. *IEEE Journal of Biomedical and Health Informatics*, pages 1–1, 2018. ISSN 2168-2194. doi: 10.1109/JBHI.2018.2799738
- M.S. Sultan, N. Martins, E. Costa, D. Veiga, M.J. Ferreira, S. Mattos, and M. Coimbra, “A New Method for the Anterior Mitral Leaflet Segmentation in Echocardiography Videos by Using the Virtual M-mode Space”, in *Proc. of IEEE EMBC 2018, Honolulu, USA, Jul 2018.*

- L. Pires, M.S. Sultan, N. Martins, E. Costa, D. Veiga, M.J. Ferreira, S. Mattos, and M. Coimbra, “Extracting Thickness Profiles of Anterior Mitral Leaflets in Echocardiography Videos”, in Proc. of IEEE EMBC 2018, Honolulu, USA, Jul 2018.

1.4 Thesis Outline

This PhD thesis is divided into nine chapters, each as independent as possible from the others to facilitate the reading. There was care to separate the background fundamentals from the contributions, and the structure of this thesis is thus organized as follows:

Chapter 1 - Introduction creates the basis for the entire thesis; it presents the general problem and addresses the motivation and expected contributions.

Chapter 2 - Ultrasound Imaging introduces the acquisition system and the resultant images and justifies its choice over other imaging techniques, as well as its advantages and disadvantages.

Chapter 3 - Rheumatology Background introduces the clinical point of view of the problem; it presents a brief overview of rheumatic diseases with a higher focus on rheumatoid arthritis, the MCPJ and the ultrasound imaging, which are the main topics in the study.

Chapter 4 - Image Processing and Analysis Background presents the main frameworks used in this work. This chapter is composed of existing works, which are used as bases for the proposals in the following chapters.

Chapter 5 - Metacarpus and Phalange Identification addresses the automatic segmentation of the metacarpus and phalange in the ultrasound images of the MCPJ using Localizing Active Contours.

Chapter 6 - Extensor Tendon Identification addresses the segmentation of the extensor tendon in the ultrasound images of the second MCPJ using a new open-ended active contours framework with phase symmetry pre-processing and anatomical prior knowledge constraints.

Chapter 7 - Joint Capsule Identification addresses the segmentation of the joint capsule in the ultrasound images of the second MCPJ using two approaches. One is using a split and merge approach, where the Simple Linear Iterative Clustering algorithm is used, followed by region growing with shape constraints. The other is based on convolutional neural networks, or more precisely, the *UNet* model.

Chapter 8 - Synovitis Detection and Grading addresses the detection and quantification of synovitis in the ultrasound images of the second MCPJ using a set of proposed masks and features in an SVM model.

Chapter 9 - Conclusions and Future work is the final chapter and the one that aggregates all of the main conclusions achieved in this work. It points out the main ideas that can be extracted from the work, and based on that, it proposes possible paths for future work.

Chapter 2

Ultrasound Imaging

Contents

2.1 Introduction	9
2.2 Acquisition Equipment	10
2.3 Image Characteristics	12
2.4 Doppler Mode	13

2.1 Introduction

Ultrasound Imaging is a diagnostic tool used to visualize the internal structure of the human body using sound waves with high frequency (above 20 kHz). The main principle is that a portion of the sound waves' energy is reflected when they hit an acoustic transition, while the remaining energy continues on its path. An acoustic transition exists when materials with different acoustic impedances are in contact, for example, blood and cardiac muscle or muscle and bone. The percentage of reflected energy is dependent on the characteristics of the acoustic transition and can be used to characterize tissues or other structures [18]. It is one of the most used medical imaging systems due to its advantages when compared with other techniques. The low cost, highly informative outputs and real-time visualization and interaction with the patient are among them. Additionally, ultrasounds do not use ionizing radiation, while the results are comparable with the remaining techniques [19]. Looking at ultrasound disadvantages, the images are difficult to read because of several artifacts that the user needs to be aware of. The

depth to which the ultrasound can penetrate is limited, and the spatial resolution of the image is inversely proportional to the depth. It is also impossible to visualize bones' interiors, because they are acoustic barriers. Finally, the lack of agreement regarding the metrics and parameters to be used in specific cases is problematic. With research and technology progression, most of these limitations are expected to be reduced.

2.2 Acquisition Equipment

The ultrasound equipment has a specific set-up that is needed to acquire the ultrasound images. In Fig. 2.1, the typical apparatus used in medical ultrasound imaging is shown.

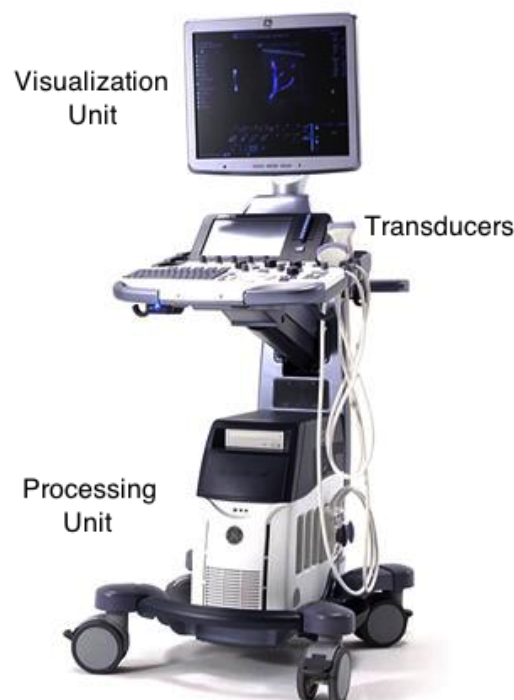


FIGURE 2.1: Typical acquisition system used in medical ultrasound imaging.

It consists of: a processing unit (a computer or embedded device), which is also responsible for the data storage; a visualization unit (screen); and a transducer. Regarding the processing unit, it should be able to handle and store all of the information, and thus sufficient processing power and data space need to be available for these tasks. The visualization unit should have high contrast so that small changes in the gray-scale image are easily detected by the human eye. Finally, the transducer is a critical component

and should be chosen according to the structure we want to visualize. In Fig. 2.2, some ultrasound transducers are shown.

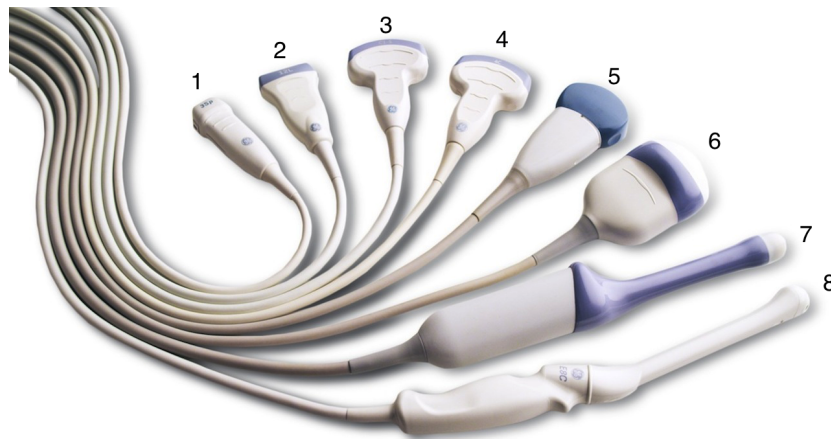


FIGURE 2.2: Example of ultrasound transducers, adapted from [20]. 1 and 2 - 2D linear probes used to visualize small body parts; 3, 4 and 5 - 2D convex probes used to visualize larger body parts; 6 - 4D probe used to visualize larger structures in 3D + time; 7 and 8 - 4D and 2D endocavity probes used to visualize internal structures using the patient's cavities.

The transducers have different shapes, characteristics and purposes, which need to be taken into account during the diagnosis. The first feature that needs to be selected is the probe shape, with the most common being linear, curvilinear and endocavity. The first is used for smaller parts and typically has higher frequencies values, providing higher details and lower depths. The curvilinear probes are used in obstetrics and general abdominal applications, and they have lower frequency values and higher depths compared to linear probes. The endocavity is a special type used for vaginal, rectal and transesophageal applications, where the probe needs to be positioned inside the cavities of patients. For each probe type, it is also necessary to select the most suitable probe size and other characteristics, such as Doppler support. One additional and fundamental component is the ultrasound gel. It is used to reduce the acoustic impedance between the transducer and the skin, allowing the sound waves to travel from the transducer to the interior of the patients without significant signal loss. In rheumatology, the most commonly used probes are linear ones with high frequencies so that higher detail is obtained. For the case of small joints, it is preferable to use smaller probes so that contact with the surface of the skin is easier.

2.3 Image Characteristics

Ultrasound images have a characteristically noisy appearance, which is known as speckle. It is considered a multiplicative noise and, therefore, proportional to the local gray-level intensity of the signal. Moreover, there are several artifacts that may occur during the image acquisition [21]. Both are considered limitations of ultrasound acquisition systems and will be discussed in more detail next.

Speckle noise:

Every acquisition system is susceptible to noise interference, and in ultrasound imaging devices, this problem is especially relevant. The speckle noise appears as a granular structure and results from the interaction between scattered sound waves and signal sound waves. It is typically found in other acquisition systems, such as sonar, synthetic aperture radar and optical coherence tomography. Objects smaller than the wavelength of the transmitted wave are also promoters of speckle noise, resulting in the cancellation or enhancement of the signal [22]. There are some ways to reduce the speckle noise. One possible method is to increase the transducer frequency, which decreases the wavelength of the transmitted signal, making it less prone to these effects. The problem, when using ultrasounds is that higher frequencies result in lower penetration depths, which can be problematic in some applications. Additionally, the frequency is limited by the equipment and the technology development. Another method is to use filters to remove or minimize the noise. In Fig. 2.3, the effect of a digital speckle reduction filter is shown.

Artifacts:

One additional characteristic of ultrasound imaging systems is the presence of artifacts. Artifacts differ from noise in that they are not caused by the random interference of the sound waves but rather are due to interactions that result in visible structures in the image that do not exist in the real world. Artifacts can be divided into two main groups: those that carry information and are useful in the diagnosis and those that do not carry meaningful information and only make the diagnosis more difficult. Both can be minimized with proper selection of the acquisition parameters, but the former are sometimes created on purpose to help in the diagnosis. In Fig. 2.4, some examples of possible artifacts are shown.



FIGURE 2.3: Example of image visualization enhancement by speckle noise reduction. Image courtesy of ContextVision.

In Fig. 2.4 - A, a shadow cast by the patella can be seen. This artifact occurs when the ultrasound waves find a transversal high acoustic impedance transition, such as a bone, calcification or gas collection. In these cases, the structures beyond the high transitions are not visible because no ultrasound waves can reach them. In Fig. 2.4 - B, a lateral acoustic shadowing can be seen due to the presence of a tangential high acoustic impedance transition; it works in a similar manner to the acoustic shadowing, but because the acoustic transition is tangential, the shadowing is only visible in the laterals, parallel to the sound waves. In Fig. 2.4 - C, a posterior acoustic enhancement can be seen due to the presence of an anechoic region (liquid). These structures create a signal response that is over-enhanced by the ultrasound device, resulting in a bright structure after the anechoic region. Finally, in Fig. 2.4 - D, a reverberation artifact is shown. It occurs when the sound waves travel back and forth reflecting several times between two structures. As the sound wave goes back and forth, the transducer receives several echoes that are seen as sequential acoustic transitions parallel to each other.

2.4 Doppler Mode

The Doppler Mode is a special ultrasound acquisition mode that uses the Doppler effect to capture motion patterns. The Doppler effect is observed when a wave hits a moving

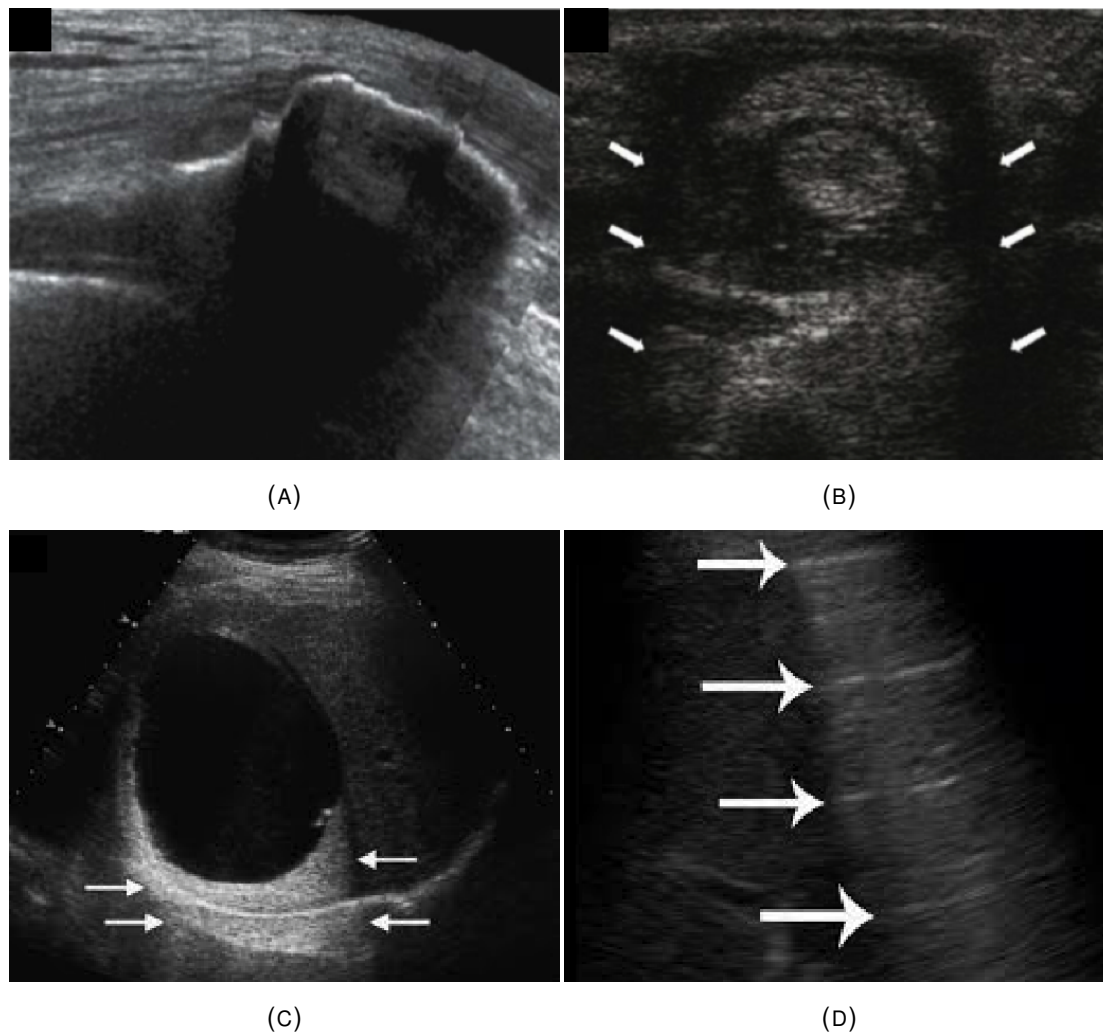


FIGURE 2.4: Examples of artifacts found in ultrasound images; A) acoustic shadowing on a patella - dark region; B) lateral acoustic shadowing on a finger extensor tendon - darker regions pointed by the arrows; C) posterior acoustic enhancement - brighter region pointed by the arrows; D) reverberation - successive parallel lines. Adapted from [19].

object. When that happens, the reflected wave has its frequency changed proportionally to the velocity of the object. For instance, if the object is moving towards the wave source, the echo will have a higher frequency than the sent wave. For the opposite case of an object moving away from the wave, the frequency of the echo will be lower. This mode is commonly used to capture the blood direction and velocity, as well as inflammatory activity. The power Doppler mode is a variation of the Doppler mode in which the direction of the object is discarded and only the magnitude is used to capture the presence of small movements. This is especially interesting in inflammatory activity, where the direction is not as important as the presence of movement. Fig. 2.5 presents three images acquired with the B-mode, Doppler and Power Doppler.

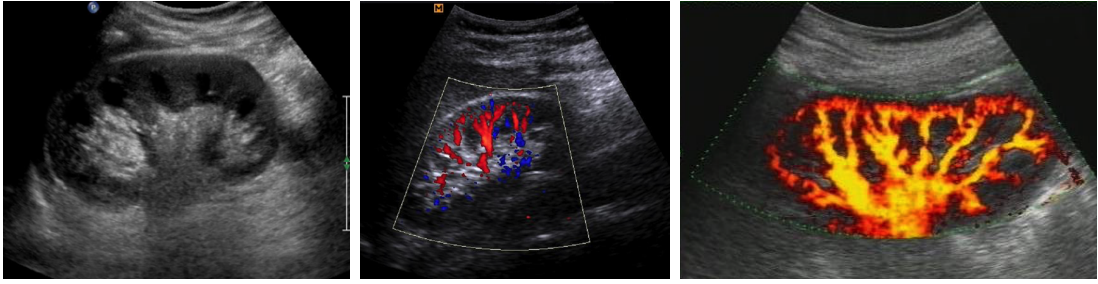


FIGURE 2.5: Examples of a B-Mode, Doppler and Power Doppler images of the kidney from different patients.

From Fig. 2.5, it is possible to gain an idea of the output of the Doppler and power Doppler modes. On the right, a B-Mode image is presented in which different tissues are visible. In the middle, the Doppler signal is visualized over the B-Mode; the pseudo-color presented in the Doppler mode represents the direction, where blue is moving away and red is moving towards the probe. The power Doppler is on the right and only has one color, representing the presence or absence of movement.

Chapter 3

Rheumatology Background

Contents

3.1 Introduction	17
3.1.1 Causes	18
3.1.2 Epidemiology	19
3.2 Rheumatoid Arthritis	20
3.2.1 Diagnosis	21
3.2.2 Treatment	24
3.2.3 Costs	25
3.3 Ultrasound for Rheumatoid Arthritis	25
3.3.1 Metacarpophalangeal Joint	26
3.3.2 Selected Acquisition Protocol	28

3.1 Introduction

Rheumatology is a sub-specialty of internal medicine that focuses on the diagnosis and therapy of rheumatic diseases. These diseases involve clinical problems in the musculoskeletal system, such as bones, muscles and connective tissues. They can affect individuals of all ages, with greater incidence in older people [23]. If not diagnosed and treated properly and in a timely manner, they can cause severe and irreversible physical deformations that considerably reduce the quality of life of those who suffer from these diseases. In fact, they are the main cause of impairment and pain in developed

countries, which makes them a serious health problem with high social and economic implications [2].

The most common manifestations of rheumatic diseases are pain, inflammation, swelling, movement stiffness and fatigue, but other symptoms can be present in the eyes, mouth and skin, such as dryness and rashes. The location and distribution of the symptoms will dictate the first diagnosis, and the complementary exams will confirm or disprove the prognosis. A correct early diagnosis is crucial to the treatment, significantly reducing the long-term consequences and possible functional limitations [2]. To achieve this, diagnostic tools are essential to achieve evidence-based decisions. The follow-up is as important as the initial diagnosis, ensuring that the condition of the patient is improved or kept as stable as possible over his/her life. Here, the diagnosis tools also play an important role and will be discussed in more detail later in this document.

3.1.1 Causes

Gout, ankylosing spondylitis, osteoarthritis, Sjogren's syndrome and rheumatoid arthritis are some examples of rheumatic diseases. Their causes are related to lifestyle, age, genetics, autoimmunity, trauma, biochemical abnormalities, inflammation and other factors [23, 24]. Lifestyle is a cause also associated with the age, since a negative lifestyle will accelerate the natural aging processes of the body. For instance, smoking, drinking alcohol and a sedentary lifestyle are potential promoters of early rheumatic problems since they create additional unwanted stress in the body. A patient's profession may also be a risk factor. Whenever a worker needs to perform repetitive tasks or when they have bad posture in the work-place, it is common to have higher incidences of some rheumatic conditions. The previous causes are time-related, and the first signs can take years to manifest; because of that, they are often overlooked. Genetic causes are associated with the individual predisposition to these problems. Here, the patient's family history is important because genetic causes can be traceable to the same disease in other family members. The autoimmune causes are not yet fully understood and are often associated with genetics, where an abnormal response of the immune system triggers an attack on healthy body structures. The trigger for this abnormal response is still to be discovered and is an open research topic. Trauma concerns the conditions

associated with accidents, falls, hits, and their severity is variable. They can be temporary or cause permanent disability or even death. Finally, biochemical abnormalities may trigger rheumatic diseases due to the lack or excess of a given component in the body. Lifestyle can also play an important role here, but it is not the only contributor. One example is that the presence of abnormal quantities of uric acid in the blood may trigger joint pain because of crystal deposits or gout [24].

3.1.2 Epidemiology

Although the associated mortality is low, the mid- and long-term effects of rheumatic diseases are huge. A study [25] concluded that, in the United States of America, rheumatic diseases affect approximately 21.6% of the population, with higher incidence in women than men (25.4% vs. 17.6%). An incidence of 50% was observed in the population over 65 years old. The authors also presented the incidence results for two age groups, those 18-44 years old and 45-64 years old, with incidence values of 7.9% and 29.3%, respectively. Finally, it was also concluded that, of these cases, 34% and 41% experience some physical limitation, which is especially problematic because it affects the active population. In Portugal, there are some projects that have studied the population, such as RheumaPT [26] and EpiReumaPt [27, 28]. These studies suggest that 30% of the Portuguese population have some type of musculoskeletal rheumatic manifestation. Of these, 20% have some significant problem, which means that they are classified as ill, 7% have some degree of incapacity, and 0.5% are dependent on others. The study of Monjardino et. al. [29] concluded that the prevalence of rheumatic diseases in Portugal is between 16% and 24%, affecting more women than men and increasing with age. According to the Second National Program for the Rheumatic Diseases [2], these diseases represent approximately 16%-23% of all General Practice consultations, are the first cause of temporary incapacity, are responsible for 30% of all domicile mobility limitations and 40%-60% of prolonged limitation for daily activities and are the cause of most early retirements (35%-41%). In Table 3.1, the epidemiology of the most frequent rheumatic diseases is shown.

TABLE 3.1: Epidemiology of major rheumatic diseases [30].

Disease	Point prevalence /1000	Incidence /1000	Age Ration (65:25yr)	Gender ratio (F:M)
Rheumatoid arthritis	8.0	0.5	6:1	2.5:1
Juvenile chronic arthritis ^a	0.7	0.1	N/A	2:1-7:1
Osteoarthritis (knee) ^b	100	N/A	0	2:1
Ankylosing spondylitis	2.0	0.07	0	1:3
Systemic lupus erythematosus	0.4	0.05	1.5:1	3:1-9:1
Systemic sclerosis	0.1	0.01	3:1	4:1
Gout	N/A	1.0	2:1	1:6

^a Children <15yr^b Prevalence among persons aged 35-74 yr.

3.2 Rheumatoid Arthritis

Rheumatoid arthritis (RA) is a chronic, systemic, autoimmune disorder that manifests itself as a disabling inflammatory activity, mostly in the synovial joints. This inflammatory activity has the long-term effect of joint destruction, which, in later stages, creates high disability (Fig. 3.1). The cause of this disease is not yet fully understood, but it is known that it involves the immune system targeting healthy joints, resulting in the thickening of the joint capsule and the destruction of the cartilage and bone. It affects from 0.5% to 1.0% of the population, with a higher incidence in women than in men (approximately 3 to 5 times more often) [31]. The first manifestations are mostly between 40 and 50 years of age in women and later in men. Given its progressive nature, it is crucial to achieve early detection so that the inflammatory activity is reduced and the damage to the joints minimized.



FIGURE 3.1: Example of the progression of Rheumatoid Arthritis on the hands.

Other complications may also appear due to RA [32]. Because some of the medications used target the immune system, it gets less effective, and the patients are more prone to infections, which are among the main causes of death in people with RA. A higher risk of heart attacks and strokes is also associated with the pericardium being attacked by the RA. Anemia, osteoporosis, leg ulcers, and pleural effusions are also possible complications [32]. They all lead to a loss in a patient's quality of life, which is also a trigger for depression, which is increased in these patients.

3.2.1 Diagnosis

Medical diagnosis is a difficult task involving the analysis of different sources of information. The rheumatology field is no different, and the gathering of information is fundamental for a good diagnosis and follow-up for the diseases. Regarding RA, early diagnosis is crucial for a proper follow-up to slow the disease progression. The challenge lies in the identification of the subtle changes that the RA causes in the patient. To help the rheumatologist in this task, the American College of Rheumatology (ACR) and the European League against Rheumatism (EULAR) created a set of guidelines in 2010. In Table 3.2, a short version of the criteria is presented:

From Table 3.2 different sources of information can be seen. Some are objective, such as the *serology*, and others are more subjective, such as the *duration of symptoms*. The *joint involvement* is dependent on the rheumatologist's experience and the patient deployment. Given its importance, the use of imaging techniques is advised so that a more precise diagnosis can be obtained. Note that the joint involvement can add up to 5 to the score, which is especially important since 6 is the threshold for a decision. Among all joints, it is also possible to verify that the small joints are more important than the large ones. An example is that a large joint adds 0 to the score, while a small joint can add 2.

The criteria presented in Table 3.2 are enabled by the diagnosis tools available to the rheumatologist. In general, they can be divided into three categories, each with a specific role and importance: physical exams, imaging techniques and sample-based analysis. Next, each one will be addressed in more detail:

TABLE 3.2: 2010 ACR/EULAR classification criteria for RA [3].

Feature	Details	Score
Joint Involvement	1 large* joint	0
	2-10 large joints	1
	1-3 small** joints (+/- large joints)	2
	4-10 small joints(+/- large joints)	3
	>10 joints (at least 1 small joint)	5
Serology	Negative RF and anti-CCP antibodies	0
	Low [#] positive RF or anti-CCP antibodies	2
	High ^{##} positive RF or anti-CCP antibodies	3
Acute Phase Reactants	Normal CRP and ESR	0
	Abnormal CRP or ESR	1
Duration of Symptoms	<6 weeks	0
	>6 weeks	1

Classification criteria are applicable to patients with synovitis in at least one joint without an alternative clear explanation. Joint involvement includes any tender or swollen joint.

* Large joints refer to shoulders, elbows, hips, knees and ankles.

**Small joints refer to metacarpophalangeal, proximal interphalangeal, second to fifth metatarsophalangeal, thumb interphalangeal joints and wrists.

[#] Low positive refers to values higher than and up to three times the upper limit of normal.

^{##} High positive refers to values three times above the upper limit of normal.

A score of at least 6/10 is needed for classification of a patient as having definite RA.

CCP cyclic citrullinated peptide, CRP C-reactive protein, ESR erythrocyte sedimentation rate, RF rheumatoid factor

Physical exams test the mobility of the joints and the presence of swelling, pain, and other movement characteristics, such as morning stiffness. The rheumatologist also takes into account the patient's history, such as professional or family disease history, since it carries important information. It is the first line of examination and of major importance considering that it creates the first prognosis and will dictate which steps to take next. Its limitations are the subjectivity of the exam since it is heavily dependent on the experience of the rheumatologist. Another limitation is the patient's complaints, since some of the indicators are dependent on their opinions. Pain, discomfort and other sensorial aspects are subjective concepts and hard to quantify for different patients. Nevertheless, it is crucial for the diagnosis and is not replaceable by any other diagnostic tool.

The laboratory tests look for deviations from the normal homeostasis and, the presence and concentration of specific antibodies and other substances in the body fluids, such as blood, urine and synovial fluid. These tests include, for example, the presence of immunoglobulin M-rheumatoid factor, in rheumatoid arthritis and inflammatory analysis, such as the erythrocyte sedimentation rate and the C-reactive protein [4]. The main advantage is that quantitative results are obtained, independent from the patient and medical experience. The disadvantages are that the results are not obtained during the consultation time. This may require the patient to visit several locations, wait for the results and book additional consultations. This not only disrupts their normal personal and professional routines but creates anxiety in the patients since these results can have a huge impact on their lives. Another limitation is the fact that some results are not specific, and conclusions can only be taken if the results are, for example, positive or supported by other evidence [3].

The imaging techniques test the visual aspect of the anatomical structures based on their expected normality pattern [33]. The typical modalities used in rheumatology are radiography, magnetic resonance imaging (MRI) and ultrasound. Radiography is helpful for understanding the internal structure of the patient's body to identify bone pathology and some soft-tissue problems. Its limitations are the use of ionizing radiation and that some indicators, such as synovitis and bone erosion, are only visible in more advanced stages, which makes it more interesting for monitoring than for early diagnosis. The MRI is used to study the presence of pathology in tendons, ligaments and cartilage. It is especially important to study the spine since it can create a 3D model of the anatomical structure. Its disadvantages are the cost of the equipment and the exam, which make it harder to perform in smaller medical institutions or as a mass diagnosis tool. Unlike radiography or CT, it does not use ionizing radiation, but in some cases, intravenous contrast is used, which is prone to other unwanted reactions. Ultrasound imaging is useful for studying soft tissues, specifically the more superficial ones. It is used to identify early synovitis, effusions and bone erosions and abnormal changes in the tendons, muscles and ligaments. The power Doppler mode can also be used to identify blood flow and obtain further information regarding the presence of active inflammation. Further advantages are the low cost of the equipment and exams, which is lower than those of other imaging modalities, such as MRI. It can be performed during the consultation, it is well accepted by patients, and some studies hint that the

results are comparable to MRI and better than radiography [5, 6]. These advantages have increased its acceptance to the point that it is now an important tool for rheumatologists [34, 35], for example, in the assessment of rheumatoid arthritis (RA)[5, 36]. The limitation of an ultrasound exam is that it is complex to perform, both its gesture and image understanding, require intensive training and experience [35]. Computer systems could help doctors in this interpretation by using image processing and analysis techniques, enhancing the visualization, and helping in the identification of structures and extraction of objective clinical measures.

3.2.2 Treatment

RA has no known cure, but some treatments are helpful in slowing the disease progression and stabilizing it. There are different treatment possibilities that need to be discussed with the rheumatologist. They differ considerably in their efficiency, side effects and cost, and the idea is to get to a point of remission or a stable condition with low symptoms so that the patients can perform their daily activities. The ACR and EULAR provide guidelines to help with the treatment planning because a single treatment that works for all patients does not exist. The drug groups used in the management of RA are:

- Analgesics
- Non-steroidal anti-inflammatory drugs
- Disease-modifying anti-inflammatory drugs
- Corticosteroids
- Biologic agents

Disease-modifying antirheumatic drugs (DMARDs) are the most effective, and it is recommended that the treatments start early and aggressively [37]. DMARD is an umbrella term for a somewhat unrelated set of drugs that are known to improve the condition of patients with rheumatoid arthritis. The most-used DMARD is methotrexate, but other alternatives exist, which are normally used when methotrexate does not produce the desired effects. Because they are unrelated, combinations of DMARDs are frequently

used and combined with analgesics to help with pain management. Same patients benefit from the use of non-steroidal anti-inflammatory drugs to reduce the inflammatory activity, but due to their side effects, some precautions must be taken, such as gastro protectants. Corticosteroids are also used to decrease the inflammatory activity, acting quickly and improving the condition of the patients. Some side effects are attributed to this treatment, and oral administration is not advised over long periods of time. It is often used as a primary drug in the beginning of the treatment and as a bridging agent, when a therapy change occurs [38]. An alternative to DMARDs are biologic agents, also known as biosimilar DMARDs, bsDMARDs or simply Biologics. They are protein-based drugs engineered to inhibit or augment part of the immune system [38]. They have a higher probability of secondary effects, such as infections and are more expensive, which justifies their use only in special cases, for which methotrexate and other DMARDs do not produce the desired effects.

3.2.3 Costs

The cost of RA treatment is an important aspect to take into consideration when choosing a treatment. It is estimated that the medication costs can range from \$1,500 to \$30,000 annually for a single patient in the USA. Other expenses might also be associated, such as movement helpers, hospital expenses and missing work days for the patient and family. A study from GBI Research estimated that, in the USA, the RA market might grow from \$6.4 billion in 2013 to \$9.3 billion in 2020. This clearly indicates that the burden of such diseases is not negligible, and because of that, measures to reduce their impact need to be taken. This growth is associated with new and more expensive treatments and with the increase in longevity and negative lifestyles, which promote the occurrence of such diseases.

3.3 Ultrasound for Rheumatoid Arthritis

The use of ultrasounds in rheumatology, and more precisely in the assessment of RA, is a relatively recent topic when compared with other well-established protocols, such as neonatal and cardiovascular assessment. Advances in equipment capabilities, with

the spread of probes with higher frequencies (above 10 MHz), have allowed the acquisition of images with better spatial resolution. These high frequency probes allow the differentiation of the small structures present in the joints [34, 35, 39]. Note that, in the case of Rheumatoid Arthritis assessment, the loss in depth penetration due to the increase in frequency is not a problem since the structures to be analyzed are mostly superficial.

Regarding the use of ultrasounds to infer the presence of RA, some studies were found in the literature. In [40], a scoring system was proposed where the implication of the proximal interphalangeal (PIP) and metacarpophalangeal joints (MCPJ) was studied on the palmar and dorsal sides. The "3 finger method" achieved the best results (using the index, middle and ring fingers), and the palmar acquisitions lead to better sensitivities, while no differences were found between the PIP and MCPJ. Alternately, [41] studied the second MCPJ on 8 different planes and concluded that the longitudinal dorsal was more important than the palmar view. These articles are somehow contradictory and show that the assessment of RA is not yet fully understood. Another finding in MCPJ US images related to RA is bone erosion. The authors in [42] studied the specificity of US to detect bone erosions. The results showed that, although bone erosions are not specific to RA, large erosions in the second and fifth MCPJ, fifth metatarsophalangeal joint (MTPJ) and distal ulna were highly specific to RA assessment.

Based on the literature review and the advice of expert rheumatologists, the longitudinal view of the second MCPJ was selected. Nevertheless, the remaining MCPJ, as well as the PIP joints, are very similar, and thus most of the work proposed here might be easily applicable to them.

3.3.1 Metacarpophalangeal Joint

As previously shown, the small joints play a vital role in the assessment of RA. Among all small joints, the MCPJ was also shown to be an important structure, and thus it was selected for this work as the object of study. In Fig. 3.2, a healthy MCPJ can be visualized.

¹Adapted from: Blausen.com staff. Medical gallery of blausen medical 2014. *WikiJournal of Medicine* 1 (2), 2014

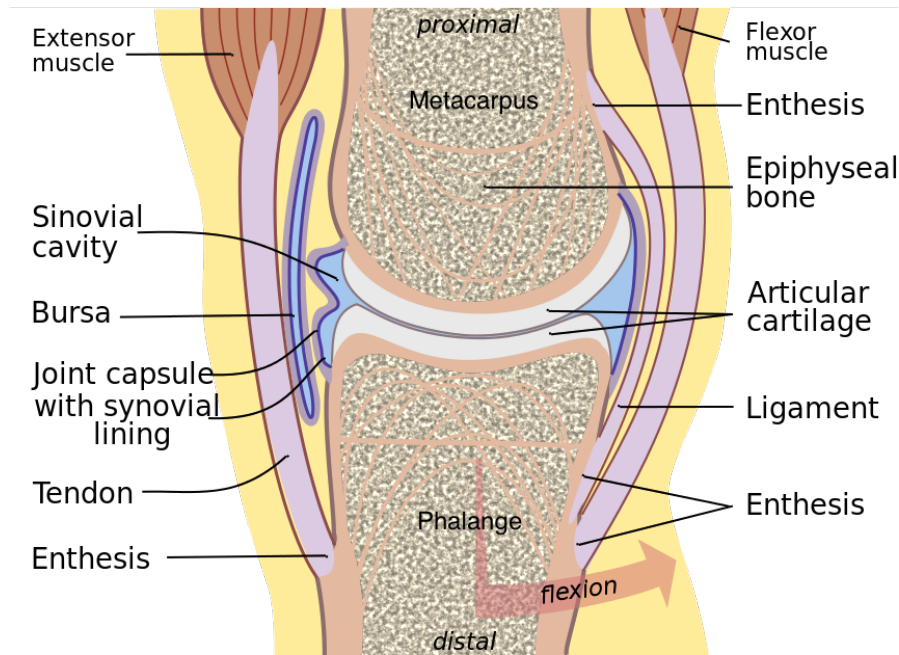


FIGURE 3.2: Illustration of a healthy metacarpophalangeal joint.¹

From Fig. 3.2, it is possible to visualize the metacarpus in the proximal region and the phalanx in the distal. Their main function is to give mechanical support to the hand and finger. The extensor and flexor muscles, together with the tendons, are responsible for the movement of the finger, where the first generates force and the second transfers this force from the muscles to the bones. The ligaments are responsible for limiting the movements so that unwanted movements are prevented. The bursa is a structure that reduces the friction between the tendon and the bones. Finally, the articular cartilage, joint capsule, synovial membrane and synovial fluid all work together to ensure that the bones do not deteriorate with movement by reducing the friction between them. The synovial membrane is responsible for keeping the synovial fluid in good condition, i.e., responsible for its renewal, so that the lubrication between the bones is proper maintained.

One of the indicators of RA is the presence of inflammatory activity in the synovial membrane, known as synovitis [4]. In Fig. 3.3, it is possible to see a representation of a normal joint and a joint affected by RA.

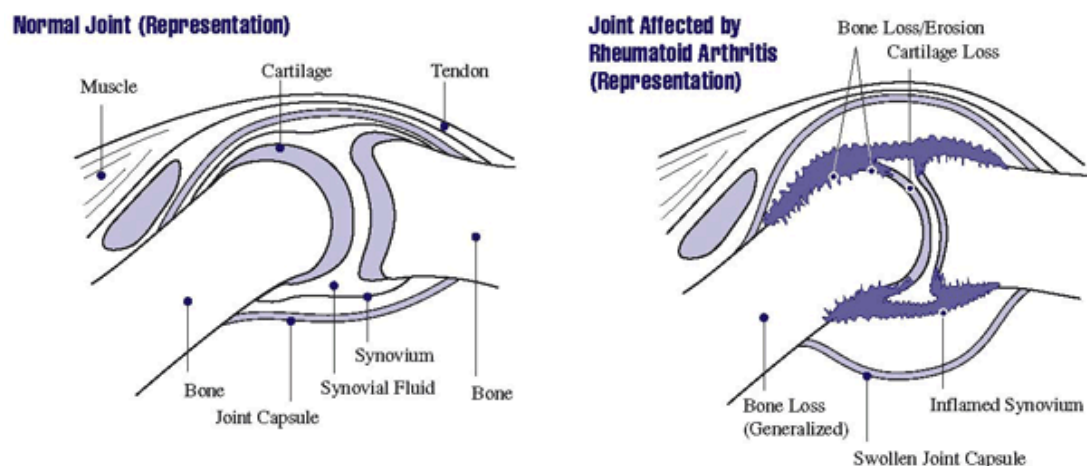


FIGURE 3.3: Illustration of a normal joint and a joint with synovitis.²

In RA, the immune system attacks the synovial membrane, making it thicker and compromising the normal lubrication process. The first effect is that the joint movement is less efficient due to the increase of the friction between the bones, which promotes cartilage deterioration. The other effect is the joint capsule swelling due to the accumulation of synovial fluid, which creates a distention of the joint capsule. [4]. These problems have long-term effects as already discussed in this chapter.

3.3.2 Selected Acquisition Protocol

The study of the MCPJ using ultrasound images can follow different acquisition protocols. In this work, the dorsal longitudinal view was selected, because it allows the visualization of the joint capsule, cartilage and possible bone erosions and other tendon problems, in a single acquisition. In Fig. 3.4, the probe positioning and the resultant image are shown.

The metacarpus are the bones of the hand that are connected to the first bone of the fingers, called phalanges and have high brightness in the ultrasound images. The skin line is the most external organ and the first protection layer of the hand, and it also has a bright appearance with a darker region below it due to the presence of a fat layer. The extensor tendon, responsible for the opening of the hand, is situated between the bones and the skin and is darker than its surroundings. Between the metacarpus, phalange and extensor tendon is the joint capsule, responsible for the friction reduction of the

²Adapted from: February 2016. URL https://www.niams.nih.gov/Health_Info/Rheumatic_Disease/default.asp

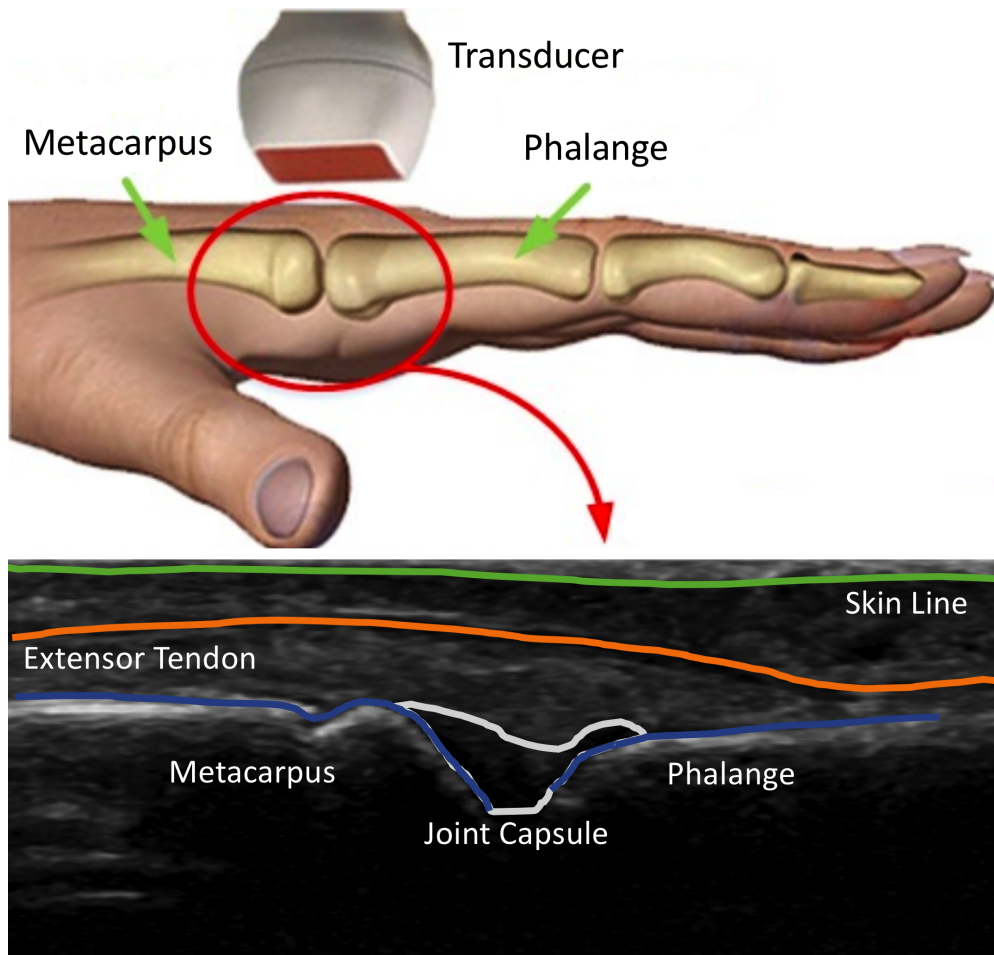


FIGURE 3.4: Representation of the second MCPJ ultrasound image acquisition.
Adapted from [12]

fingers due to their natural movement. In the ultrasound images, it is not always visible, but when synovitis is present, it is usually visible as a darker structure. In the head of the metacarpus, it is also possible to see a small darker region, which is the cartilage.

The use of the power Doppler mode is also important in the analyses of synovitis. A high Doppler signal is indicative of inflammation activity since more blood movement is involved in these cases. Our choice in this work was to focus on the B-mode ultrasound images, but power Doppler should be considered by future computer vision researchers given its usefulness to rheumatologists.

Chapter 4

Image Processing and Analysis Background

Contents

4.1	Ultrasound Imaging Existing Work	32
4.1.1	Pre-Processing	32
4.1.2	Segmentation	33
4.1.3	Feature Extraction and Classification	34
4.2	Active Contours Model	35
4.2.1	Internal Energy	36
4.2.2	External Energy	36
4.2.3	Optimization	37
4.3	Localizing Active Contours	37
4.4	Phase Symmetry	39
4.5	Simple Linear Iterative Clustering	41
4.6	Convolutional Neural Networks	42
4.7	Performance Metrics	45
4.7.1	Modified Hausdorff Distance	45
4.7.2	Dice Similarity Coefficient	47

4.1 Ultrasound Imaging Existing Work

As stated in Chapter 2, ultrasound images are usually corrupted by noise and artifacts. This characteristic increases the interest in image processing techniques that allow better image interpretation [44]. Several areas already use image processing techniques with success, such as echocardiography, breast ultrasound, vascular ultrasound and gynecological and obstetric ultrasound [21, 45–48]. The typical image processing pipeline can be divided into four main steps: pre-processing; segmentation; feature extraction and classification. The literature review, which will be presented next, follows this division.

4.1.1 Pre-Processing

Image pre-processing can be defined in several ways, but the main idea is to start from an image and obtain a new image with less variability and greater information availability with the aim to improve and simplify the overall processing pipeline. This includes cropping, color/gray-scale rearrangements, noise removal, and others. Because ultrasound images are frequently corrupted by noise, it will be the focus of this subsection. The main type of noise found in ultrasound images is speckle. Some studies suggest that it can be modeled as a Rayleigh distribution [44, 49], while others suggest that Gamma or Fisher-Tippett distributions are good approximations [21, 50, 51], or even a multiplicative model [52]. Several methods to remove speckle have been proposed: some use the wavelet transform [53–56], others are based on anisotropic diffusion [57–59], and others are based on partial differential equations [60]. In [61], three pre-processing methods were tested in the classification of ultrasound kidney images: the median filter, wiener filter and histogram equalization. The results showed that the median filter was the best among these three.

In [62], a modeling system is proposed to recover the radio frequency original signal from the log-compressed ultrasound images outputted by the equipment systems. The authors argue that the post-processing, used for display, results in better looking images but compromise the use of realistic models in image analysis. It is reported that, with the proposed framework, the observed data is better represented for modeling purposes using a Rayleigh distribution.

In [63], a decomposition method for ultrasound images is proposed, where the input image is separated into two distinct images, the de-speckle and the speckle image. The authors argue that the speckle, often considered as noise, carries information regarding the underlying tissues scanned with the ultrasounds and should be considered as signal instead of noise.

4.1.2 Segmentation

Regarding the segmentation of structures in musculoskeletal ultrasound images, several works can be found in the literature. [64] proposed a bone fracture detection system where the segmentation of the bone was achieved using phase symmetry and Log-Gabor 2D filters. These filters are used to find local properties of the image by calculating the signal magnitude and phase at a given scale and location. The authors in [65] developed a bone contour detection algorithm based on ultrasound principles. Each pixel is mapped to a probability of belonging to the bone surface, whose value is based on the combination of the shadow effect and reflection of the bone. After that, a line model is created where the probabilities are used to minimize a cost function that converges to the bone surface using dynamic programming techniques. Because the ultrasound images are very granular and the contours are not easily identifiable, some authors have noted that segmentation methods that incorporate prior knowledge, such as shape, movement and localization, achieve better results. In [66], the use of a Bayesian approach is proposed to segment the prostate, while other contributions used active contours and active shape models [67, 68], as well as level-sets [69]. Other segmentation techniques are based on machine learning methods, where a model is trained with known cases and later used for unknown images. Some works have been published in this field using AdaBoost, Probabilistic Boosting Trees and Random Forests [45, 70]. The use of Deep Learning (DL) approaches is increasing in all fields, and in a survey conducted by [71], some of the most recent developments in the medical field were addressed. There, the authors divided the approaches into several categories, and regarding the musculoskeletal field, only one work was presented, focusing on ultrasound images [72], while the remaining ones focused on MRI, CT and X-ray images. The field in which ultrasound imaging and DL are more explored is the

analysis of cardiac images [73–75]. In [76], CNN's were used to identify the edges in musculoskeletal ultrasound images of the shoulder.

4.1.3 Feature Extraction and Classification

The detection and quantification of findings in musculoskeletal ultrasound images is a relatively recent topic when compared with other well-established ultrasound acquisition protocols. Regarding the detection of findings, several works used first-order statistics, Haralick and Law features, Hu moments, wavelet- and Fourier-based features and others [61, 77–81]. In these works, Neural Networks and/or SVM were used to create the decision models, with no apparent advantage or significant preference for one or the other. These techniques were used to propose solutions in kidney, heart and liver ultrasound images.

In [61], a classification system for kidney ultrasound images is proposed. It assumes a 5-class classification problem (normal, cyst, stone, tumor and failure) and starts with a pre-processing step to remove noise and reduce the variability between images. The authors compare two features sets, one based on the Haralick features [82] and the other using wavelet features. The principal component analysis (PCA) is used in the wavelet features to reduce the dimensionality of the feature vector. Finally, a Neural Network is used to create the decision model. A correct classification rate of 97% was achieved using the wavelet-based features, which outperforms the 95% achieved with the Haralick features.

A more recent work proposes the use of CNNs to automatically characterize plaque composition in carotid ultrasound images [83]. Approximately 90,000 patches with 15x15 pixel sizes were used to create the model. The authors reported a correlation of 0.9 with the clinical assessment for the estimation of the lipid core, fibrous cap and calcified tissue areas.

In [81], an ultrasound imaging system for fatty and normal liver classification is proposed. The authors reported 95% accuracy and 100% sensitivity using a pre-processing step to recover an image with the speckle information and using the Haar wavelet decomposition on that speckle image.

Regarding the musculoskeletal ultrasound classification methods, [84] proposed a carpal tunnel syndrome classification system and compared electromyography with ultrasound imaging. The author concluded that the first is better, but the ultrasound achieved comparable results, which makes it an interesting alternative for daily use.

4.2 Active Contours Model

The active contours model is a well-known framework used to solve image processing problems, such as segmentation and tracking. The algorithm looks for a set of connected points, v_i , that define a line around the object of interest. These points are allowed to move under certain restrictions, which are referred to as energies. The objective is to find the the points v_i that minimize the energy, $E(v_i)$:

$$\arg \min_{v_i} E(v_i) \quad (4.1)$$

There are several ways to formulate the energies, but they are usually divided into two main groups. The internal energy, E_{int} , referring to the line itself, is responsible for imposing the expected shape of the object, and the external energy, E_{ext} , referring to the image, I , is responsible for the adaptation of the expected shape of the object to the image. Equation 4.2 presents the general energy measurement equation using the continuous representation of v_i as $v(s)$.

$$E(v(s)) = \int E_{int}(v(s)) ds + \int E_{ext}(I, v(s)) ds \quad (4.2)$$

Note that, to obtain $v(s)$ from v_i , we need to define an interpolation model. The typical method is piece wise linear interpolation, but other methods can be used as well, such as polynomial and spline interpolation. Next, the internal and external energy calculation will be explained, and possible solutions to optimize 4.1 will be presented as well.

4.2.1 Internal Energy

The internal energy, E_{int} , refers to the energies that only take into consideration the solution, $v(s)$. The most common way to express the internal energy is by using the first and second derivative of $v(s)$, which are respectively known as the continuity and the curvature:

$$E_{int}(v(s)) = \alpha(s) \left\| \frac{dv}{ds}(s) \right\|^2 + \beta(s) \left\| \frac{d^2v}{d^2s}(s) \right\|^2 \quad (4.3)$$

where $\alpha(s)$ and $\beta(s)$ are weight values used to control the importance of the continuity and the curvature, respectively. The continuity ensures that a continuous line is obtained, instead of scattered points, by limiting the distance between successive points. The curvature penalizes oscillations, making the line more or less smooth.

4.2.2 External Energy

External Energy refers to the energies that take into consideration the solution, $v(s)$, and the information present in the image, I . The most common external energies are related to the image intensity, edges or other pre-processing techniques:

$$E_{ext}(I, v(s)) = w_L E_{Line}(I, v(s)) + w_E E_{Edge}(I, v(s)) + w_F E_{Filter}(I, v(s)) \quad (4.4)$$

where w_L , w_E and w_F are, respectively, the weights of the line, edge and filter energies used to control the relative importance of each. The E_{Filter} component is a generic formulation of E_{Edge} , where the filter is chosen for the specific problem. The E_{Edge} energy is obtained from the application of an edge detector to the image. In this field, the number of possible formulations is huge and dependent on the image characteristics, but the main idea is to enhance the objects of interest by creating regions of low values that $v(s)$ should move to.

4.2.3 Optimization

After the energies formulations, equation 4.1 needs to be solved to obtain the final solution. Gradient descent based methods are the most common approaches [85]. They use the derivatives of the energies functions to conduct a pre-defined coarse segmentation to a local minimum, taking into consideration that, at the minimum, the gradient should be zero. Other methods use the Euler-Lagrange equation, which is also based on the energies derivatives, but with a different formulation [85]. Both methods work well in cases where the number of local minimums is small or the confidence in the initialization is good. However, these two characteristics are not always met, so other optimization methods or strategies might be needed. For these cases, Grid search, Monte Carlo and Evolutionary Algorithms [86] are possible solutions since no precise initialization is needed and local minimums can be avoided to a certain degree.

4.3 Localizing Active Contours

The Localizing Active Contours technique (LAC) was proposed by [87], and like most active contours methods, it starts from a coarse segmentation and iteratively, converges to a more precise segmentation. In the LAC method, the energies are calculated locally, allowing a natural adaptation to background non-uniformities. Following the definition of [87], the localizing active contour model is defined as the sum of all the local energies along the contour line:

$$E(\phi) = \int_{\Omega_x} \delta\phi(x) \int_{\Omega_y} B(x, y) \cdot F(I(y), \phi(y)) dy dx + \lambda \int_{\Omega_x} \delta\phi(x) \|\nabla\phi(x)\| dx. \quad (4.5)$$

where, $\delta\phi(x)$ is a portion of the line centered in x , $B(x, y)$ is the area around the point (x, y) and is used to define the localizing area. F is the internal energy function, and the last term refers to the continuity of the contour line and is scaled by λ . To help understand each variable, a representation of each is presented in Fig. 4.1.

To use gradient descent methods, the derivative of the energy with respect to ϕ is used to obtain an evolution equation:

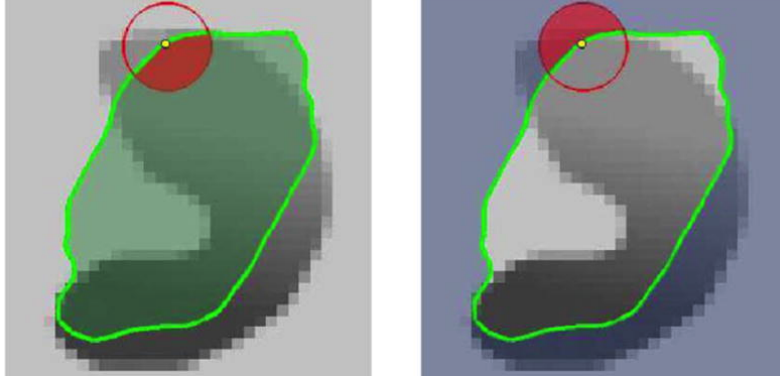


FIGURE 4.1: Representation of the Localizing Active Contour algorithm; In green is the contour line, that separates the interior and exterior of the image. The red circle is the neighborhood, B , of the point (x, y) , which is represented in yellow.

$$\frac{\partial \phi}{\partial t}(x) = \delta \phi(x) \int_{\Omega_y} B(x, y) \cdot \nabla_{\phi(y)} F(I(y), \phi(y)) dy + \lambda \delta \phi(x) \operatorname{div} \left(\frac{\nabla \phi(x)}{|\nabla \phi(x)|} \right). \quad (4.6)$$

The internal energy, F , used was the Uniform Model (UM) defined as:

$$F_{UM} = H\phi(y) (I(y) - u_x)^2 + (1 - H\phi(y)) (I(y) - v_x)^2 \quad (4.7)$$

where u_x and v_x are, respectively, the mean intensity of the exterior ($1 - H\phi(y)$) and the interior ($H\phi(y)$) of the contour in a given point neighborhood:

$$u_x = \frac{\int_{\Omega_y} B(x, y) \cdot H\phi(y) \cdot I(y) dy}{\int_{\Omega_y} B(x, y) \cdot H\phi(y) dy} \quad (4.8)$$

$$v_x = \frac{\int_{\Omega_y} B(x, y) \cdot (1 - H\phi(y)) \cdot I(y) dy}{\int_{\Omega_y} B(x, y) \cdot (1 - H\phi(y)) dy} \quad (4.9)$$

For more implementation details or other internal energy formulations please refer to [87].

4.4 Phase Symmetry

Symmetry is an important characteristic of objects and signals. The general definition states that an object is symmetric if it remains invariant under some transformation. The most common case is bilateral symmetry, where one part of the object is equal to the reflection of the other part, as shown in Fig. 4.2.

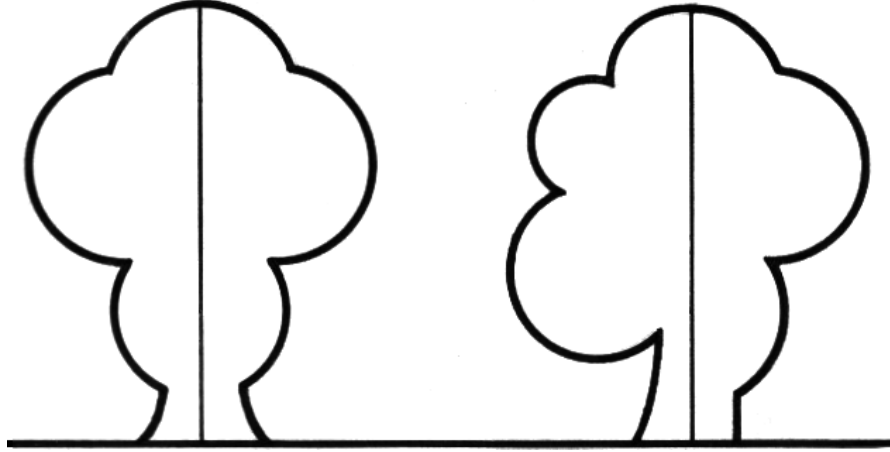


FIGURE 4.2: Example of a symmetric object in the left and an asymmetric object in the right.

The problem with finding symmetry in an object is that the object needs to be identified first. In [88], a method for symmetry identification without prior segmentation was proposed. The main idea is to use wavelet filters to study the signal in the phase domain. These filters must preserve the phase information and, therefore, must be an even-symmetric odd-symmetric pair (low-pass and high-pass, respectively). The complex-valued Gabor Functions wavelets, or more precisely, the LogGabor functions, meet this requirement and were used in this work. Let S denote a 1D signal and M_n^e and M_n^o denote even-symmetric and odd-symmetric wavelet filters at scale n . The convolution of the signal with the filters is represented as:

$$[e_n(x), o_n(x)] = [S(x) * M_n^e, S(x) * M_n^o] \quad (4.10)$$

where $e_n(x)$ and $o_n(x)$ are, respectively, the even and odd responses. Assuming that they are orthogonal to each other, the first being the real and the second the imaginary, it is possible to calculate the amplitude $A_n(x)$ and the phase $\Phi_n(x)$ as follows:

$$A_n(x) = \sqrt{e_n(x)^2 + o_n(x)^2} \quad (4.11)$$

and

$$\Phi_n(x) = \text{atan2}(e_n(x), o_n(x)) \quad (4.12)$$

For the case of a symmetric object, the amplitude response of the even filter is expected to be high and the amplitude response of the odd filter low, which in practice implies that the phase, Φ , will control the proximity of the response to the even or odd component as shown in Fig. 4.3.

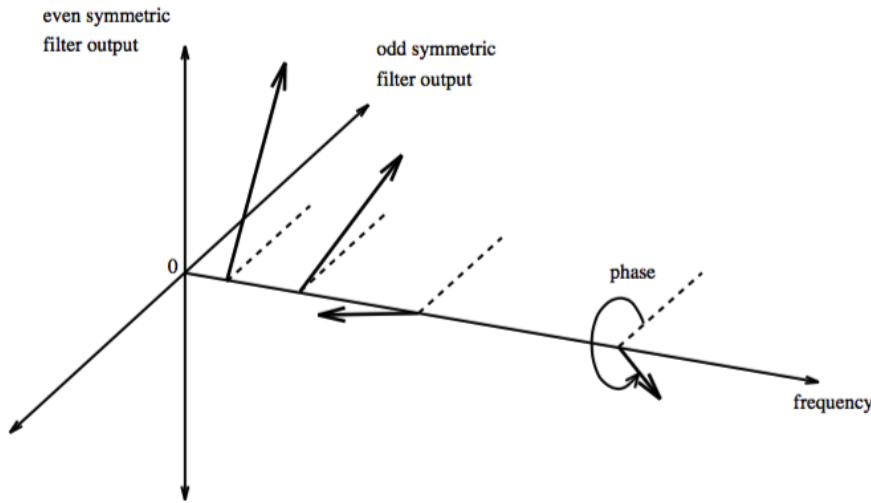


FIGURE 4.3: Representation of the odd and even orthogonal filters. When an object is perfectly symmetric the response will be a vector parallel to the even axis, and the other way around for the asymmetric case. Note that the frequency axis is related to the wavelet multi-resolution

Thus it is possible to use the cosine and the sine of the phase to measure the symmetry, that is, subtract the absolute value of cosine from the sine. To include the multi-scale component, n , the weighted sum of all scales is done, resulting in equation 4.13.

$$Sym(x) = \frac{\sum_n [A_n(x)[|\cos(\Phi_n(x))| - |\sin(\Phi_n(x))|] - T}{\sum_n A_n(x) + \epsilon} \quad (4.13)$$

After Simplification¹, the final equation is obtained:

¹ $\cos(\text{atan2}(x, y)) = \frac{x}{\sqrt{x^2 + y^2}}$ and $\sin(\text{atan2}(x, y)) = \frac{y}{\sqrt{x^2 + y^2}}$

$$Sym(x) = \frac{\sum_n [|e_n(x)| - |o_n(x)|] - T}{\sum_n A_n(x) + \epsilon} \quad (4.14)$$

Here, a new term, T , is introduced that acts as a noise reduction component, and its value should be equal to the maximum noise response. To extend 4.14 from 1D signals to 2D, the 1D approach is applied to multiple orientations and weight sum all in the same manner used in the scale. Hence, it is also possible to control in which directions test the symmetry. This property is particularly interesting in ultrasound images since they tend to have more horizontal structures, and consequently, the vertical symmetry is more important than the horizontal.

4.5 Simple Linear Iterative Clustering

The Simple Linear Iterative Clustering (SLIC) method was proposed by [89] and consists of an adaptation of the k-means algorithm that uses the intensity and spatial information, ensuring connectivity between the pixels in each cluster, which are referred to as super-pixels. The term super-pixel is used to identify a set or cluster of pixels with similar characteristics. Because they can incorporate several pixels, they are a more compact representation of the image, reducing redundancy and possibly accelerating existing algorithms.

The SLIC algorithm was first proposed to process color images in the CIELAB color space, but the extension to gray-scale images is trivial. The algorithm starts with an initialization phase similar to k-means, where a set of k centroids, $C_i = [l_i \ a_i \ b_i \ x_i \ y_i]^T$, is calculated from the input image. Here, the first three coordinates are the CIELAB color centroid, and the last two are the spatial centroid. The samples are extracted using a grid approach with a step of $S = \sqrt{\frac{N}{k}}$, where N is the number of pixels in the image. Next, a rearrangement of the centroids is performed in which the centroid is moved to the lowest gradient position in a 3x3 neighborhood. This step reduces the noise sensitivity of the algorithm. Next, each pixel is iteratively assigned to the closest centroid, and the centroids are recomputed at the end of each iteration in a manner similar to k-means. The main differences lie in the search space and the distance metric. The search space for each cluster is limited to a window of size $2S \times 2S$, ensuring that the

pixels assigned to the centroid are in a controlled neighborhood and considerably reducing the computational time. The distance metric incorporates the spatial distance, 4.15, and the color distance, 4.16:

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (4.15)$$

$$d_c = \sqrt{(l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2} \quad (4.16)$$

where j and i refer to the pixels and the centroids, respectively. The distance, D , is the combination of these two distances:

$$D = \sqrt{\left(\frac{d_c}{m}\right)^2 + \left(\frac{d_s}{S}\right)^2} \quad (4.17)$$

where m is a parameter that controls the relative importance of d_s and d_c . If m is small, the color prevails over the spatial distance, making the super-pixels less regular. If m is large, the super-pixels get more regular but less uniform in terms of color. When using the CIELAB color space, values between 1 and 40 are advised for m . Equation 4.16 can be adapted from the CIELAB color space to gray-scale by changing d_c to:

$$d_c = \sqrt{(g_j - g_i)^2} \quad (4.18)$$

where g is the gray level at the respective indexes of j and i .

4.6 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are the current state-of-the-art approach for solving complex classification, identification and segmentation tasks. The concept is an extension of the well-known Neural Network framework, and it uses the same principles and tools, such as inspiring in the human brain and multiple layers of artificial neurons. One key aspect of the CNNs is that the numbers of layers and neurons are much larger than those typically used in Neural Networks. Because of that, CNNs are

able to create complex nonlinear relationships between the input data and achieve results comparable to humans. CNN models can take several forms, but they all use the same basic modules, which are the convolutional layers. These layers have special properties that are interesting for image processing tasks. They replace the typical fully connected layers, allowing the models to have much fewer parameters and therefore a more efficient and easy training. In Fig. 4.4, the difference between a fully connected and a convolutional layer is visualized.

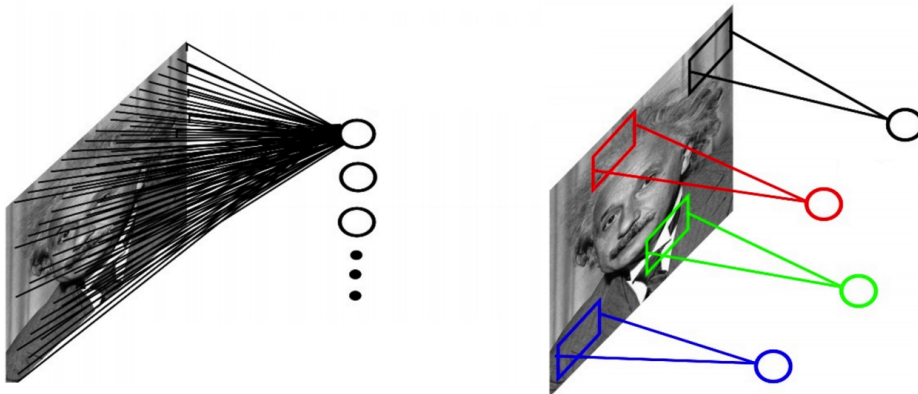


FIGURE 4.4: Difference between a fully connected layer (left) and a convolutional layer (right).

In a convolutional layer, each neuron only looks to a limited portion of the image at a time, while a fully connected layer looks to all pixels at once. This difference implies that convolutional layers are only able to capture local information while fully connected layers capture global information. To overcome this limitation and obtain global relations between the pixels (context), several convolutional layers are concatenated, which increases the receptive field of deeper layers in relation to the input image, as depicted in Fig. 4.5.

This concatenation has an additional effect of successively capturing more complex information from the image. In fact, it is possible to observe that the first layers capture low-level features, and deep layers capture high-level features Fig. 4.6.

From Fig. 4.6, it is possible to see that, in the first layers, the information extracted is mostly based on gradients, and as successive layers are concatenated, more complex information is obtained. Next, a brief explanation of the main building blocks of a CNN model will be presented.

CNN building blocks

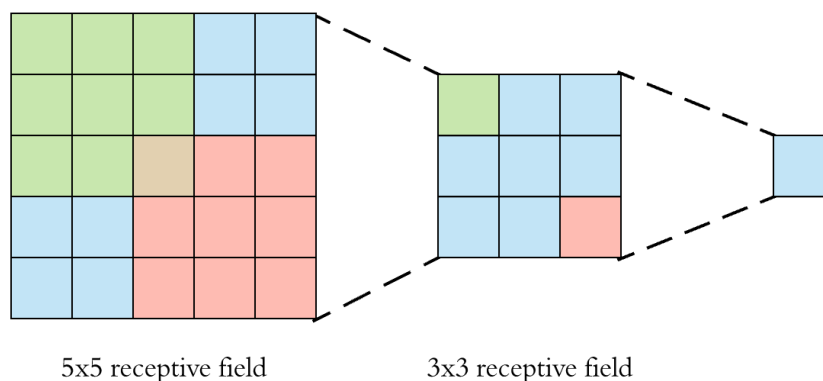


FIGURE 4.5: Example of receptive field growth in deeper layers.

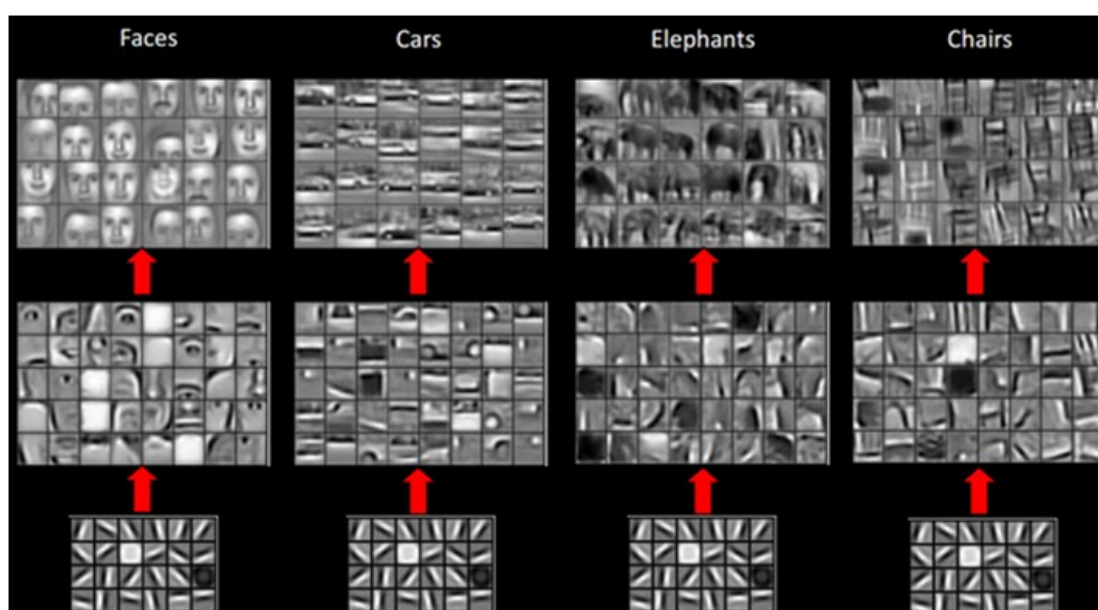


FIGURE 4.6: Example of the low level and high-level features learn in shallow and deep layers (from the bottom to the top).

To create a CNN model, there are several blocks that can be used to solve each specific problem. In this subsection, a brief introduction to the most common blocks is presented.

Convolutional Layer: this layer is the most important one in CNNs, and it is responsible for the information enhancement observed in Fig. 4.6. They have two main parameters and several secondary parameters. The two main parameters are the convolutional filter size and the number of neurons. The first is used to define the receptive field of each neuron and is normally set to a small odd number, and the second is dependent on the application but is typically set to a base-2 number. The other secondary parameters

are the stride, which defines the jumps that the neuron does in the convolutional operation, and typical values are 1 and 2. The use of padding so that the dimensions are preserved and the use of dilated filters to increase the receptive field without increasing the number of parameters are also used in this type of layer.

Activation Layer: this layer adds non-linearity to the model. Note that a set of successive linear operations result in a linear operation, and thus by introducing a non-linear operation, complex relations between components can be created. Common activation functions are:

- Rectified Linear Unit - Relu
- Exponential Linear Unit - Elu
- Sigmoid
- Softmax

Pooling Layer: this layer samples the resulting features from the convolutional layers so that the number of parameters is decreased, as well as the model complexity. This layer also helps increase the receptive field of the next neurons, because it compresses local information. Typical pooling layers are the max-pooling and average pooling with a down-scaling factor of 2.

Batch Normalization: this layer is used to normalize the features created by the convolutional layers. This allows the inputs of the following layers to have standardized input values, which helps in the convergence process, making it more stable and faster.

4.7 Performance Metrics

4.7.1 Modified Hausdorff Distance

The Hausdorff Distance is a metric used to quantify the closeness between two lines, planes or sets of points. It is defined as the maximum minimum distance between two sets, A and B , as shown in Fig. 4.7:

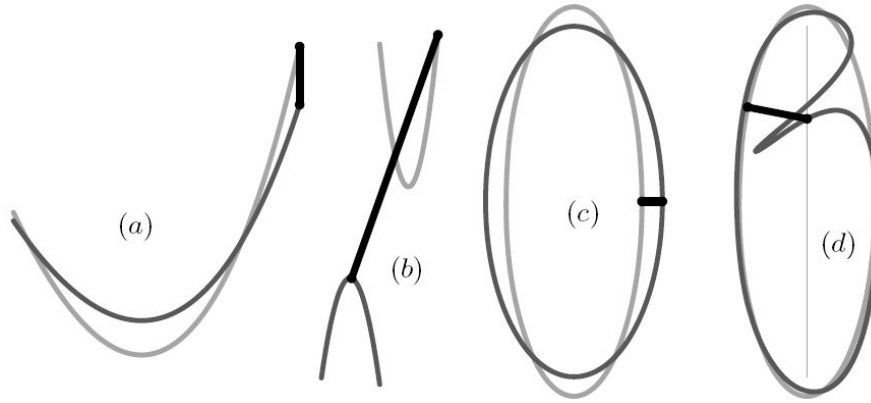


FIGURE 4.7: Examples of the Hausdorff distance between two lines. Black - Hausdorff distances; Gray - Lines. Adapted from [90]

Mathematically, it is defined as:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (4.19)$$

with:

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (4.20)$$

where $\|\cdot\|$ is the Euclidean distance, between two points.

One limitation of the Hausdorff Distance is that it is highly sensitive to outliers. To overcome this limitation, other variations of the Hausdorff Distance have been proposed. Here, the Modified Hausdorff Distance (MHD) [91] will be presented in more detail.

As previously said, the MHD [91] aims to reduce the impact of outliers in the tested sets, and thus $h(A, B)$ was replaced with $h_{mod}(A, B)$, which is defined as:

$$h_{mod}(A, B) = \frac{1}{|A|} \sum_{a \in A} \min_{b \in B} \|a - b\| \quad (4.21)$$

In this manner, instead of the maximum minimum distance between the point sets, the MHD calculates the average minimum distance between points, which reduces the impact of a single point since all points are taken into account.

4.7.2 Dice Similarity Coefficient

The Dice Coefficient is a metric that quantifies the overlap between two regions. It is commonly used in segmentation problems to compare the performance of different algorithms and parameterizations. Given two regions, A and B , the Dice Coefficient, $DICE$, is defined as:

$$DICE(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (4.22)$$

From 4.22, it is possible to conclude that, if the regions are almost coincident, $DICE$ will be 1, and if they are totally separated, $DICE$ will be 0. Fig. 4.8 provides a visual representation of these special cases.

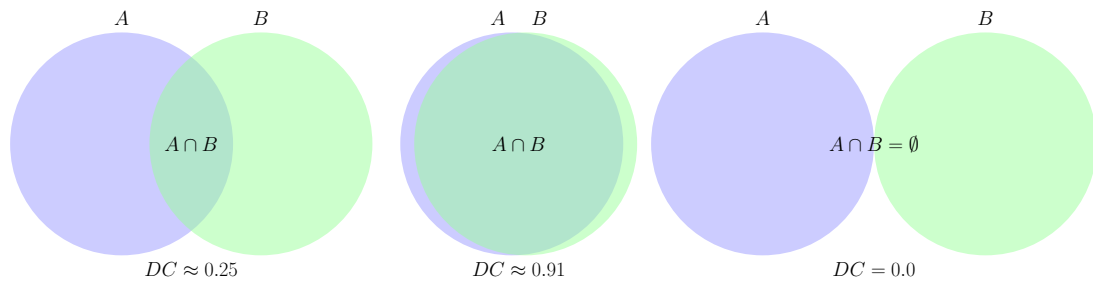


FIGURE 4.8: Visual representation of the Dice Coefficient in three different cases.

Chapter 5

Metacarpus and Phalange Identification

Contents

5.1	Introduction	49
5.2	Proposed Work	51
5.2.1	Pre-processing	52
5.2.2	Contour Initialization	52
5.2.3	Contour Refinement - LAC	52
5.2.4	Upper Line Extraction	53
5.3	Results and Discussion	54
5.4	Complementary Results	55
5.5	Conclusions	57

5.1 Introduction

In this chapter, the segmentation of the metacarpus and phalange bones is presented, which is the first contribution of this thesis. This chapter is mostly based on the work published in [9], along with a new contribution involving additional extended results over an enlarged dataset. In ultrasound images, the only visible part of the bones is their surface. Due to their high density when compared to the surrounding structures, they reflect all the sound waves that reach them. Therefore, the only information that

might appear below the bones is noise or artifacts. This is an interesting property since it allows for the reduction of the search zone of other structures once the bone surface is detected. Moreover, correct identification of the bones can be used to extract information regarding the presence of bone erosion, osteophytes or other bone-related conditions visible on their surface. Fig. 5.1 presents some examples of ultrasound images of the dorsal view of the second MCPJ and the expected challenges of the segmentation of the metacarpus and phalange.

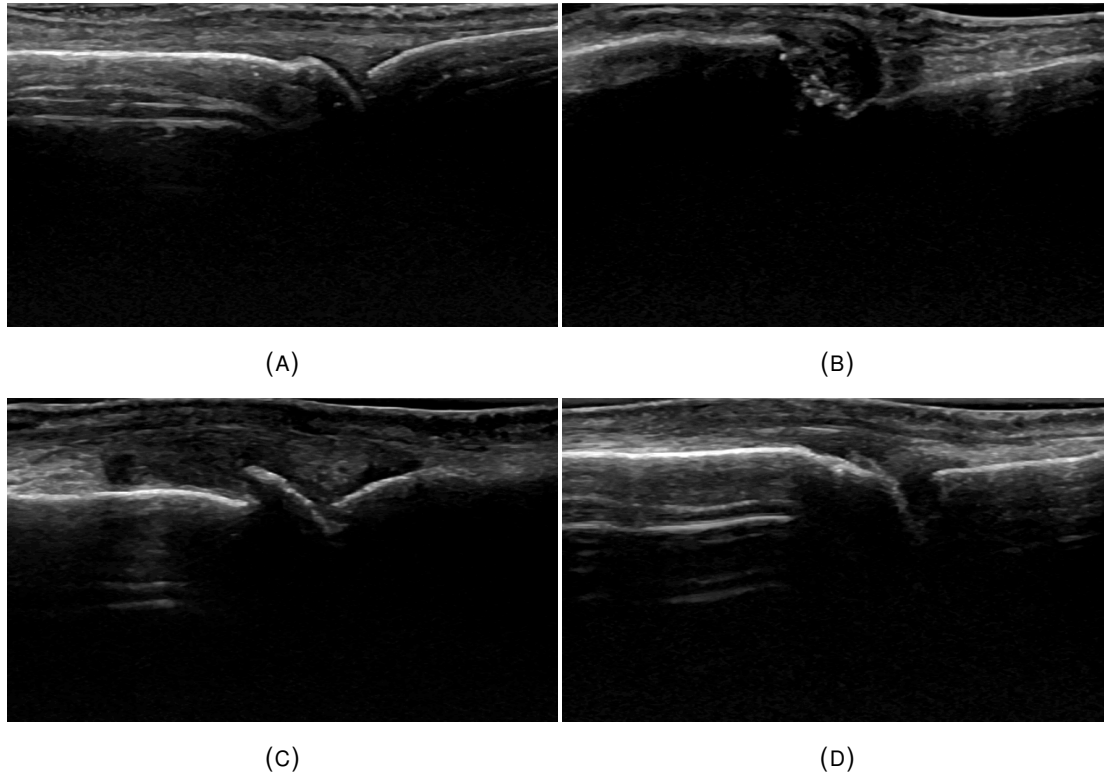


FIGURE 5.1: Examples of ultrasound images of the dorsal view of the second metacarpophalangeal joint. A) normal image; B) bone erosion with subluxation; C) osteophyte; D) reverberation artifact.

Findings, such as osteophytes and erosions, create discontinuities in the bone surface, the presence of reverberation can create structures similar to the bones, and subluxations change their normal and expected position. An automatic tool, such as the one proposed here, needs to be able to handle these findings and achieve results similar to those achieved in normal images. In the case of anatomical findings, a proper segmentation of the bones can be a major step in the characterization of the patient's condition (e.g., the quantification of missing bone when bone erosion is present). Previous works proposed some solution for this problem: In [92], the author proposes the use of confidence maps improved with the use of a Trimmed Non-Local Means filtering to improve

the results of a conventional segmentation method. The confidence maps are directly related to the likelihood of the ultrasound being transmitted and reflected along the tissues, and trimmed Non-Local Means filtering was used to reduce the image noise and improve the confidence maps results. The segmentation is then achieved using the well-known Otsu method on the confidence map, followed by morphologic operations on the resulting blobs to improve the contours. In [11], the log Gabor filter is used to reduce noise and enhance ridge-like structures (bones), and a global threshold is then used to extract the bone candidates. Using prior knowledge of the bone shapes, a score system is used to remove outliers. These two last works are the closest found in the literature, addressing the segmentation problem of the metacarpus and phalange.

5.2 Proposed Work

The segmentation of the metacarpus and the phalange was divided into 4 main steps, as depicted in Fig. 5.2: pre-processing, contour initialization, the Localizing Active Contours (LAC) algorithm, and finally, upper line extraction. Next, each step will be addressed in detail.

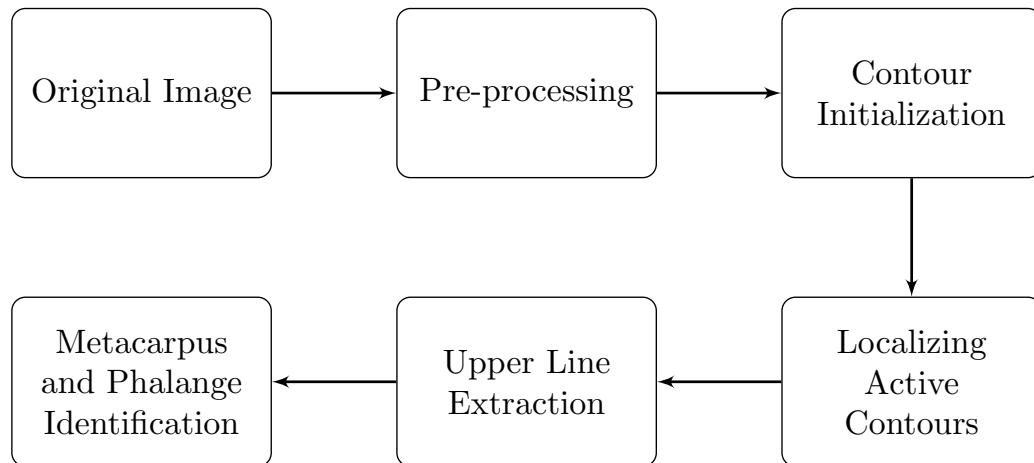


FIGURE 5.2: Proposed pipeline for the identification of the metacarpus and phalange.

5.2.1 Pre-processing

The proposed method starts with the trimming of the bottom half of the image, followed by a down-sampling by a factor of 2. These procedures reduce the amount of information to be processed and increase the processing speed. From Fig. 5.1, it is possible to see that the bottom halves of the images have no useful information since the ultrasounds are not able to reach so deep into the tissue. Regarding the down-sampling, there was no visual evidence that it interfered with the overall segmentation performance. These steps were empirically selected, and further down-sampling or trimming can be performed if the processing speed is critical.

5.2.2 Contour Initialization

For the initialization of the LAC algorithm, a template approach was used. It is known that the bones are brighter than the background and have a horizontal linear shape. Moreover, they do not overlap at any point, and the end of the metacarpus is close to the beginning of the phalange. Based on these assumptions, a rectangular structuring element was used to locate possible bone regions. To do so, the template was moved vertically, looking for the region of maximum integral (inside the template), which is set to the foreground at the end of each vertical sweep. Thereafter, the template is moved horizontally to the right to perform another vertical sweep. Based on the continuity assumption, we also impose that, from one template to the next, there should be some overlap, and thus the horizontal step is set to half of the structuring element width. The vertical step was set to 1 pixel, and the template size was 5% of the trimmed image size (height and width).

5.2.3 Contour Refinement - LAC

After contour initialization, the LAC algorithm is used to improve the segmentation. This technique was based on the work of [87], which was briefly explained in Chapter 4. Like most of the active contours methods, it starts from a coarse segmentation and iteratively converges to a more precise segmentation. In the LAC method, the energies are calculated locally, allowing a natural adaptation to non-uniformities of the

background. This characteristic is useful because, in the MCPJ images, the background and foreground are different along the image and within the same objects. For instance, the joint capsule area is usually darker than the peripheral zones (Fig. 5.1). The use of other techniques, such as Chan-Vese, that take into consideration the characteristics of all interior and exterior regions of the contour would require a pre-processing stage to reduce the image non-uniformities.

The parameters were empirically selected as follows: 800 iterations and a neighborhood radius of the local component of 5 pixels. The resulting mask is up-sampled to the original size, and artifacts (small structures) are removed. Fig. 5.3 shows the evolution of the LAC algorithm, from column 1 and 2 (iteration 10 and 800).

5.2.4 Upper Line Extraction

The upper line extraction module aims to identify the bone surface. It uses the mask obtained from the LAC and assumes that there are no structures under the bones since they are highly reflective of ultrasound waves. Following this idea, the vertical first derivative of the mask is computed, and only the positive values are kept (background to foreground transitions). The resulting image is dilated with a 3x3 square kernel, and the centerline is extracted to ensure connectivity and slightly smooth the result. Afterwards, the two biggest lines are selected as the bones, and lines in the same columns as these two lines are removed. As previously said, there is no bone overlap, and therefore only one line should be present in the vertical direction. At this point, it is expected that only two lines are present in the image, one for the metacarpus and another for the phalange. To separate them, the line on the left is assigned to the metacarpus and the one on the right to the phalange. In some cases, the algorithm returns more than two structures due to discontinuities in the bone, as in the case of osteophytes or bone erosions. In these cases, the lines between the metacarpus and phalange lines are connected to the closest one until only two lines remain. The closeness measure is the horizontal Euclidean distance from the tips of the lines. This choice was based on visual inspection of the results. In some cases, the metacarpus discontinuities create a vertical gap larger than the typical horizontal gap between the phalange and the metacarpus (5.3 - case 3); consequently, the inclusion of the vertical distance would force the connection of the tip of the metacarpus to the phalange.

5.3 Results and Discussion

To test the proposed method, a database of 164 images from the second MCPJ was used. The images were acquired with the collaboration of Unidade Local de Saúde do Alto Minho and manually annotated by two rheumatologists using dedicated software. The annotations consisted of the identification of the metacarpus and phalange upper line and the tendon. The images were acquired with a high-frequency ultrasound device, a GE Healthcare LOGIC-S8, and saved in anonymized DICOM format. Regarding the image characteristics, its size is $\{448 \times 760\}$ (rows, columns) pixels with 8 bits and a spatial resolution variable between $40 \mu m$ and $55 \mu m$.

Fig. 5.3 shows some examples of intermediate stages and the final results of the proposed method.

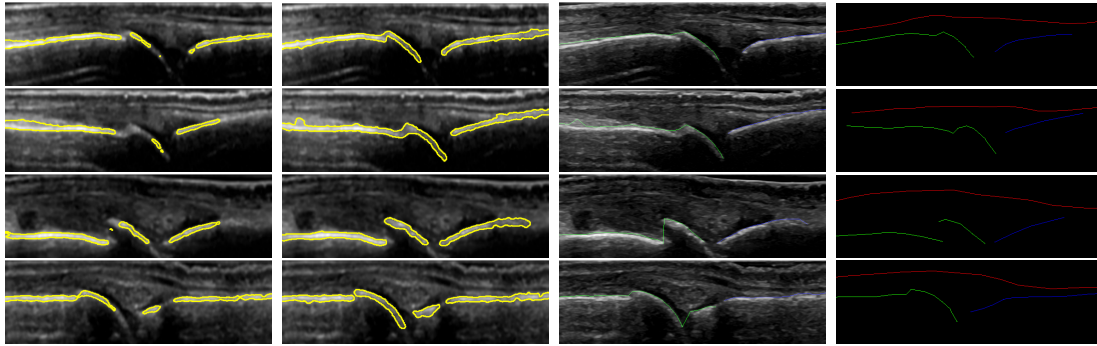


FIGURE 5.3: Visual results of the metacarpus and phalange segmentation using the LAC algorithm. In the lines it is possible to see different cases; in the first column, the results after 10 iterations, in the second the final LAC results, in the third the upper line extraction and in the last column the manual segmentation.

Attending to the results, it is possible to conclude that, for cases 1, 2 and 3, the algorithm performs as expected. In 3, the effect of the metacarpus discontinuity was compensated by the proposed change in the line detection. Alternately, in case 4, the horizontal distance was not a good measure to overcome the phalange discontinuity.

To test the algorithm's capacity to detect the bones, the average segmentation error was calculated. The metric used was the pixel-wise vertical Euclidean distance. If this distance was smaller than a given threshold, it was considered a true positive; otherwise, it was considered a false positive. A set of thresholds was used to create a plot with the segmentation error and the respective *confidence*. In this manner, it is possible to gain a better understanding of the *confidence* intervals for each error and select the desired operation point.

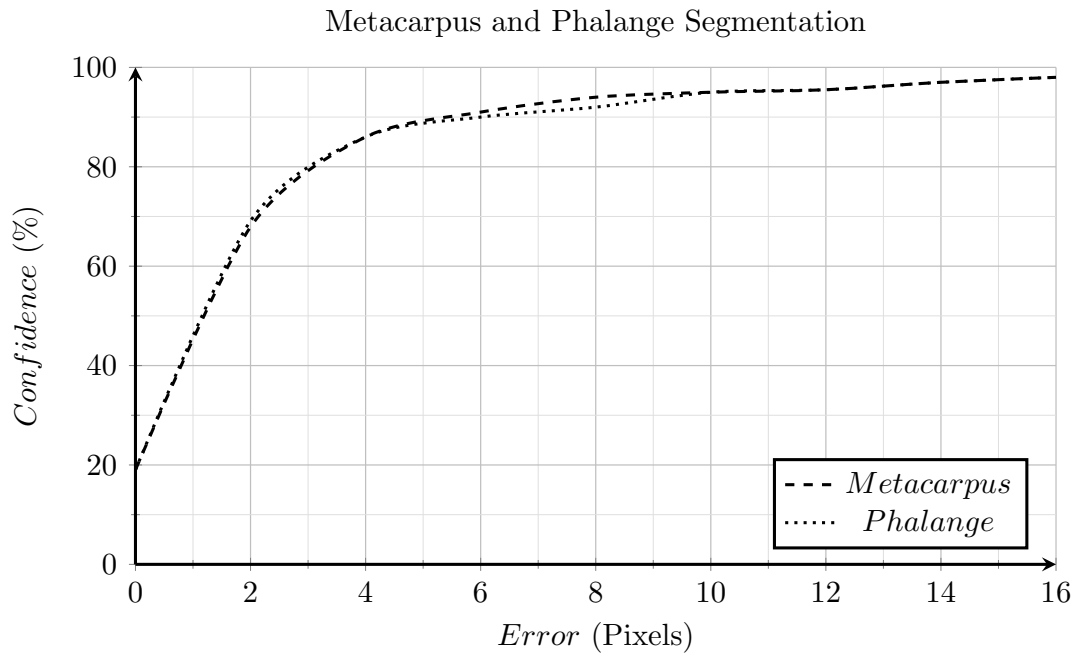


FIGURE 5.4: Segmentation error for the metacarpus and phalange, using the vertical distance in pixels.

Looking at Fig. 5.4, it is possible to conclude that the segmentation of the metacarpus and the phalange led to similar results. Approximately 95% of the pixels were segmented with an error smaller than 10 pixels (corresponding to approximately 0.5 *mm*).

5.4 Complementary Results

Throughout the development of the work presented in this thesis, the database got larger with the inclusion of additional data and annotations. Because of that, an additional section was added to this chapter so that the results of the entire database are known. To allow better comparison with the previous work, the same performance measure was used, which is the vertical distance in pixels between the resulting segmentation and the ground truth. After that, and to achieve an improved metric, the modified Hausdorff distance, 4.7.1, in *mm* was used.

Fig. 5.5 shows the results for all 240 images, expanded from the original 164 published in [9].

The results for all 240 images show that the segmentation of the metacarpus was very close to that presented in [9], with the same 95% *confidence* for an error of less than

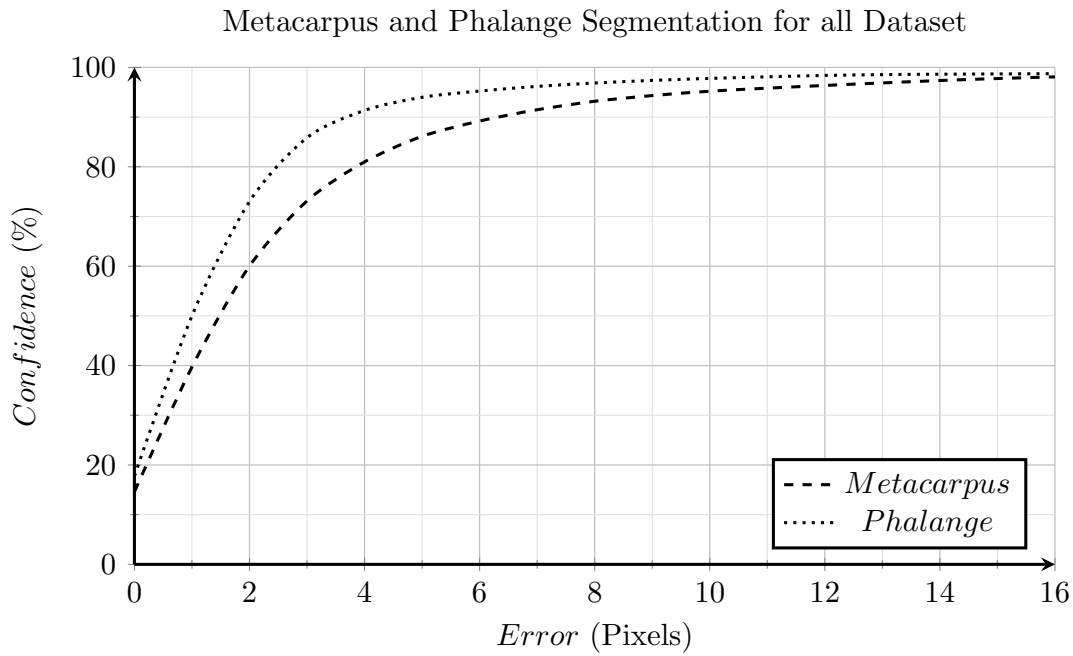


FIGURE 5.5: Segmentation error for the metacarpus and phalange using the vertical distance in pixels, in all dataset.

10 pixels. For a smaller error of 4 pixels, the results were slightly worst on the new dataset (80% vs 85% *confidence*). Looking at the phalange segmentation results, they were better in the new dataset. For 95% *confidence*, an error of 6 pixels was obtained, where [9] achieved an error of 10 pixels for the same *confidence*. These results prove that the proposed algorithm was robust enough to be used on new images without new parameterizations.

An improved metric is also presented here, using the modified Hausdorff distance, 4.7.1, instead of the vertical distance. This metric is more realistic because both dimensions are taken into account; moreover, the image analysis is more interesting because the number of pixels is not constant in all images, and merging them into one result could thus shadow some results. In this manner, it is possible to obtain a better idea of the percentage of images that are well segmented and identify the ones with the worst results. In Fig. 5.6, the improved metric results are shown.

The results presented in Fig. 5.6 reinforce the ones previously obtained. The phalange was segmented with higher *confidence* than the metacarpus for small errors. For an error of 0.3 *mm*, both structures present similar results, with slightly better results for

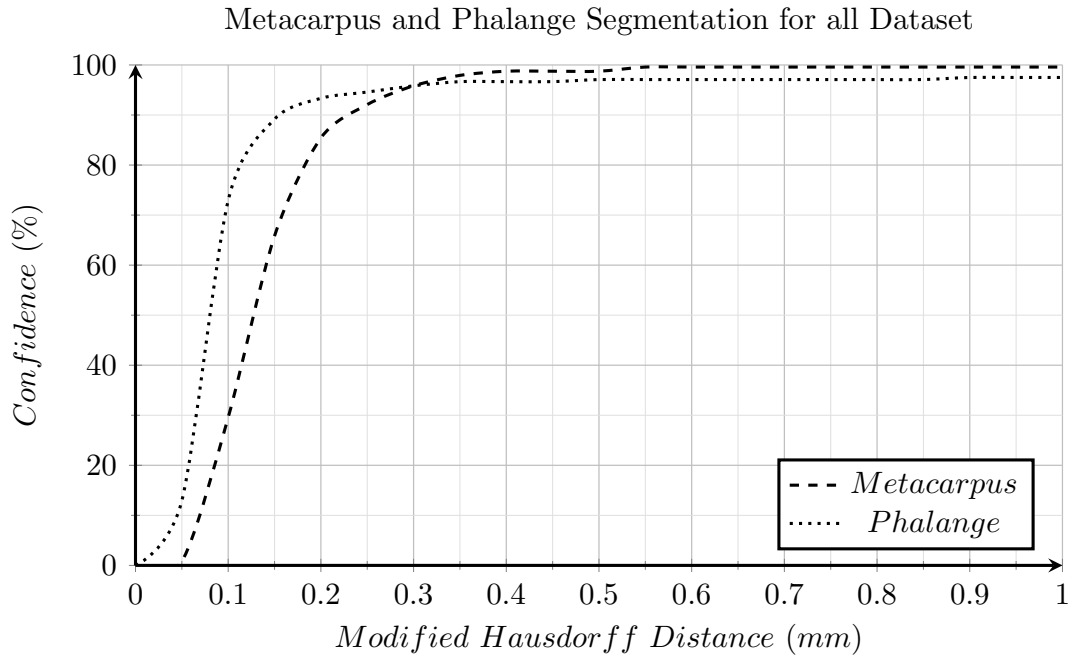


FIGURE 5.6: Segmentation error for the metacarpus and phalange using the Modified Hausdorff Distance in millimeters, in all dataset.

the metacarpus on errors above 0.3 *mm*. Overall, 95% of the images were segmented, both with errors of less than 0.3 *mm*.

5.5 Conclusions

A new method for the segmentation of the metacarpus and phalange in images of the second MCPJ was proposed. The results obtained showed that it is possible to automatically identify these structures, with a *confidence* of 95% for an MHD error below 0.3 *mm*. Additional tests in new data and with new measures reinforced the robustness of the proposed segmentation method. We believe that these are acceptable results, but an extensive study should be conducted to understand the clinical implications of such values. From the present results it is also possible to observe that, in some cases, the results obtained on the phalange are better than the ones obtained on the metacarpus. One possible justification is the fact that the metacarpus has higher shape variability and that more pathologies are found on it, namely bone erosions and osteophytes. Nevertheless, they are sufficient to reduce the search area of other structures, such as the extensor tendon and joint capsule, which are still open problems, as

well as the classification of MCPJ images as pathological or non-pathological. The use of this technique in other ultrasound segmentation problems is possible, and this work can be used as a starting point due to its simplicity and relatively fast processing.

Chapter 6

Extensor Tendon Identification

Contents

6.1	Introduction	59
6.2	Segmentation Framework	61
6.2.1	Internal Energy	62
6.2.2	External Energy	63
6.2.3	Area Constraints	65
6.2.4	Additional Constraints	66
6.2.5	Implementation	67
6.3	Results and Discussion	68
6.3.1	Dataset and Metrics	68
6.3.2	Number of Points Optimization	69
6.3.3	Weights Optimization	70
6.3.4	Visual Results	72
6.4	Complementary Results	73
6.4.1	New Data Results	73
6.4.2	Fully Automatic Segmentation	74
6.5	Conclusions	77

6.1 Introduction

The extensor tendon is the structure responsible for the opening of the hand by connecting the extrinsic extensor muscles of the hand to the fingers. In the ultrasound

images, they are visible as a dark line between the skin and the metacarpus and phalange (Fig. 6.1). The segmentation of this structure is one of the most challenging steps of this work since the tendon is not always visible and has an appearance similar to other structures present in the image, such as the fat layer below the skin. Moreover, the presence of pathologies can change its appearance and position, as shown in Fig. 6.1.

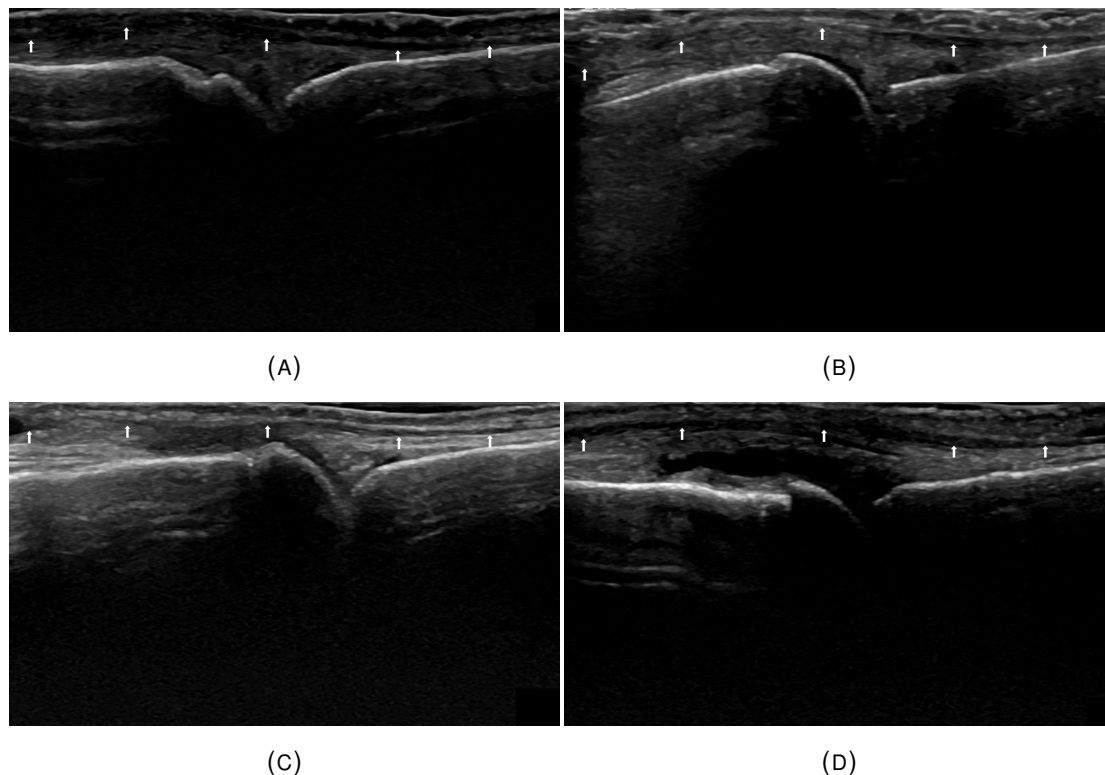


FIGURE 6.1: Examples of ultrasound images of the dorsal view of the second MCPJ. The white arrows indicate the position of the extensor tendon.

From Fig. 6.1, it is possible to visualize the previously discussed challenges. In Fig. 6.1 - A, C, the tendon and a fat layer are visible. They are very similar in echogenicity and shape, and the most distinguishing aspect between them is that the fat layer is always above the tendon. In Fig. 6.1 - B, the tendon is only visible in the right half of the image; in the left half, the rheumatologist typically infers the rest of the tendon, assuming continuity. In Fig. 6.1 - C, an artifact known as tendon anisotropy can be seen, where the tendon creates a shadow effect below itself, making the structures darker and with an echogenicity similar to the joint capsule. This artifact can also be confused with synovitis since the region that gets darker might be the same that gets affected in synovitis. In Fig. 6.1 - D, it is possible to see the effect of the synovitis in the position of the tendon, moving it closer to the skin. The correct identification of this

structure on the MCPJ images can be used to separate the joint capsule region from the skin region, making it easier to identify and analyze the joint capsule.

Attending to the literature, it is possible to find several works regarding the segmentation of tendons in ultrasound images [48, 93–95], but most of them target large tendons, such as the Achilles and knee tendons. Unfortunately, these tendons are much thicker and have a different appearance from the extensor tendon of the fingers in the ultrasound images. Because of that, the frameworks used for large tendons are not applicable for the problem presented here. Regarding the identification of the extensor tendon, only a preliminary approach was found in the literature [12]. In this work, the segmentation was achieved using a pre-processing based on the Log Gabor transform to enhance ridge-like structures and reduce speckle noise. Then, a medial axis transform is used, followed by a filtering of the vertical connected lines. The resulting lines are adjusted using morphological operators, and small lines are discarded. Finally, the tendon is selected from the remaining lines using anatomical references. The limitation of this method was related to the fact that the tendon was not always fully identified. In some cases, the tendon was only segmented in one portion, and the remaining (less visible) part was discarded. Based on the problems identified in [12] and the image characteristics, a segmentation by fitting using an active contours framework was chosen. This framework is robust against speckle noise and allows the incorporation of shape knowledge, which is helpful to enforce the model to missing structures in the image. Next, the complete framework will be discussed in detail, which is based on the work presented in [8].

6.2 Segmentation Framework

To identify the extensor tendon in an image $I(x, y)$, we modeled the active contour line as a set of N two-dimensional points, $v_i = \{(v_{x1}, v_{y1}), (v_{x2}, v_{y2}), \dots, (v_{xN}, v_{yN})\}$, spaced along the image. These points are used to obtain the complete tendon line, $v_s(x)$, by applying an interpolation function. This model is interesting for reduction of the optimization search space, because v_i is a subset of v_s . Knowing that the tendon is a continuous structure and that it behaves like a string under horizontal tension with vertical forces applied due to the joint capsule, bones and skin, we selected a spline

interpolation. However, other interpolation methods could be used, such as linear interpolation, without significantly changing the algorithm implementation. We also introduce here the neighborhood of $v_s(x)$, as $v_s(x, j)$, where j is the perpendicular distance from the point $v_s(x)$, negative values of j are assumed to represent below or inside and positive values above or outside the contour, and a value of zero is the line itself, i.e., $v_s(x, 0) = v_s(x)$. The energies are calculated using $v_s(x, j)$, and the challenge lies in identifying the position of each point in v_i , as shown in Fig. 6.2.

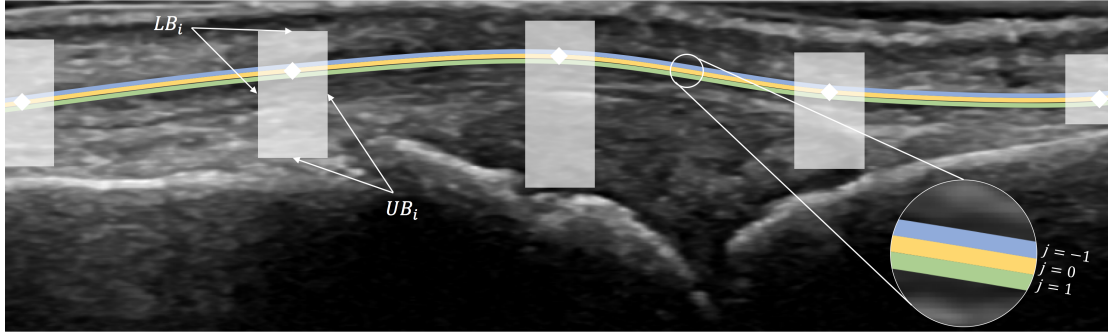


FIGURE 6.2: Proposed open ended active contour model framework. The white diamonds represent v_i , the horizontal blue (above), yellow (middle) and green (below) lines are, respectively, $v_s(x, -1)$, $v_s(x, 0)$, $v_s(x, 1)$. The transparent squares are the allowed moving zones of each point v_i , which are delimited by the lower and Upper bounds, LB_i and UB_i (explained in more detail in Section 6.2.4).

To identify the coordinates of the points v_i , we used a modified version of the active contours formulation briefly presented in Section 4.2. The identification of the proper metrics for the internal energy (E_{int}), external energy (E_{ext}) and additional constraints are the key points of this work. Next, each one will be addressed in more detail.

6.2.1 Internal Energy

In this model, the connectivity is guaranteed because of the use of a sparse representation and interpolation, which simplifies the E_{int} formulation to some degree. However, it is still necessary to impose some constraints. The first one is to limit the slope of the tendon. Given the hand anatomy and acquisition protocol, we know that the tendon should be mostly horizontal. Due to pathologies, bad hand posture or even natural variations, the tendon can have higher curvatures, so we imposed a tendon slope limit and modeled it as the soft constraint (6.1).

$$E_{Slope}(v_s(x)) = \sum_x \max(0, |v'_s(x)| - \text{tg}(\theta)) \quad (6.1)$$

where $v'_s(x)$ is the first derivative of $v_s(x)$ and θ is the slope threshold. Based on the anatomical aspect of the tendon and acquisition method, a value of 25° was used in this implementation. The second constraint is that the line, $v_s(x)$, should not have more than 2 concavities, one due to the joint capsule and the other due to the metacarpus or phalanx. One way to quantify this is by counting the number of zero-crossings of $v'_s(x)$ and applying a decision function (6.2).

$$E_{Concavity}(v_s(x)) = \frac{1}{1 + e^{-p_1(ZC(v'_s(x)) - p_2)}} \quad (6.2)$$

where p_1 is the parameter that controls the variation of the function and p_2 is the point where the energy is equal to 0.5. $ZC(v'_s(x))$ is the number of zero-crossings of $v'_s(x)$; because it is an integer and, in this case, has a limited number of allowed values $\{0, 1, 2\}$, we set the parameters p_1 and p_2 to 10 and 2.5, respectively. This ensures that all allowed values have energy close to 0 and the others close to 1. In practice, this makes the energy behave like a hard constraint.

6.2.2 External Energy

Regarding the external energy formulation, the ultrasound image is very noisy, and some structures may be missing due to inadequate probe position; because of that, gradients are not good energy measures. To overcome this problem, pre-processing techniques are used to reduce the noise and enhance the objects of interest. Here, we propose the use of phase symmetry to enhance the extensor tendon. As previously shown, the phase symmetry has several parameters, with the most important being the angle span and the bandwidth. The angle span, ϕ , was set to 90° because the tendon is a horizontal structure. For the bandwidth, we chose to analyze 4 scales, starting from 0.2 mm (4 pixels in size) and a step of 0.15 mm (3 pixels), performing the analysis up to 0.8 mm (16 pixels). These values were empirically chosen, based on the spatial resolution ($\sim 50 \mu m$) and typical tendon thickness. Finally, we also choose to use both positive and negative concavity structures (the ridge and valley-like structures,

respectively). Note that both valleys and ridges are symmetric, but they have inverse concavities. Given that the problem formulation implies a minimization, we choose to penalize the ridges and enforce the valleys via (6.3).

$$I_{Sym}(x, y) = \frac{\text{Sym}_{\oplus}(x, y) - \text{Sym}_{\ominus}(x, y) + 1}{2} \quad (6.3)$$

where $\text{Sym}_{\oplus}(x, y)$ is the positive concavity, $\text{Sym}_{\ominus}(x, y)$ is the negative concavity, and $I_{Sym}(x, y)$ is the transformed input image $I(x, y)$. Fig. 6.3 shows the resulting images for 3 cases.

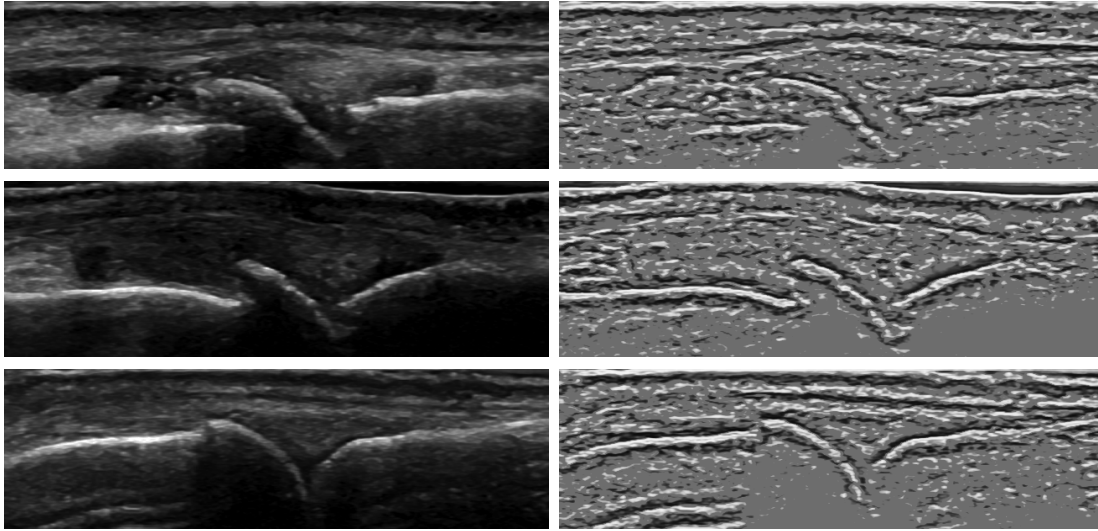


FIGURE 6.3: Example of the phase symmetry pre-processing results; on the left the original image, $I(x, y)$, and on the right the resultant image, $I_{Sym}(x, y)$.

Finally, the $E_{PhaseSymmetry}$ is equal to the sum of the values of I_{Sym} along v_s , (6.4).

$$E_{PhaseSymmetry} = \frac{1}{N} \sum_j w_n[j] \sum_x I_{Sym}(x, v_s(x, j)) \quad (6.4)$$

where $w_n[j]$ is the weight of each neighborhood line j . Here, we used $j = \{-1, 0, 1, 2\}$ with the respective weights of $\{1.0, 1.0, -0.5, -0.5\}$. This choice was based on the fact that we wanted to identify the tendon from its upper limit, and thus the neighborhood lines are used to detect bright-to-dark region transitions. Even with a noisy image, we decided to include the original image in the energy formulation to verify if it is possible to use it as a metric despite the known limitations. Additionally, it will be used to test

the hypothesis that the phase symmetry is a valid pre-processing method. In (6.5), the energy calculation for this is presented as:

$$E_{Intensity} = \frac{1}{N} \sum_j w_n[j] \sum_x I(x, v_s(x, j)) \quad (6.5)$$

Once again, $w_n[j]$ are the neighborhood weights, and the image, I , is assumed to be normalized between 0 and 1. The values of j and $w_n[j]$ are the same as the $E_{PhaseSymmetry}$.

6.2.3 Area Constraints

The position of the tendon is expected to be somewhere between the skin line and the metacarpus and phalanx bones. To incorporate this knowledge in our model, we propose the use of the area between the tendon and the skin line, A_{Up} , and the area between the bones and the tendon, A_{Down} , as references. It is expected that these areas have similar values with some variations due to natural anatomical differences. One way to quantify this is to calculate the percentage of the area above the tendon, referred to as A_{Ratio} , using (6.6).

$$A_{Ratio} = \frac{A_{Up}}{A_{Up} + A_{Down}} \quad (6.6)$$

Given the previous assumptions and a safety margin of 30%, we assume that A_{Ratio} is normal if its value is approximately 50% +/- 15%. To validate our premise, we used the expert annotations to extract measurements in a controlled environment. The distributions of these areas are shown in Fig. 6.4.

From Fig. 6.4, it is possible to conclude that A_{Up} and A_{Down} have similar area distributions and that A_{Down} has one outlier. Alternately, A_{Ratio} has no outliers, it is independent in spatial resolution, and the distribution is slightly more compact, with all its values between 34% and 64% (49% +/- 15%). Based on this, we concluded that our assumptions regarding the relative position of the tendon are correct. These values were used to create the energy function (6.7).

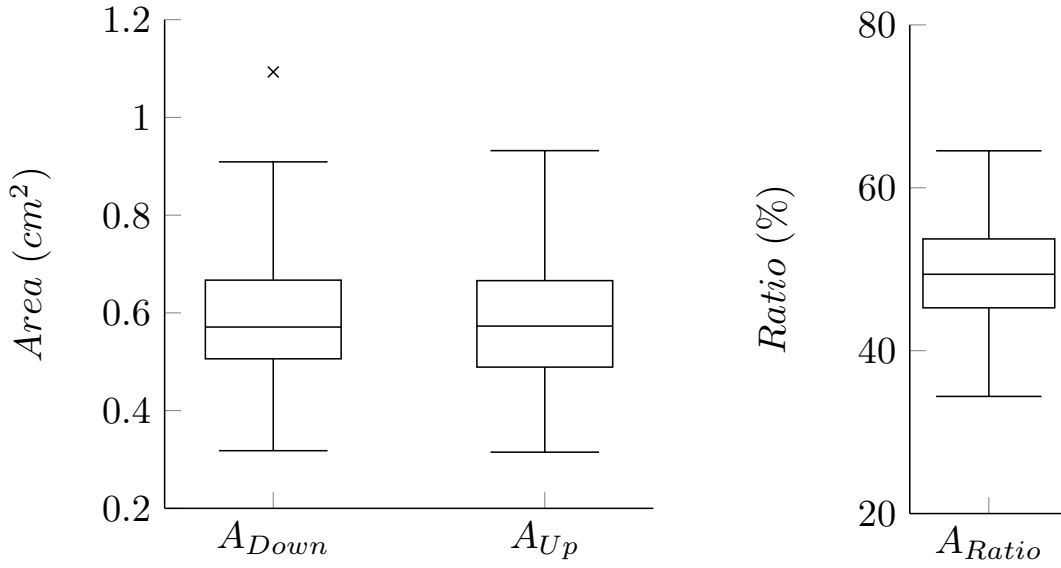


FIGURE 6.4: Box plot of the area measurements, the cross indicates the value of the biggest outliers.

$$E_{AreaRatio} = \max(|A_{Ratio} - p_3| - p_4, 0) \quad (6.7)$$

where $E_{AreaRatio}$ is the energy metric, p_3 is the minimum energy point, and p_4 is the variation of the energy. Based on Fig. 6.4 and using a $[0, 1]$ interval, we set the values of p_3 and p_4 to 0.5 and 0.1, respectively. This ensures that $E_{AreaRatio}$ is low (< 0.05) in the 50% \pm 15% interval and that it increases as A_{Ratio} moves away from this region.

6.2.4 Additional Constraints

To promote better results and faster convergence, we decided to impose some additional constraints. The lower and upper bound limits are used to reduce the search space of the points v_i . Letting LB_i and UB_i represent the lower and upper bounds of v_i , respectively, we impose that:

$$LB_i < v_i < UB_i, \forall i \in \{1, 2, \dots, N\} \quad (6.8)$$

Fig. 6.2 shows the regions delimited by LB_i and UB_i . Note that, in this case, the horizontal limits (upper and lower) are equally spaced because no relevant information

can be extracted from this component. Regarding the vertical component, both limits are dependent on the anatomical structures. For instance, we know that the tendon should be between the skin line and the bones, so these two structures are used to impose the limits.

6.2.5 Implementation

Based on the previous formulations, the final energy is presented in (6.9).

$$E(v_s) = \sum_k w_e[k] * E_k \quad (6.9)$$

where $w_e[k]$ are the weights used to scale and control the importance of each energy term, E_k , with $k = \{PhaseSymmetry, Intensity, Slope, Concavity, AreaRatio\}$, subject to (6.8).

Once the energy formulations are completed, (6.9) needs to be solved. The gradient descent method [85] is commonly used, where the derivative of the energy is used to conduct a pre-defined coarse segmentation to a finer segmentation (local minimum). Another alternative involves the Euler-Lagrange equation, which uses the same idea of the derivative, moving the contour with the gradient, but with a different formulation [85]. Both methods work well in cases where the number of local minimums is small or the confidence in the initialization is good. However, these two characteristics are not always met, as is the case for this work. Furthermore, the implementation of a gradient-based approach is not always possible since the derivative of the energy must be calculated. A Genetic Algorithm (GA) approach was chosen since no explicit initialization is needed and local minimums can be avoided to a certain point because of the multiple random initializations [96]. Other optimization methods could also be used, such as greedy or Monte Carlo approaches, but since the optimization method is not a key feature of this work, we decided to use GA to validate our energy formulation. The GA implementation of the MathWorks MATLAB® R2015b "Global Optimization Toolbox" was used, with the default parameters.

6.3 Results and Discussion

The proposed method was tested on a 4 GHz iMac with 32 GB of RAM. For this study, we used the manual segmentation of the metacarpus, phalanx and skin line to remove possible interferences caused by automatic segmentation methods, which are themselves prone to errors. In that manner, we measure the isolated performance of the automatic tendon segmentation.

6.3.1 Dataset and Metrics

The proposed algorithm was tested on images of the dorsal view of the second MCPJ extracted from 83 patients. Images with no visible tendon were discarded, and the remaining were manually annotated by two expert rheumatologists, resulting in a total of 175 images. The annotations consisted of the identification of the metacarpus, phalanx, skin line and extensor tendon. The images were acquired with a high-frequency ultrasound device (12 MHz), a GE Healthcare LOGIQ-S8, and saved in anonymized DICOM format. Regarding the image characteristics, their size was cropped to 448x760 (rows, columns) pixels with 8bits/pixel and a spatial resolution between 40 μm and 55 μm . The set of images is unbalanced, with more cases without pathology (137) than with pathology (38). There was no special care in the balancing of the gender, age or other person-related characteristic. The MHD (4.7.1) was used as metric to compare the results obtained with different parameters. It outputs a segmentation error for each image, and so it is possible to create a cumulative distribution by calculating the percentage of correctly identified images (*confidence*) for different MHD threshold values. The correctly segmented images will be the ones with a MHD below the threshold while the incorrectly segmented are the ones with a MHD above threshold. In this way, it is possible to easily compare different configurations by looking at the plots. Additionally, the best method should be the one with the higher area under the curve (AUC), which is easy to measure and visualize in a cumulative graph.

6.3.2 Number of Points Optimization

First, we looked for the number of points, N , knowing that high values will allow more complex shapes but might also promote local minimums and increase the computational time. Alternately, low values will make the line less deformable, and it might not be able to fit all needed shapes. Given this trade-off, the value of N should be the lowest possible without compromising the results. Note that local minimums can be reduced using additional constraints, such as the E_{Slope} and $E_{Concavity}$. To study N , we extract equally spaced points from the manual segmentation. These points are then used to create a new approximated tendon using spline interpolation. The error is calculated using the MHD between the interpolated and reference tendon, and the results are shown in Fig. 6.5.

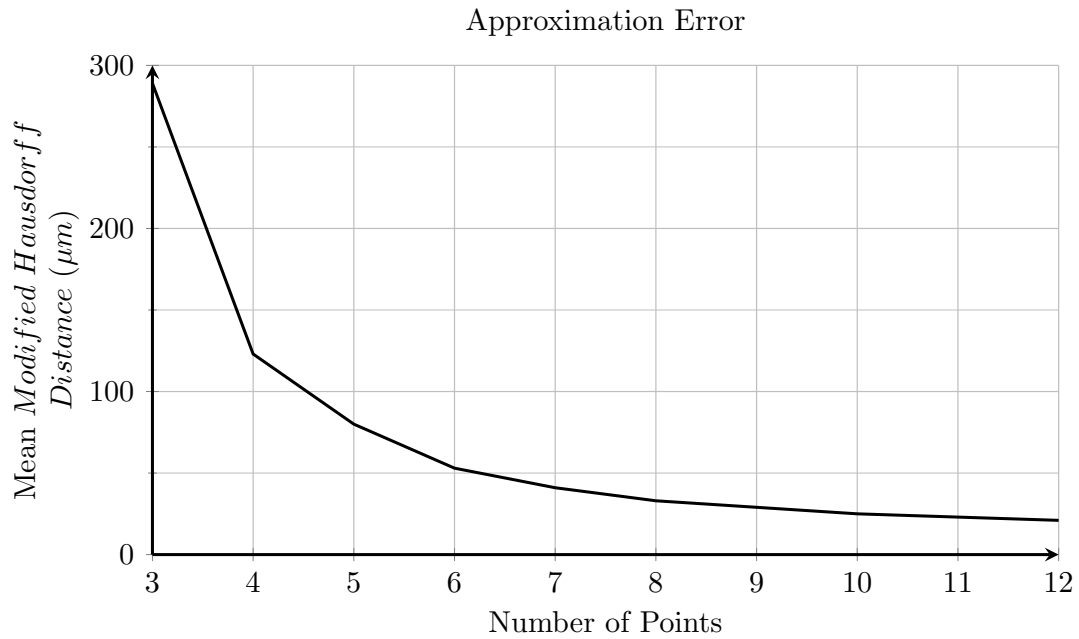


FIGURE 6.5: Approximation error using different number of points.

From Fig. 6.5, it is possible to verify that the error decreases with increasing number of points. Assuming a mean error equal to the mean spatial resolution ($50 \mu m$), a minimum of 6 points should be used, and because of that, N , was set to 6.

6.3.3 Weights Optimization

In this section, we want to verify the performance of the proposed method and obtain the configuration that yields the highest segmentation performance. To achieve this, we created two main test sets, one where the external energy was calculated with *PhaseSymmetry* and the other with *Intensity*. In this manner, we can verify which one works best. Then, for these two main sets, we tested the inclusion of the line neighborhood information, $w_n[j]$, (hereafter referred to as *NeighborhoodWeights*), and optimized the weights, $w_e[k]$, for $k = \{Slope, Concavity, AreaRatio\}$. These weights were learned from a random balanced set of 50 samples (25 pathological and 25 non-pathological) using grid search, and the remaining samples (125) were used in the tests. The grid search parameters were set between 0 and 2 with a step of 0.2 for the *Slope* and *AreaRatio*; for the *Concavity*, the allowed values were 0 and 1 since it is a hard constraint. The metric used to compare each configuration was the AUC, as presented in Section 6.3.1, using a threshold range from 0.0 *mm* to 1.0 *mm*, with a step of 0.005 *mm*. The results obtained on the test set are shown in Fig. 6.6.

Based on Fig. 6.6, it is possible to correlate the segmentation error and the *confidence* in that error. In this manner, it is possible to verify if the proposed method is suitable or not for a given problem. Regarding improvement proposals for the open-ended active contours framework, it is possible to conclude that the *PhaseSymmetry* clearly improved the segmentation results in all configurations, achieving the best AUC of 0.786. Alternately, the best *Intensity* achieved an AUC of 0.698. Regarding the *NeighborhoodWeights* the results showed a clear improvement in all cases. For instance, in the *PhaseSymmetry*, the AUC increased from 0.694 to 0.786. The *NeighborhoodWeights* forces the curve to move to transition zones, which is more restrictive than low-intensity zones, since it is dependent on the tendon thickness and more prone to ultrasound artifacts (tendon anisotropy). The *Slope*, *Concavity* and *AreaRatio* also improved the overall results, but their impact is more noticeable on the *Intensity* formulation, increasing from 0.623 to 0.698. In the configuration with *PhaseSymmetry*, the improvement was residual, approximately 0.015 in the AUC, and more significant for errors between the 0.3 *mm* and 0.6 *mm*. This observation was attributed to the fact that, in the *Intensity* formulation, the problem is much more loose since the information in the image

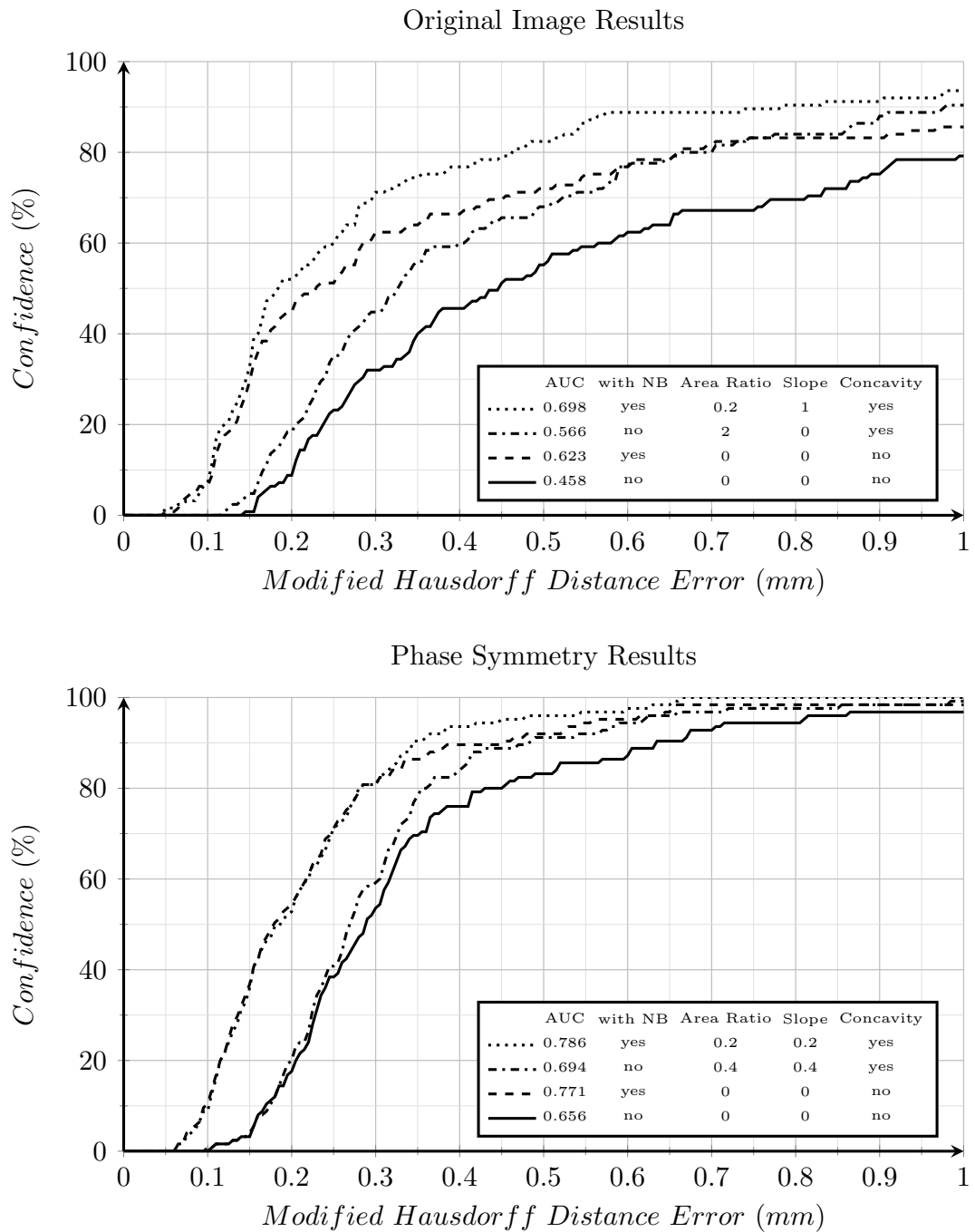


FIGURE 6.6: Results obtained for different parameter configurations of the proposed method. On the left are the results of the method using no pre-processing (*Intensity*), and on the right are the results with phase symmetry pre-processing (*PhaseSymmetry*). The parameters used in each configuration are in the respective tables. MHD is Modified Hausdorff Distance, NB the *NeighborhoodWeights*, (6.4) and (6.5), and AUC is the area under the curve.

has more noise, and thus additional constraints help in the correct convergence. Alternately, in the *PhaseSymmetry* formulation, the images have less noise, and thus

the line converges more easily to the correct position without the need for additional constraints. Assuming a *confidence* of 95%, the best *PhaseSymmetry* configuration was able to identify the tendon with an error below 0.5 *mm*, whereas methods with *Intensity* achieved a 90% *confidence* for an error below 0.75 *mm*. This enforces the *PhaseSymmetry* as a valid pre-processing method in this application.

Even though the number of pathological images is smaller than the non-pathological, we decided to separate them and measure the algorithm performance for the best *PhaseSymmetry* and *Intensity* configurations, giving us better insights into the performance of our proposed algorithm. The results show that the AUC is smaller for pathological images, but in the *PhaseSymmetry* configuration, the reduction was much smaller (from 0.786 to 0.732) than in the *Intensity* (from 0.698 to 0.386). These results reinforce the utility of the *PhaseSymmetry* pre-processing.

6.3.4 Visual Results

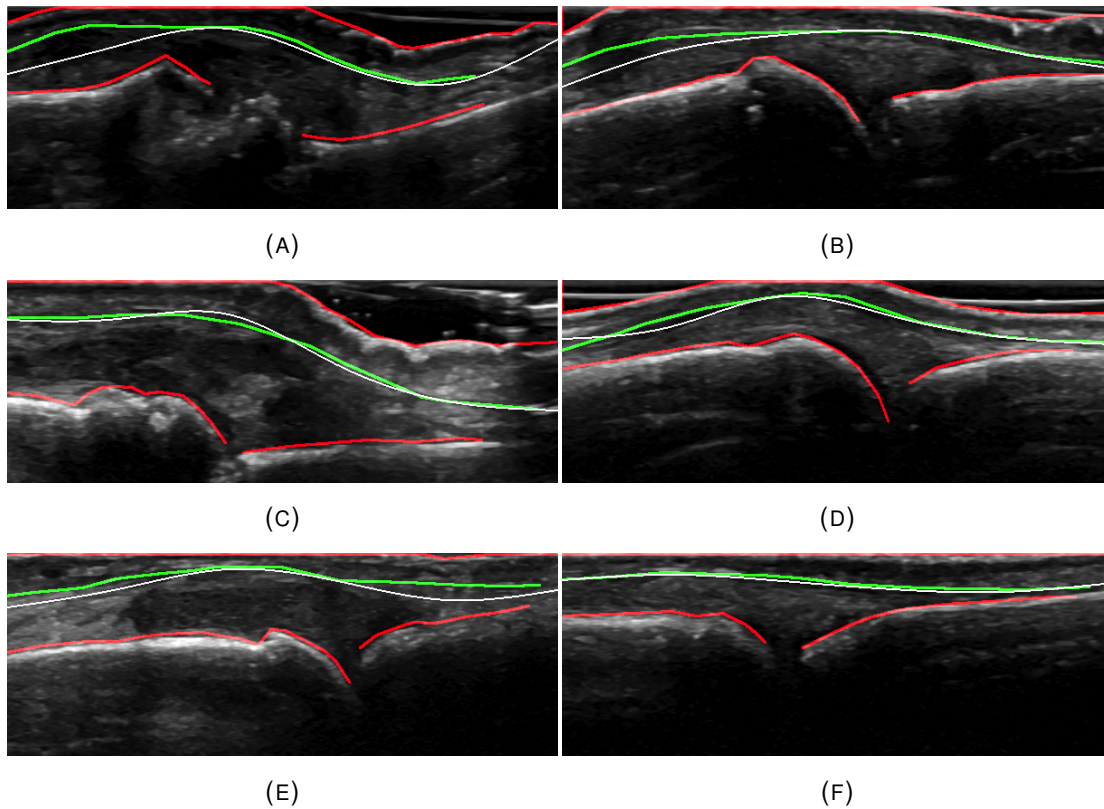


FIGURE 6.7: Visual results of the extensor tendon segmentation. In red are the upper and lower limits, in green (thicker) the reference segmentation and in white the results; Left column - pathological; Right column - non-pathological; A, B, E with MHD above 6 pixels and C, D, F with MHD below 6 pixels.

Finally, some visual examples are shown in Fig. 6.7. It can be seen that the tendon position and shape are highly variable, as well as the size of each finger joint. Nevertheless, the segmentation was achieved even when the image was more complex (Fig. 6.7 - A, C). In Fig. 6.7 - F, the image is considered to have very good quality, and the subsequent segmentation was almost perfect. Looking at images with worst visual results, Fig. 6.7 - A, B, E, we can observe that the error was higher in the peripheral regions, which correspond to the areas of less clinical interest.

6.4 Complementary Results

Following the same idea of Chapter 5, it was decided to extend the results so that the all new data acquired along the project is included in this work. In [8], the proposed method was tested on 175 images (50 to train the parameters and 125 to test), and the results that follow include 65 new images, for a total of 240 images. The best parameter configuration obtained in [8] is used here without further parameterizations, and the manual annotations will also be used. After that, a fully automatic method is presented, where the manual annotations are replaced by automatic methods.

6.4.1 New Data Results

Fig. 6.8 presents the results for the 65 new images added after the publication of the work presented in [8]. The metrics used are the same so that a direct comparison is possible.

From Fig 6.8, it is possible to conclude that the results obtained on the new dataset are comparable with those presented in [8]. For 95% *confidence*, the error obtained on the new dataset was 0.6 *mm*, 0.1 *mm* higher than the results of the original dataset. These differences are acceptable and still proves the efficiency of the algorithm. A study on a larger database would be a plus for this work since small databases are more susceptible to the variability of the data.

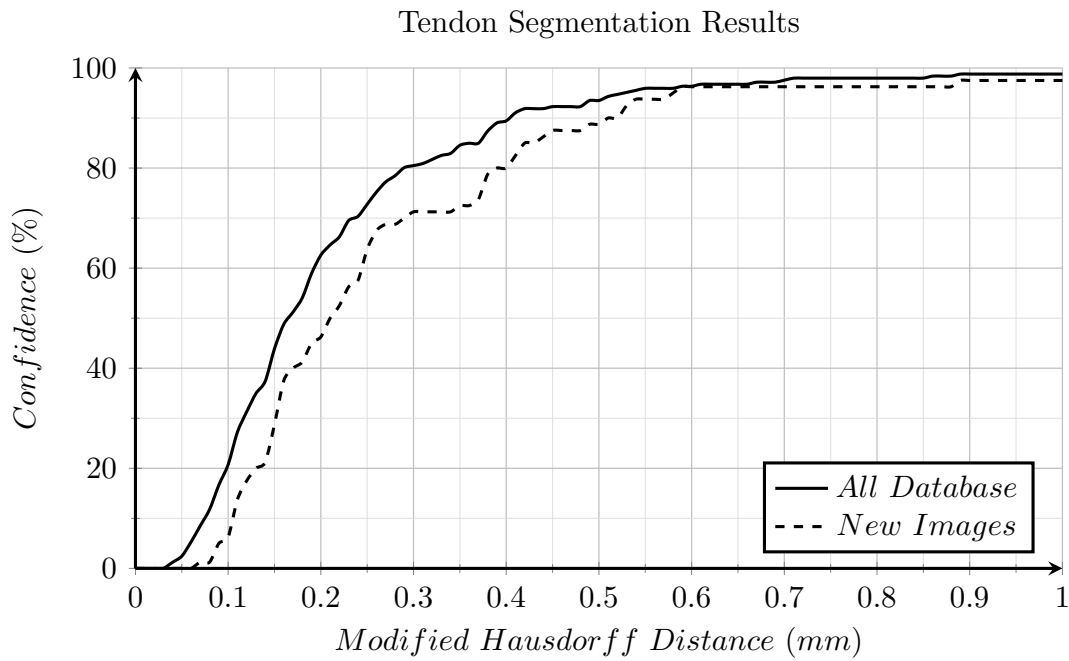


FIGURE 6.8: Results obtained for the segmentation of the extensor tendon, in all database.

6.4.2 Fully Automatic Segmentation

The segmentation of the extensor tendon was achieved with satisfactory results using manual annotations of the bones and skin line. However to achieve a fully automatic system, these two components need to be obtained automatically. Automatic bone segmentation was proposed in Chapter 5 and will be used. The skin line was not addressed, and a simple approach based on [97], is used here.

In Fig. 6.9 is presented the diagram of the proposed fully automatic extensor tendon segmentation method.

The input image is processed independently to extract the skin line and the bones (metacarpus and phalange). The metacarpus and phalange segmentation follow the proposal of Chapter 5, and the skin line will be briefly explained next. The extensor tendon segmentation uses the information of the input image and the results of the bones and skin line segmentation to obtain the final result. Regarding the skin line, the method starts with a pre-processing phase, using the Total Variation Denoising method (TVD) [98] with lambda equal to 4.0 and 40 iterations. Then, a global threshold equal to the 30% percentile is used, followed by an area blob filter (50 pixels) and a region filter

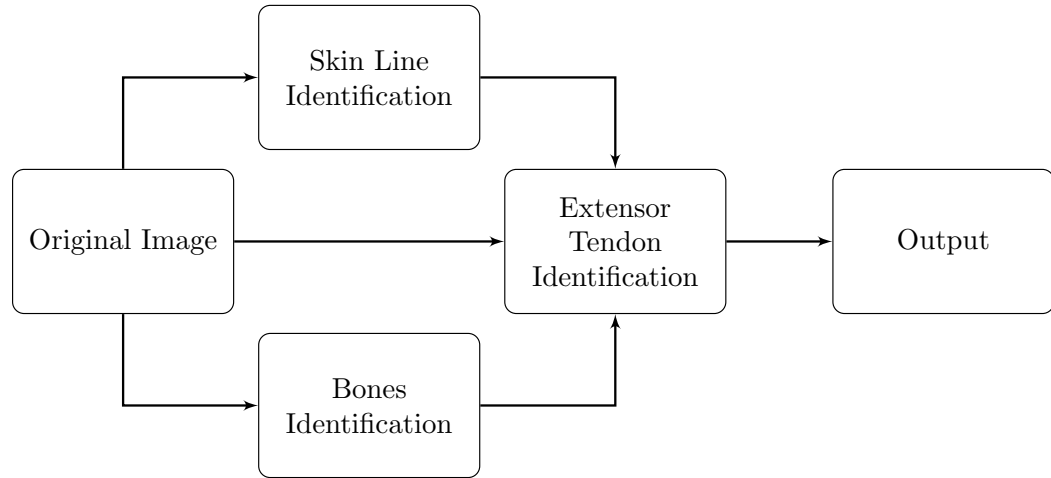


FIGURE 6.9: Diagram of the fully automatic extensor tendon segmentation method.

that keeps only the objects in the top of the image (10% of the top rows). The resulting objects are used as the input of the LAC (4.3) to refine the contours, resulting in the final skin line segmentation.

Fig. 6.10 shows the results for the fully automatic segmentation method on the original 125 test images.

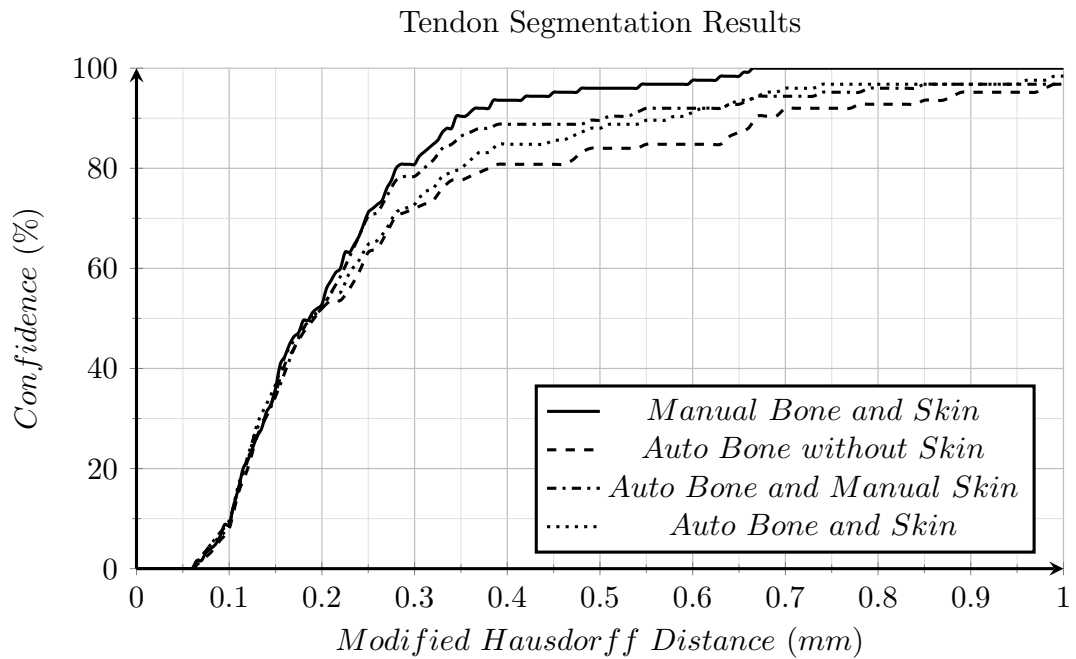


FIGURE 6.10: Results obtained in the test data for the different automatic and manual segmentation settings.

From Fig. 6.10, it is possible to see that, as expected, the results are best when the ground truth is used. Alternately, when automatic segmentation methods are used, the *confidence* decreases, which is justifiable due to the bone and skin line not being perfect. Nevertheless, the decrease in performance is acceptable, dropping from approximately 95% to 88% for a 0.5 mm error. From these results, it is also possible to verify the importance of the skin line in the tendon segmentation of the proposed method. When no skin line information is used, the *confidence* dropped from 89% to 84% for the same 0.5 mm error. In Table 6.1, we present three specific MHD errors.

TABLE 6.1: *Confidence* obtained for the different configurations and for three MHD.

	0.2 mm	0.5 mm	0.8 mm
Manual Bone and Skin	52%	95%	100%
Auto Bone and Manual Skin	52%	90%	95%
Auto Bone and Skin	52%	88%	96%
Auto Bone Without Skin	52%	84%	91%

From Table 6.1, it is easy to see that the *confidence* for small errors (less than 0.2 mm) are equal for all configurations, indicating that some images are not affected by the inclusion of the automatic segmentations. After visual inspection, it was concluded that they represented images with less challenging characteristics, that is, no pathology and well-defined structures, which lead to perfect bone and skin line segmentations. Alternately, for high errors, only the configuration with manual bones and skin line lead to 100% *confidence*, where the other methods stayed at 96%, 95% and 91%. This is due to failings in the bone or skin line segmentation because of extreme deformations in the joints. We consider these extreme cases to be out of the scope of this work because they represent patients with advanced disease stages where early detection is not possible and follow-up is not as important as in other stages. In Fig. 6.11, some of the visual results obtained are presented.

In the first and second rows, it is possible to observe the importance of the skin line to avoid local minimums related with the sub-cutaneous fat layer. In the second row, the automatic skin line detector failed to detect the left portion of the skin line and the fake skin line (present on both sides), which compromise the final results on the left side. In the third row, it is possible to see an example where an inefficient segmentation of the bone and skin line did not interfere significantly with the extensor tendon segmentation. This result was only possible due to the sparse representation of the tendon, using a

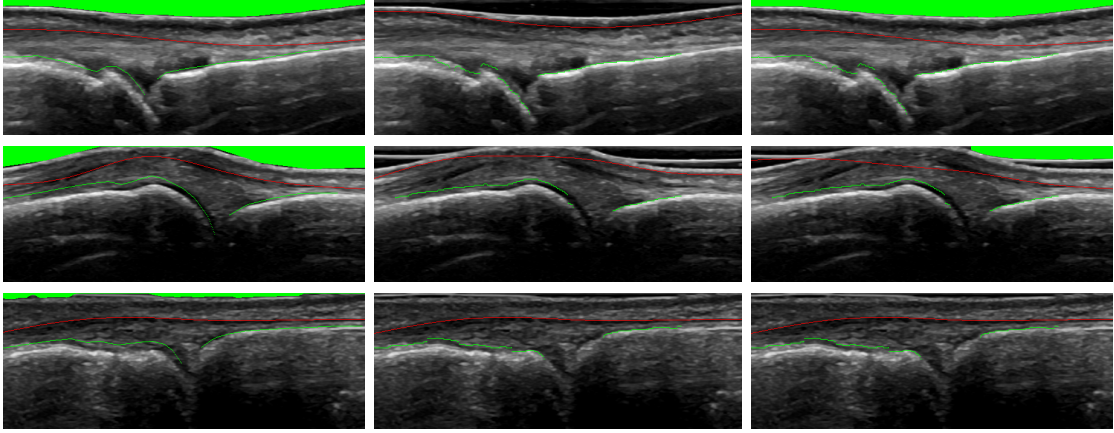


FIGURE 6.11: Visual results of the fully automatic segmentation of the extensor tendon for different configurations. In the rows are different cases and, in the columns, the different configurations, respectively from left to right: manual bone and skin line; automatic bones without skin line; all automatic. The green top region is the region above the skin line, the green lines are the bones and the red line is the automatic tendon segmentation.

small number of points in the interpolation and the use of additional restrictions that impose continuity and shape constraints. In this case, due to the small size of the skin line region, its identification was not fundamental to a proper segmentation.

6.5 Conclusions

A new algorithm for the segmentation of the extensor tendon using an active contours framework with phase symmetry pre-processing and prior knowledge energy formulations was proposed. It has the advantages of not requiring an initialization and that the derivative of the energies does not need to be computed, which allows more flexible energy formulations. The results showed that segmentation is possible with an error below 0.5 mm for a *confidence* of 95%. Although lacking more robust medical evidence, this not only seems to be a small error from a clinical perspective but a surprisingly good one from a visual segmentation problem perspective since the tendon is not always visible and the number of similar structures in the image creates several local minimums. Nevertheless, further work needs to be done, to verify if this error is indeed acceptable when extracting information from images. Regarding the pre-processing, we concluded that the *PhaseSymmetry* is a valid option for this application, considerably improving the segmentation. Attending to the energies, the inclusion of the *NeighborhoodWeights*

also led to an improvement of the results. For the *Slope*, *Concavity* and *AreaRatio* energies, the improvement in the final solution was small but important in the *Intensity* formulation. Regarding the parameters, most of them can be obtained from the images using statistics and prior knowledge of the structures under study, for example, the *AreaRatio* and *Slope*. Alternately, the energy weights are harder to infer, and because of that, a grid search approach was used. It is possible that some of our proposed constraints may not be valid for extreme deformations. Given that, we are motivated by the screening and quantification of rheumatic diseases, and these extreme deformations are considered beyond the scope of this work. They are typically easy to identify by visual inspection alone and without the need for precise quantifications such as the ones this work aims to contribute. The database size is acceptable but unbalanced, with less pathological entries than non-pathological, which might create a bias in the results. This problem was minimized using a balanced set in the training.

An extension of the work proposed in [8] was also presented, by including additional data and proposing a fully automatic method. Quantitative and qualitative results have shown that the extensor tendon method is robust when noise from other automatic methods is added, instead of a perfect scenario where manual annotations are used. It was able to handle some missing information due to the use of different energies, which enforced the model. From the results, it is possible to conclude that fully automatic segmentation is possible, despite a loss of 7% in the overall *confidence*. Improvements in the bone and skin line segmentation could lead to an increase in the performance and approximate it to the results of [8]. Another possible improvement is the use of cross information from the different segmentation methods instead of independent skin line and bones segmentations, which might improve the performance of all methods, for instance, in the cases where one of the structures is less visible. A relaxation of the problem, allowing user interaction with the input of a small number of points, could also be interesting, since it is a middle point between the best results and the less manual effort.

In the future, more tests will be conducted on images with pathologies and integrated with other automatic skin line and bone segmentation methods aiming to improve the results of the fully automatic segmentation system. Tests on other types of images may also be conducted since this algorithm is not specific to the segmentation of the

extensor tendon. Extension of the framework for 3D images (surfaces) is also possible but out of the scope of this work.

Chapter 7

Joint Capsule Identification

Contents

7.1	Introduction	81
7.2	Proposed Work	84
7.2.1	Split and Merge	84
7.2.2	Convolutional Neural Networks	88
7.3	Results and Discussion	92
7.3.1	Metrics	93
7.3.2	Results - <i>SM</i>	93
7.3.3	Results - CNN	96
7.3.4	Comparative Results and Discussion	97
7.4	Conclusions	100

7.1 Introduction

This chapter addresses the joint capsule identification in ultrasound images of the metacarpophalangeal joint (MCPJ). The joint capsule is one of the main structures studied in this work. This structure gives clues about the presence of synovitis and bone erosions, which are related with the presence of Rheumatoid Arthritis (RA). As previously discussed in Chapter 3, RA affects primarily the joint capsule of the small joints as a persisting inflammation, which causes joint damage over time. The correct identification and measurement of the swelling created by this inflammation could

be used as a direct measure of RA progression. Given this, the main motivation for this work is the automatic identification of the joint capsule in ultrasound images of the MCPJ. Then, this information can be used to automatically identify and quantify the presence of synovitis and RA.

The segmentation of the joint capsule is not a trivial task, due to the presence of artifacts and high shape variability, as seen in Fig. 7.1.

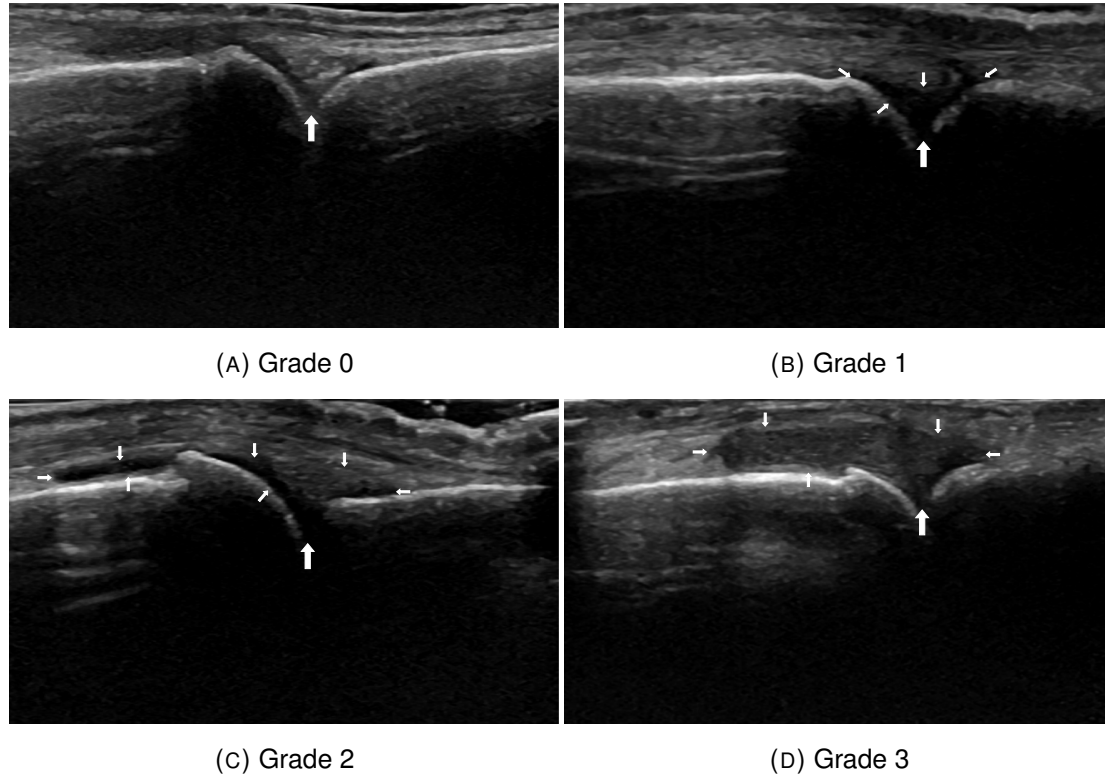


FIGURE 7.1: Example of ultrasound images of the second MCPJ with different joint capsules aspects and synovitis grades. Small arrows indicate the joint capsule extension and big arrow indicates the joint.

Fig. 7.1, gives an idea of the shape variability of the joint capsule. From Fig. 7.1-A to Fig. 7.1 - D increasing synovitis degrees are presented (degree 0 to 3, respectively) and it is possible to see the joint capsule growth pattern. It starts in the MCPJ, then grows over the metacarpus and after that over the phalange, resembling a balloon inflating. Regarding the expected variations and artifacts, Fig. 7.1 - A shows an anisotropy artifact. This artifact creates a region with the same gray-scale pattern as the typical synovitis region, which increases the difficulty of the segmentation. In Fig 7.1 - C the joint capsule is not totally connected, which creates two structures that belong to the same bone structure. Moreover, the joint capsule and the cartilage have similar

echogenicity patterns and, because of that, it was decided to identify both as the same structure leaving as future work the creation of classification criteria to separate them.

Attending to the literature, it is possible to conclude that the detection of the joint capsule is still an open problem. Given that the automatic identification of this structure is far from trivial, some authors used semi-automatic approaches, which reduce considerably the complexity of the segmentation. In [99], the authors, proposed a semi-automatic method for the detection of the synovitis contours in ultrasound images of the MCPJ and proximal interphalangeal joint. The algorithm starts with a pre-processing step, which is based on the minimization of local variations and the assumption that the noise follows a Rice distribution. Then, the user must choose three points, two in the synovial limits in the bones and the third point is used to define the synovial extension between the limits of the bones and the articulation. After that two sets of active contours are created. The first defines the contours of the hyper proliferation of the synovial membrane and is used to initialize the second contour that will separate the synovial membrane and the soft tissues. In [97] and [100], the authors proposed an automatic method to identify the joint capsule. The algorithm starts with the identification of the skin border and bones; then a coarse segmentation is performed using a global threshold obtained from the accumulative histogram (at 65%). This coarse segmentation is improved using the skin border, the bones' segmentation and a confidence map. The author reported satisfactory visual results from experts. In terms of Deep Learning techniques in the ultrasound images of the MCPJ, there were no works found in the literature. The closest one was the work of [72], where CNNs were used to identify hip dysplasia in ultrasound images. In [71] there is a broad overview of the existing methods using DL in the medical field. With respect to the segmentation of 2D images, the authors emphasize the *UNet* model, published by [101], which leads to a significant improvement in the segmentation of cells in electron microscopic stacks using a small training set. Given that the database available in the present work is also limited, it was decided to use the *UNet* as a starting point for this task.

7.2 Proposed Work

Following the ideas presented with Fig. 7.1, it was decided to create an algorithm that mimics the growth of the joint capsule when synovitis is present. A split and merge approach is proposed because it allows the use of region growing strategy that can be constrained to force the final shape to resemble the joint capsule. Additionally, an approach using CNNs, more precisely the *UNet* model, was included, because these models are the current state of the art approach for segmentation and classification. Next, both approaches will be discussed in more detail.

7.2.1 Split and Merge

The diagram of the proposed method is shown in Fig. 7.2.

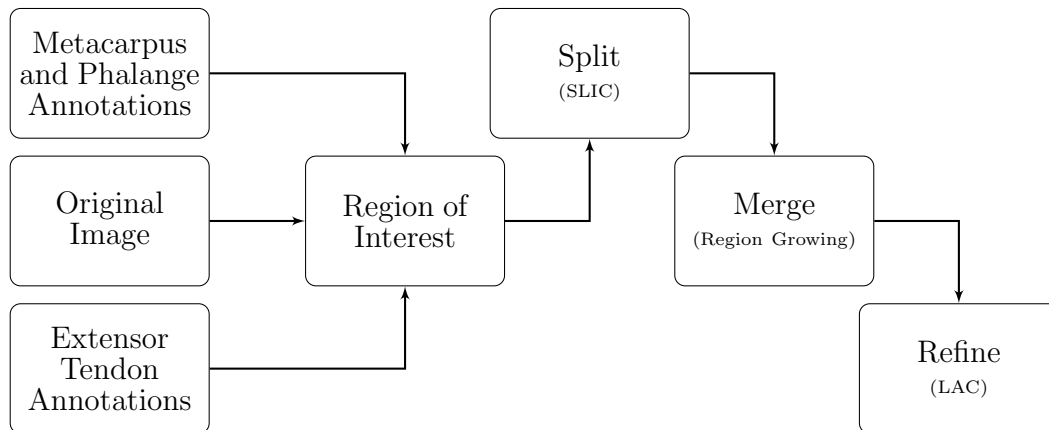


FIGURE 7.2: Joint capsule segmentation pipeline using a Split-Merge-Refine approach ($SM - R$).

The first step defines the region of interest using as input the location of the metacarpus, phalange and extensor tendon. Afterwards, the SLIC method [89] is used in the split step to reduce the number of elements in the merging step by grouping pixels with high similarity in individual sets, i.e., super-pixels (*SPs*). This step also improves the robustness, because the new sets are less related with their neighbors. The merging step is achieved with an adapted region growing method where shape constraints are added to force the final shape to be similar to the typical and expected joint capsule shape. Finally, the Localizing Active Contours (LAC) method [87] is used to smooth the final contours while adapting them to small intensity transitions.

Region of Interest

The first step of the proposed method uses the annotations of the metacarpus, phalange and extensor tendon to reduce the search region. This restriction reduces the variability between images and removes the interference from other structures, such as the skin line and the extensor tendon. The metacarpus and phalange annotations were also used to obtain the reference joint point, $P_0 = (x_0, y_0)$ (big arrow in Fig. 7.1), which will be used to start the merge process. P_0 is calculated using the metacarpus rightmost point, $P_m = (x_m, y_m)$, and the phalange leftmost point, $P_p = (x_p, y_p)$ as:

$$P_0 = \left(\frac{x_m + x_p}{2}, \max(y_m, y_p) \right) \quad (7.1)$$

In short, x is the middle point between the end of metacarpus and the beginning of the phalange and y is the deepest point between the two.

Split

Starting from the region of interest previously obtained, the SLIC algorithm [89] groups individual pixels in clusters (SPs) attending to their spatial and intensity distances. This algorithm can be seen as a k-means algorithm with an intensity and a spatial component. The idea is that closer pixels with similar intensities belong to the same SP , while pixels that are spatially distant and/or with distinct intensity values belong to different SPs . This is achieved by defining a distance metric, D , that is influenced by the spatial distance and the distance between intensities:

$$D = \sqrt{\left(\frac{d_c}{m}\right)^2 + \left(\frac{d_s}{S}\right)^2} \quad (7.2)$$

where d_s is the spatial Euclidean distance between a pixel and the SPs centroids, d_c is the intensity (or color) distance, m is a control variable that balances the contribution of d_s and d_c . S is the expected size of each SP and is defined as:

$$S = \sqrt{\frac{N}{k}} \quad (7.3)$$

where N is the number of pixels in the image and k is the approximated number of SP s. For more details, please refer to Chapter 4 Section 4.5. Fig. 7.3 shows the result of the SLIC, with $k = 2000$ and $m = 15$.

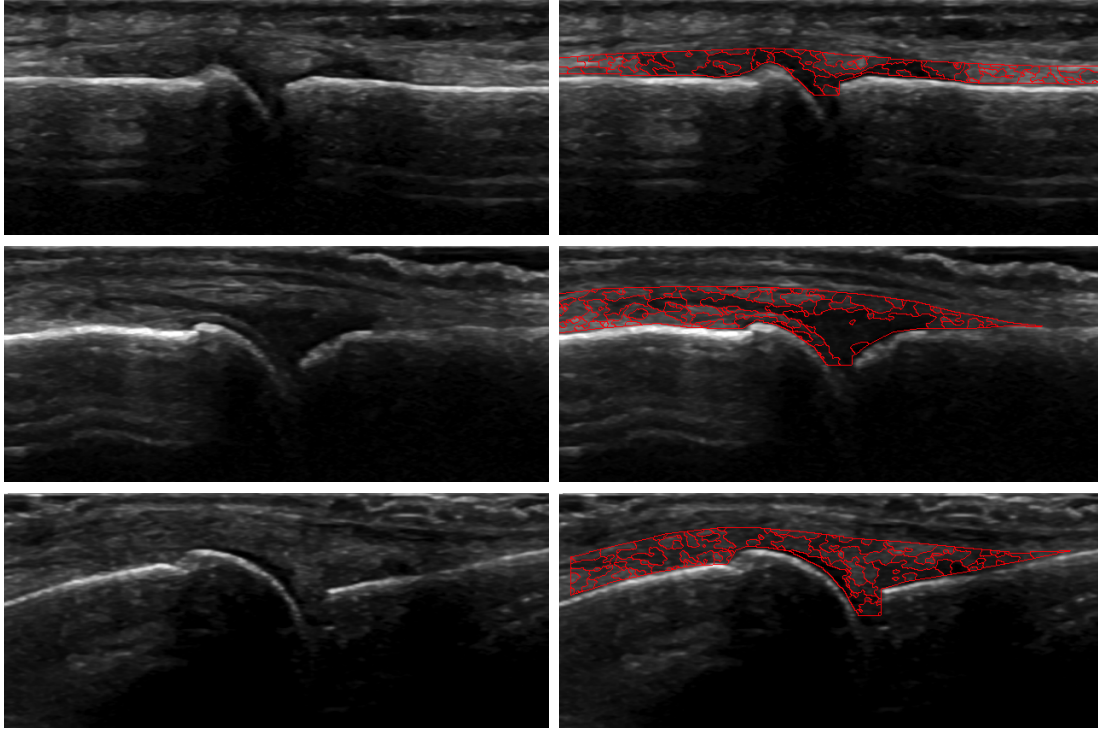


FIGURE 7.3: Results of the SLIC algorithm in the MCPJ images.

Merge

The problem with the SLIC method is that it creates many more objects than the ones present in the image (over-segmentation), consequently, the resulting SP s need to be merged into the final object. The DBSCAN [102] is typically used in these situations, but, in our case, it was not producing the expected results since it does not take into account the expected shape of the object. Because of that, an alternative approach is proposed for this specific problem. Taking into account that the bones and tendon segmentation is available, it was decided to create an adapted version of the Region Growing algorithm. It is similar to the original algorithm, but instead of pixels, the growth is done in the SP domain. To do so, the seed point P_0 is used to identify the first super-pixel, SP_0 , which is used to start the growing process. The pixel neighbors of the region growing method are replaced by a neighborhood vector, \mathcal{N} , which identifies the super-pixels that are connected to the growing object. The super-pixels are then iteratively joined until no more objects in \mathcal{N} can be added. In Algorithm 1, the pseudo-code of the merging process is presented.

Algorithm 1: Super-pixel merging process with shape constraints.

```

1 Merge ( $SP_0, SP_{All}$ );
   Input :  $SP_0$  - Seed Super-pixel;
            $SP_{All}$  - All Super-pixels;
   Output:  $SP_F$  - Final Super-pixel results;
2  $SP_F \leftarrow SP_0$ ;
3  $\mathcal{N} \leftarrow$  Get Neighbors of  $SP_F$  from  $SP_{All}$ ;
4 while  $\mathcal{N} \neq \emptyset$  do
5    $(SP_n, \Delta g) \leftarrow$  Find  $SP_n \in \mathcal{N}$  with smaller gray scale distance ( $\Delta g$ ) from  $SP_0$ ;
6   if ( $\Delta g > T_{\Delta g}$ ) then
7     break;
8   end
9   if ( $SP_F \cup SP_n$ ) results in allowed shape then
10     $SP_F \leftarrow SP_F \cup SP_n$ ;
11     $\mathcal{N} \leftarrow$  Get Neighbors of  $SP_F$  from  $SP_{All}$ ;
12  else
13     $\mathcal{N} \leftarrow \mathcal{N} \setminus SP_n$ ;
14  end
15 end
16 return  $SP_F$ ;

```

The Merge algorithm has two main components, one related with the intensity (gray-scale) and another related with the expected shape. The first is present in Algorithm 1 - Line 5 and is defined as the Euclidean distance from the median intensity of the elements in SP_n and SP_0 . If all SP s distances are greater than a given threshold, $T_{\Delta g}$, the merging process stops. If the distance of one SP is smaller than $T_{\Delta g}$ then it is verified if the inclusion of SP_n in the final object creates an allowed shape (Algorithm 1 - Line 9). This is achieved using a shape constraint defined as the *DICE* between $SP_F \cup SP_n$, and its own convex hull. This is based on the assumption that the joint capsule behaves like a balloon and, therefore, it should have a shape that is nearly convex. Given that, it is expected that the *DICE* between these two objects produces a high value, and thus, a threshold of 0.5 was empirically selected for this problem. Finally, $T_{\Delta g}$ was empirically set to 30.

Refine

The segmentation result obtained from the merge process creates a spiky appearance. Since the joint capsule is expected to be a smooth region, it was decided to add an additional shape smoothing step using the LAC algorithm. Other post-processing techniques could be used, such as morphological operators, but the LAC was selected since the results were more consistent. The uniform model [87] was used with a small

number of iterations (25) because the LAC is used here for refinement and not for segmentation.

7.2.2 Convolutional Neural Networks

One limitation of the previously method is that it needs additional information other than the image itself, to segment the joint capsule, i.e., the metacarpus, phalange, and extensor tendon locations. This increases the complexity of the method and creates a dependency on other segmentation methods. In order to overcome this problem, it is proposed the use of CNNs to identify the joint capsule in a single step. This is the current state of the art approach to solve several image processing problems. From classification to segmentation, the CNNs achieved better results than most of the previously used techniques [100, 101, 103]. One problem, is that the typical approaches using CNNs require an extensive database to properly parameterize the network. Because of that, and after the literature review, the *UNet* model seemed promising and it was included in the present study. Next it will be presented in more detail.

The *UNet* model

As previously said, the number of images available in this work is relatively small. To overcome this limitation, the work of [101] was used since it proposes an architecture that was able to train with small number of observations. This new architecture is called *UNet* and tries to capture local and global information in a single model, by adding skip connections before every *max - pool* operation. This ensures that the local information lost after the *max - pool* is kept and added later in the network. Moreover, the authors used data augmentation to create synthetic data and an improved objective function to enforce separation between classes. Fig 7.4 allows a better visualization of the *UNet* model.

The addition of skip connections (represented by the gray arrows in Fig. 7.4) increases the detail of the segmentation by preserving the local information before pooling operations. This architecture creates a multi-scale representation where coarse information is gathered in the bottom and finer details in the top of the model, as seen in Fig. 7.4. In the work of [104, 105], this effect was clearly shown.

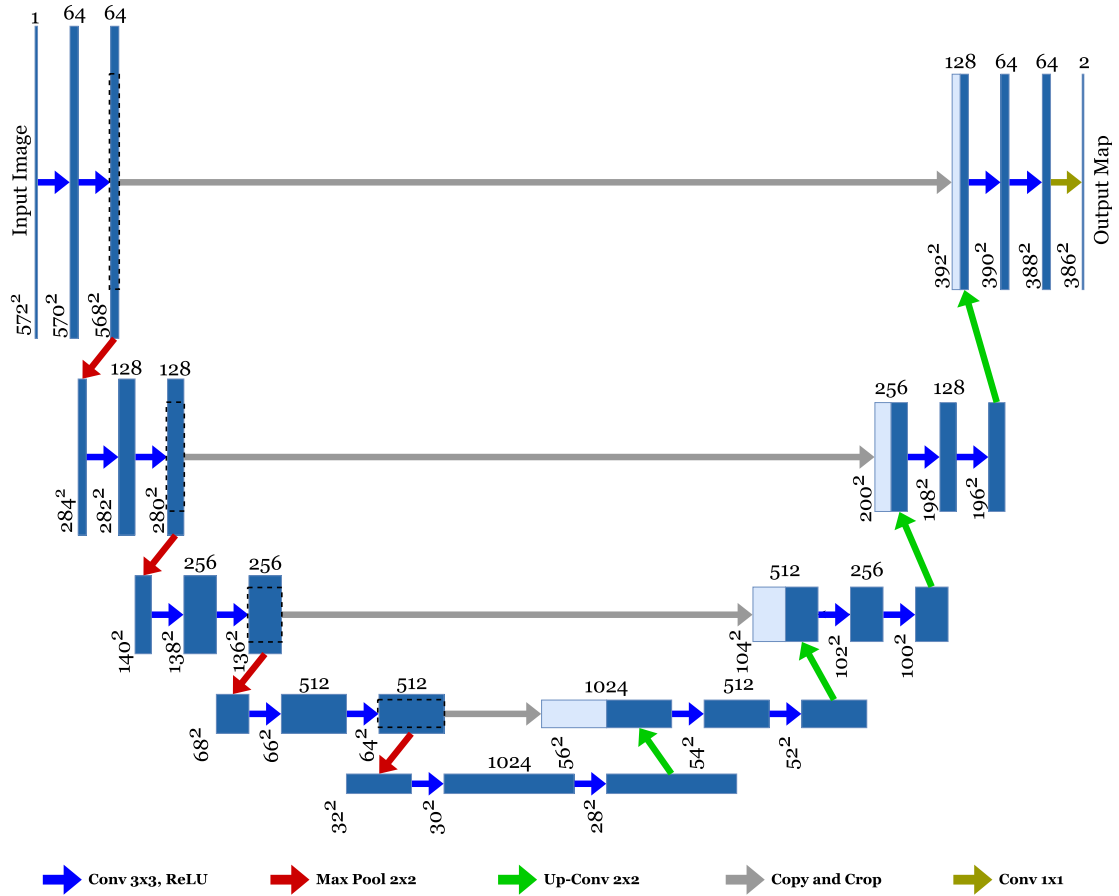


FIGURE 7.4: The *UNet* model proposed by [101], with a depth of 4 and 64 neurons in the first layer.

Fig. 7.5 shows that every additional skip connections increases the detail in the output mask.

As referred earlier, the main advantage of this approach is that only the image is used as input, reducing the dependency on other inputs. The main disadvantage is that a training phase is needed to parameterize the network and, depending on its size and on the problem complexity, the number of examples necessary can be higher than the ones available. Nevertheless, it was decided to include it in this study because there is evidence that CNNs can perform well for complex tasks and there are ways to reduce the number of examples necessary to train the network, by creating synthetic data.

Model adaptation

The *UNet* model proposed in [101] was formulated to solve a multi-object segmentation problem with touching boundaries. Here, the problem is the detection of a single object

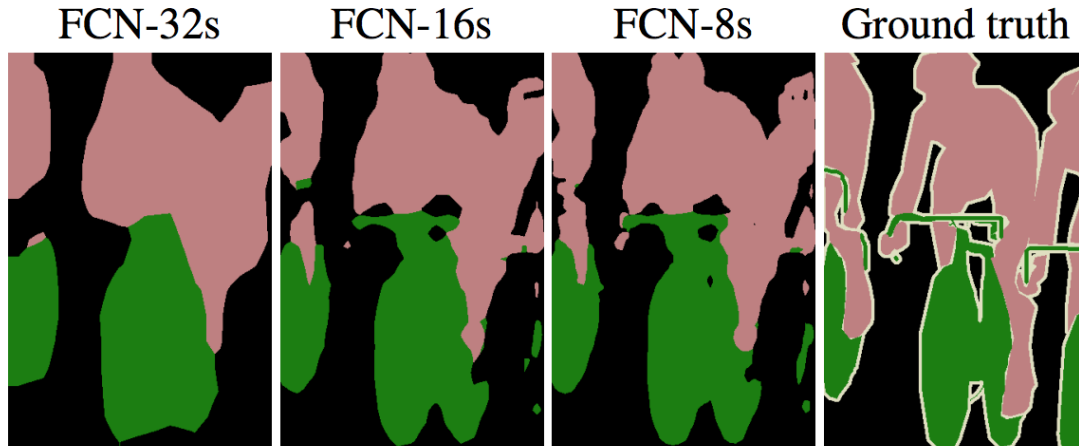


FIGURE 7.5: Result of adding skip connections. FCN-32s - no skip connections; FCN-16s - 1 skip connections; FCN-8s - 2 skip connections. Image from [105].

(joint capsule) and consequently, the *UNet* model needs to be adapted to this specific problem. Those changes will be addressed in the following subsections.

Architecture:

The architecture proposed for this problem is depicted in Fig. 7.6. The first architectural change is the use of zero padding in the convolutions instead of the valid values. The authors of [101] proposed the use of the valid values with a mirror padding pre-processing, so the final outputs have the same spatial dimensions as the original input image. They argue that this accelerates the training, however this was not observed during preliminary tests and therefore, it was decided to use zero padding in all convolutions. The second change is the use of batch normalization layers before the concatenation steps. This adds a regularization effect by ensuring that the concatenated feature maps have the same order of magnitude. In the last convolution layer, the Softmax activation function was changed to a Sigmoid function.

Data Augmentation:

Along with the *UNet* architecture, [101] used a special data augmentation technique to compensate the reduced number of training samples. In that work, elastic transforms were used to simulate the possible deformations that a cell can experience. In the MCPJ images, that same technique is not usable, since the deformation is constrained by many other factors, like the presence of bones, tendons and rotation points. Because of that, it was decided to create a different set of augmentation techniques. These

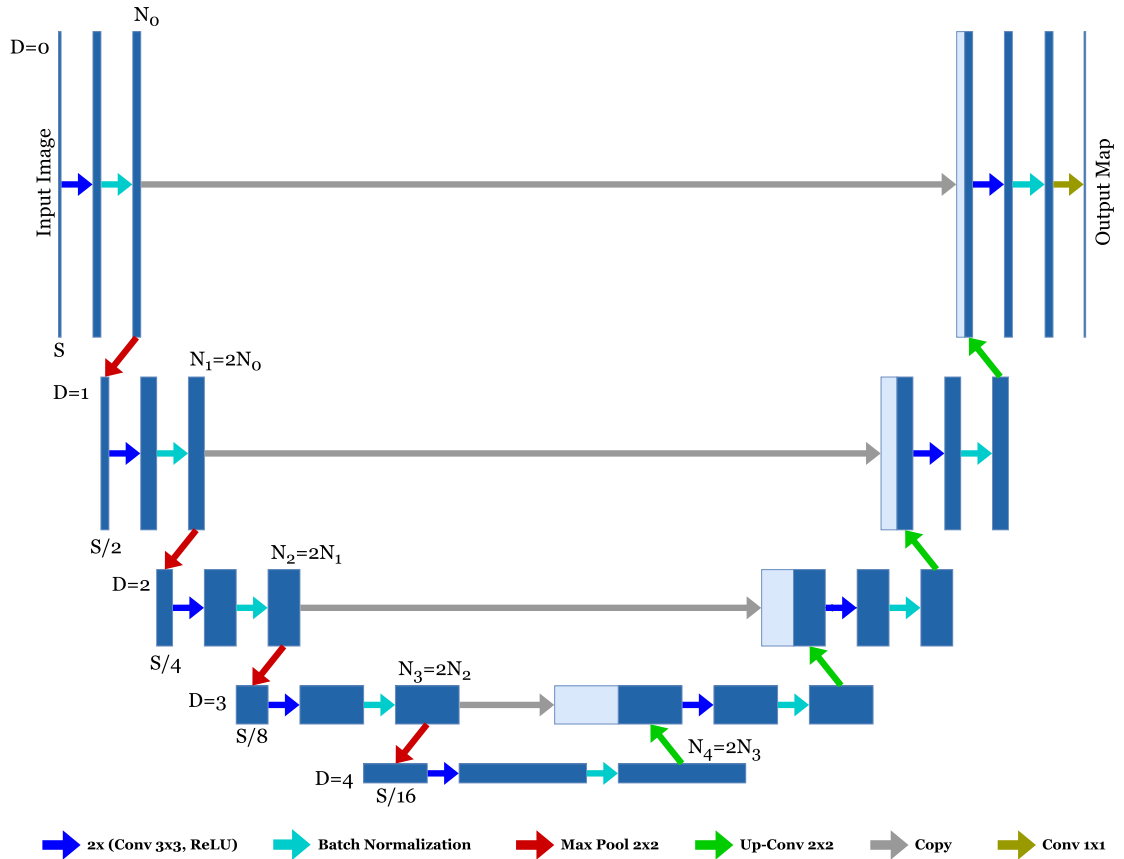


FIGURE 7.6: Proposed changes to the *UNet* model using normalizing layers before every *max - pool* operation and with *same* padding in all convolutional layers. N_x stands for the number of filters used, D for depth and S for the spatial size of the feature map.

techniques are expected to improve the generalization capabilities of the models by creating additional artificial data with different characteristics, not present in the original database.

- Scale - simulates anatomical differences in size (larger/smaller joints) and the ultrasound device acquisition parameters: depth, probe width;
- Horizontal shifts - simulates different probe positions along the joint;
- Illumination - simulates different gray-scale maps used by the ultrasound device operators;
- Noise - simulates the noise from different equipment;
- Horizontal flips - simulates the transformation that occurs when the operator rotates the probe 180° .

The input image was scaled using a bicubic interpolation, while the output mask was scaled using the nearest neighbor interpolation. The range of the scale parameter was empirically selected to be from 0.8 to 1.2. The horizontal shift was done using a random crop operation inside the valid zone of the resulting scaled image. The illumination was achieved using a random scalar value added to the input image between -20 and 20. The noise was added using a random noise generator with mean zero and standard deviation 10. The horizontal flips were done with a simple mirror operation in the image. In each epoch, 10 new images are randomly created for each original image. In this way, it is expected to increase the variability of the training data.

Other changes:

The input image was trimmed to 256 lines and 736 columns to avoid rounding errors when performing downsampling and upsampling and to reduce the bottom dark region of the images. The optimization method was the adaptive momentum (Adam) with default parameters.

The objective function was the mean squared error instead of the one proposed in [101], which took into account the existence of touching boundaries and multiple objects on the image, which is not adequate for this situation.

7.3 Results and Discussion

In this chapter, two new approaches for the segmentation of the joint capsule are proposed and compared with the work of [100] (hereinafter referred as *Ref*). In order to do so, a database consisting of 243 ultrasound images of the MCPJ was used (around 110 patients). The phalange, metacarpus, skin line, extensor tendon, and joint capsule were manually annotated in all images, using a self owned dedicated software as described in A.1.

The *SM* and *Ref* algorithms were evaluated in all images of the database and the inclusion of the post-processing was also tested. The *UNet* models were trained in a set of 193 images and tested in 50 different images. The division was done manually to ensure that different cases are present in each set. The database is relatively small and, because of that, a random split method might create bias in the results. The manual

division took into account the synovitis degree and the presence of other conditions such as osteophytes and bone erosions. The test set of the *UNet* was used to compare the three methods, since it is the only set that contains unbiased images (not used to train the *UNet*).

7.3.1 Metrics

In order to compare the proposed methods, the Boxplot of the Dice Similarity Coefficient (*DICE*) will be used.

The *DICE* Boxplots are useful to visualize the distribution of the results and the *Welch's t - test* will be used to verify the statistical significance among them. Finally, in order to understand the overall performance of each method, a plot of *DICE* vs *confidence* was created. The *confidence* is the percentage of images segmented with a *DICE* higher than a given threshold (between 0 and 1). This plot is helpful to study the possible operation points by looking at the trade-off between the *confidence* and *DICE*. More detail on the *DICE* was presented on Section 4.7.2

7.3.2 Results - *SM*

The *SM* approach was tested with and without post-processing and compared with the algorithm proposed by [100], creating a total of four possible configurations (*Ref*, *SM*, *Ref - R* and *SM - R*). From the whole dataset, 8 images were discarded since they did not have a visible joint capsule, resulting in *DICE* values of 0 in all configurations. The results for the remaining images are shown in Fig. 7.7

From the analysis of Fig. 7.7, it is possible to see that the *SM* and *Ref* have very similar results. The refinement step has a positive effect, improving the results in both with higher impact on the *SM*, which lead to the *SM - R* to outperform the *Ref - R*. All methods have several outliers in the bottom, pointing that some images were poorly handled by these methods.

The *Welch's t - test* was used to calculate the *p*-values between each configuration. In Table 7.1 these results are shown.

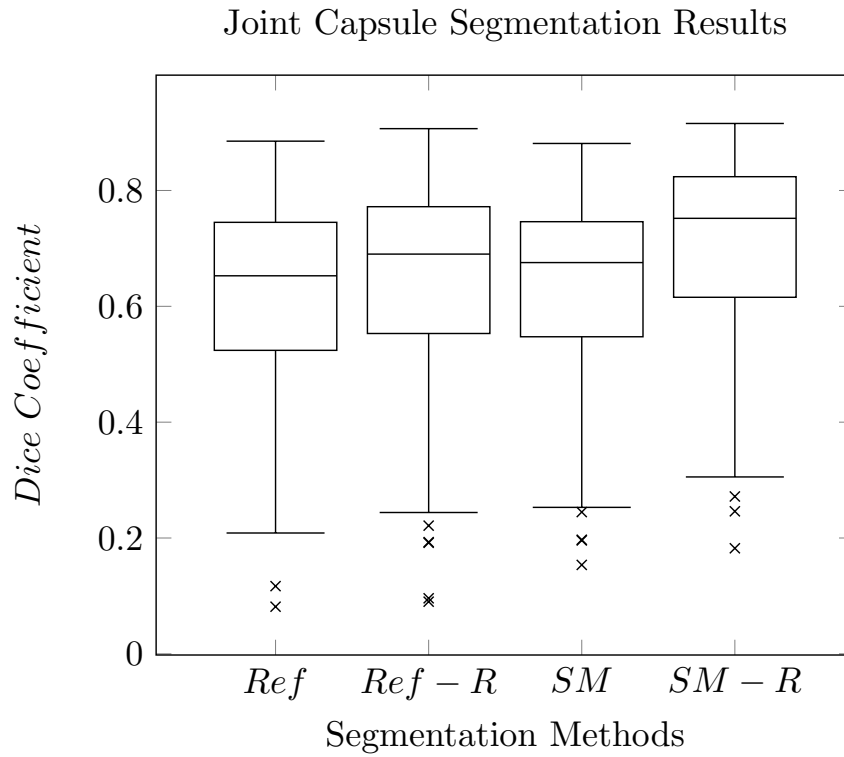


FIGURE 7.7: Boxplot of the *DICE* values obtained for the different configurations; *SM* - Split and Merge; *R* - Refine; *Ref* - [100]. The \times are the outliers.

TABLE 7.1: *p*-values resulting from the *Welch's t - test* of different pairs of methods.

Method 1	Method 2	<i>p</i> -value
<i>Ref</i>	<i>SM</i>	0.45694
<i>Ref</i>	<i>Ref - R</i>	0.04172
<i>SM</i>	<i>SM - R</i>	$3.24E^{-6}$
<i>SM - R</i>	<i>Ref - R</i>	0.00143

Table 7.1 confirms that the differences previously pointed out are indeed statistically significant ($p\text{-value} < 0.05$). It is also possible to verify that the *Ref* and *SM* are statistically equivalent ($p\text{-value} = 0.46$).

In Fig. 7.8, the plot of the *confidence* vs the *DICE* is presented.

Fig. 7.8 confirms the results shown in Fig. 7.7 and Table 7.1. Furthermore, it is also possible to verify that, for a *confidence* of 95%, the results are far from the optimal, with a *DICE* smaller than 0.2. For a *DICE* 0.7, the best *confidence* value was in the *SM - R* method with 60% of the images correctly identified.

In Fig. 7.9, it is possible to visualize some of the outputs of each method.

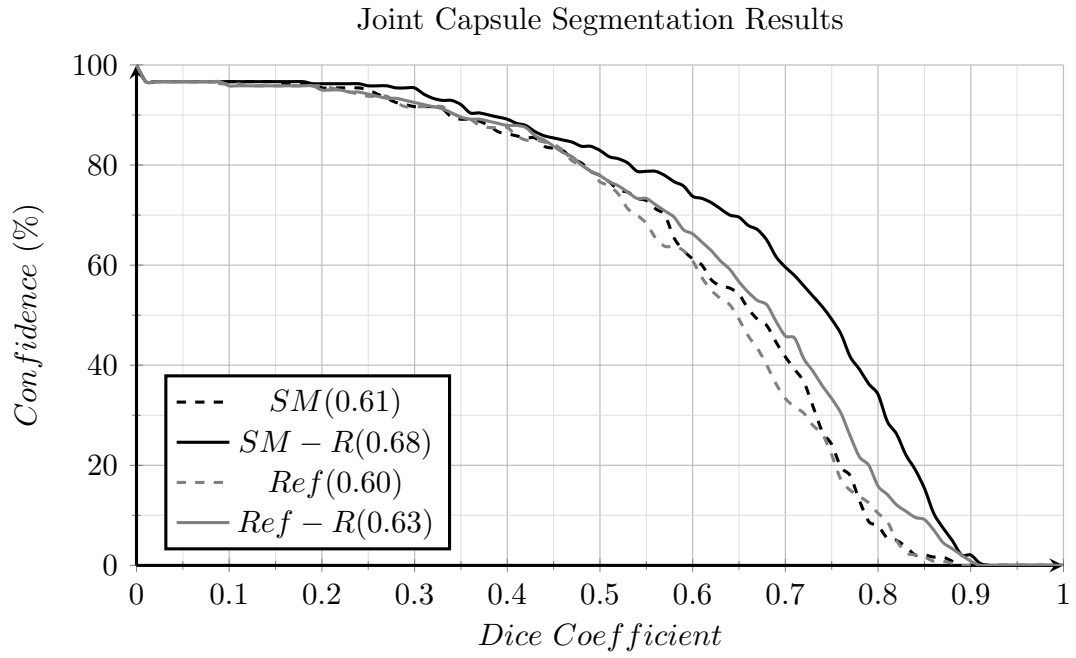


FIGURE 7.8: Segmentation results obtained for the different configurations. The values between round brackets are the Area Under the Curve (AUC); SM - Split and Merge; R - Refine; Ref - [100].

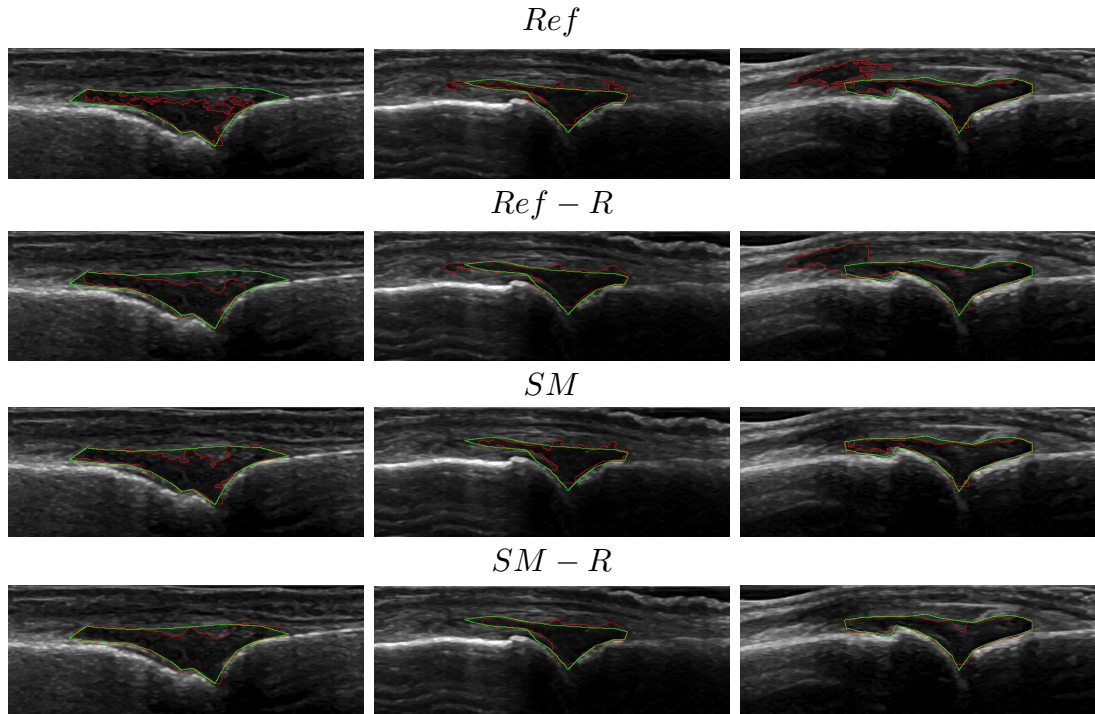


FIGURE 7.9: Visual results obtained for the segmentation of the joint capsule for different configurations. In the columns are different cases and, in the rows, the different configurations: Ref - [100]; SM - Split & Merge; R - Refine. In green is the ground truth and in red is the respective automatic results.

Fig. 7.9 shows the effect of the refinement step (from row 1 to 2 and from row 3 to 4). After the LAC, the contours are much smoother and closer to the manual segmentation/delineations. The last column puts in evidence one of the main limitations of the *Ref* method when an ultrasound shadowing artifact created by the tendon is present. The *Ref* method only takes into account the intensity values, which results in a segmentation that incorrectly includes the artifact region. The inclusion of the shape information in the *SM* method improves the results by stopping the growing process, once the shape starts getting irregular.

7.3.3 Results - CNN

The proposed changes to the *UNet* model were implemented using Python and the TFLearn API with Tensorflow as engine. In order to have a better understanding and achieve the most accurate results, a study of the depth, D , and number of neurons, N_x , of the model was done. Since the number of neurons in each layer is different, it was decided to follow the rule that after every *max - pool* operation the number of neurons is doubled. In this way, only the number of neurons of the first depth level, N_0 , needs to be parameterized. The results obtained in the test set are presented in Table 7.2.

TABLE 7.2: Mean *DICE* results in the test set. In bold are the top 3 results.

$D \backslash N_0$	2	4	8	16	32
2	0.643	0.671	0.663	0.668	0.673
3	0.714	0.702	0.715	0.728	0.739
4	0.765	0.766	0.795	0.771	0.786
5	0.820	0.795	0.739	0.786	0.731

Table 7.2 allows to conclude that the best model was the one with $D = 5$ and $N_0 = 2$, followed by the models with $D = 5$ and $N_0 = 4$ and $D = 4$ and $N_0 = 8$. In general, the results improved as the depth increases from 2 to 4, while in the depth 5 that trend is not verified. Moreover, the results are more sensitive to the depth of the network than to the number of filters. In Fig. 7.10, it is possible to visualize the outputs of the best configuration of each depth, D .

Fig. 7.10 shows clearly the effect of increasing D , for small values the output has some granularity, while higher values of D result in a smoother output. For small values of

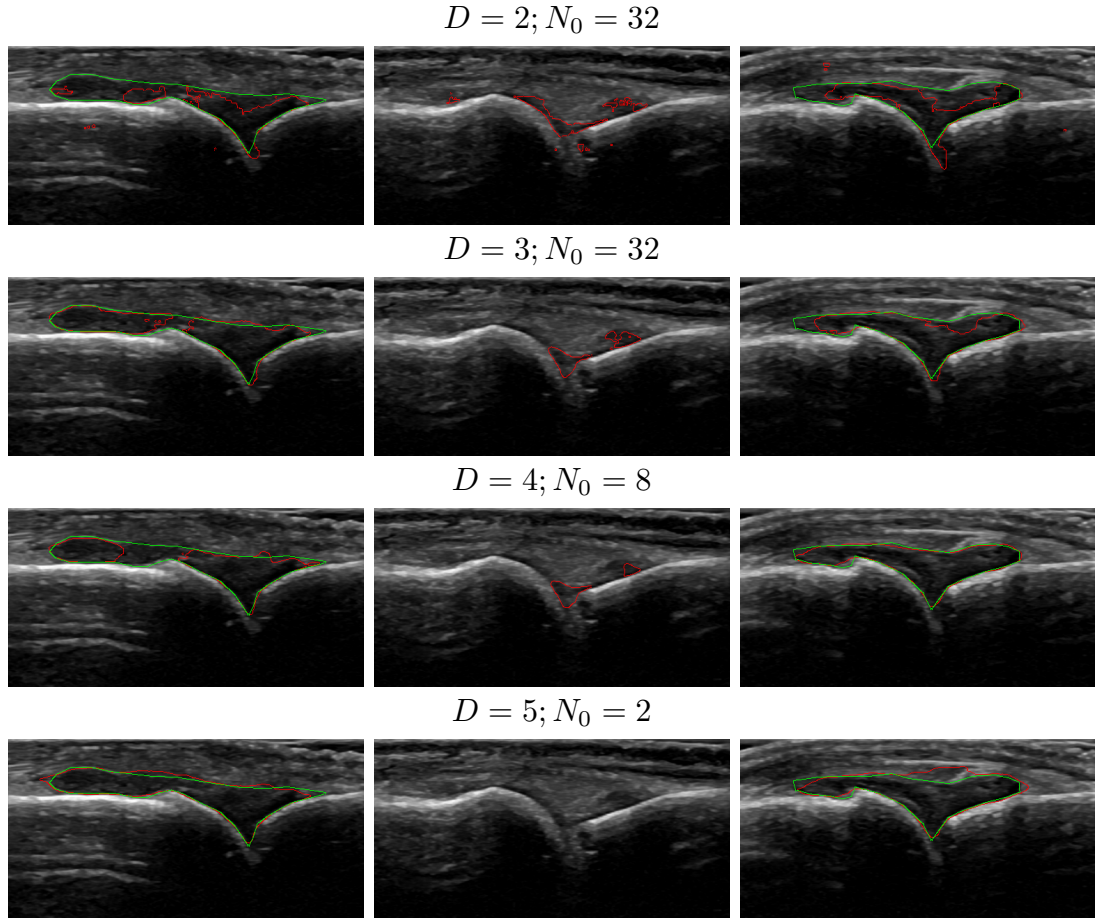


FIGURE 7.10: Visual results obtained for the segmentation of the joint capsule for different *UNet* parameterizations (on top of the images are the values of D and N_0). Each column presents a different case. The delineations in green and red refer to the manual and automatic segmentations, respectively.

D , it is also possible to observe some scattered false positives, which would be easy to remove in a post-processing step, but still they would hardly compare to the results of higher D values.

7.3.4 Comparative Results and Discussion

In this section, the best methods achieved in the previous tests are compared. Fig. 7.11, shows the Boxplot obtained from the *DICE* extracted from the 50 test images:

The boxplot graph indicates that the *UNet* model achieved better results than the *SM-R* and *Ref-R* methods. The distribution of the *UNet* is more compact and has less outliers (only one) than the other methods. Both, the *SM-R* and *Ref-R* have two outliers each with *DICE* of 0, which correspond to images without a visible joint

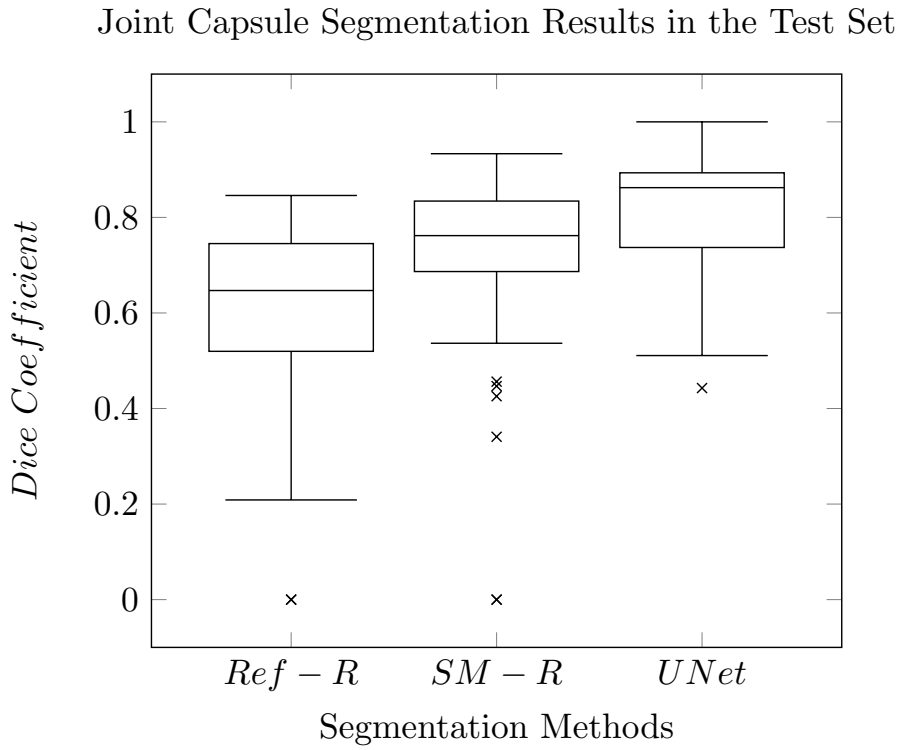


FIGURE 7.11: Boxplot of the *DICE* values obtained for the test images in the best $Ref - R$, $SM - R$ and $UNet$ methods. The \times are the outliers.

capsule. In these cases, these methods do not have the capacity to output an empty segmentation, resulting in false positives.

The statistical significance of these results was evaluated using a *Welch's t - test* and observing the respective *p*-values, which are presented in Table 7.3.

TABLE 7.3: *p*-values resulting from the *Welch's t - test* of different pairs of methods.

Method 1	Method 2	<i>p</i> -value
$Ref - R$	$SM - R$	0.01542
$Ref - R$	$UNet$	$9.21E^{-9}$
$SM - R$	$UNet$	0.00084

Therefore the $SM - R$ is statistically more accurate than the $Ref - R$ and the $UNet$ is statistically more accurate than the other two methods.

In order to understand the possible operating points of the models, it was decided to include the graphic of the percentage of correctly segmented images for different values of the *DICE*, Fig. 7.12

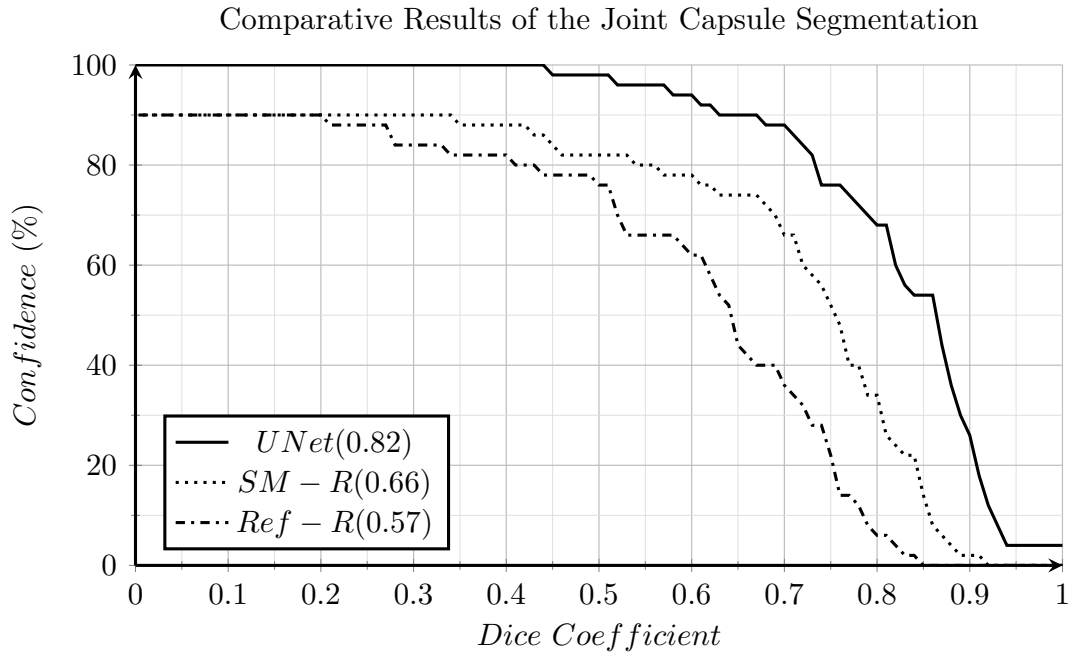


FIGURE 7.12: Results obtained for the segmentation of the joint capsule using three different methods. The values between round brackets are the Area Under the Curve (AUC);

From the analysis of Fig. 7.12, it is possible to conclude that the *UNet* outperforms by far the other two methods with a *confidence* of 90% for a *DICE* higher than 0.67, while the *Ref - R* and the *SM - R* achieved the same *confidence* for a *DICE* of 0.20 and 0.34, respectively. The *UNet* model was also the only one that achieved 100% *confidence* for a *DICE* higher than 0.44.

From the analysis of Fig. 7.13 is possible to verify that the *UNet* tends to produce results that are closer to the manual segmentations. The second row shows one example without joint capsule, which was correctly identified by the *UNet*. This image is responsible for the value 0 of the *DICEs* in the *Ref - R* and *SM - R* observed in Fig. 7.11. In the fourth row, it is shown the example where the *UNet* achieved the worst results (outlier of Fig. 7.11). In this case, the *UNet* was not able to properly identify the joint capsule, even though it seems a trivial case. In this example, the *SM - R* was able to achieve better results, but they are still far from optimal.

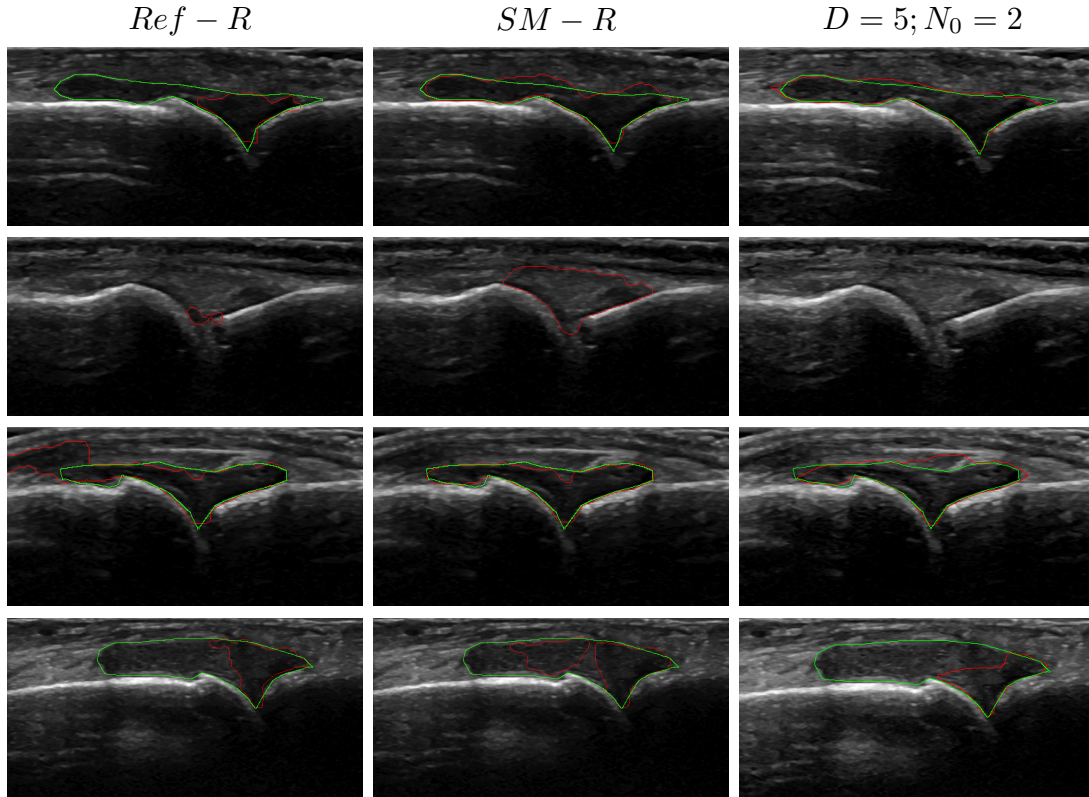


FIGURE 7.13: Visual results obtained for the segmentation of the joint capsule for the $Ref - R$, $SM - R$ and $UNet$ with $D = 5$ and $N_0 = 2$. In the rows are different cases, in green is the ground truth and in red is the respective automatic results.

7.4 Conclusions

Two new methods for the segmentation of the joint capsule in ultrasound images of the MCPJ were proposed, as well as an improvement of the [100] method. Results have shown that the $UNet$ method can segment 95% of the images with a $DICE$ higher than 0.6. The $UNet$ outperformed the $Ref - R$ and $SM - R$ methods by a large margin. Furthermore, the $UNet$ model only uses the information of the image itself, opposing the other methods. which require additional information.

Both the $SM - R$ and $Ref - R$ require annotations to create the initial region of interest. The main difference is that the $Ref - R$ needs the annotations of the skin line and the $SM - R$ method needs the annotations of the extensor tendon. The extensor tendon creates a smaller region of interest, which is more relevant for the identification of the joint capsule. The skin line might be easier to obtain manually and automatically, however it does not remove the sub-cutaneous layer and the extensor tendon, which have intensity values similar to those of the joint capsule. The refinement step used in

the $Ref - R$ and $SM - R$ leads to significant improvements in both methods, mainly because this post-processing step produces smoother contours that were less sensitive to noise and artifacts.

One advantage of the $SM - R$ method over the $UNet$ is that the formulation and parameterization are easier to understand and have direct implications on the results. On the other hand, the $UNet$ is a black box system, which is not always well accepted by some clinical personnel. The disadvantage of the $SM - R$ and $Ref - R$ methods is that the results assumed a perfect scenario since the manual annotations were used. In order to achieve a fully automatic system, the bones and the extensor tendon need to be identified first. Knowing that no perfect automatic segmentation method exists for these structures, it is expected that the results deteriorate once the automatic identifications are added.

In the future, the dataset will be expanded with new examples, with focus on images with pathology. This expansion may improve considerably the results of all methods and specially the $UNet$. These segmentation methods will also be included in the automatic synovitis detection and quantification.

Chapter 8

Synovitis Detection and Grading

Contents

8.1 Introduction	103
8.2 Proposed Work	105
8.2.1 Feature Extraction	106
8.2.2 Feature Selection	108
8.3 Results and Discussion	109
8.4 Conclusions	112

8.1 Introduction

This chapter presents the study for the synovitis detection (presence or absence) and grading in ultrasound images of the metacarpophalangeal joint (MCPJ).

As stated in Chapter 3, synovitis is one of the most important indicators of the presence of Rheumatoid Arthritis (RA). It consists on the inflammation of the synovium and the long-term effects of this condition are the joint degradation and mobility loss. It is classified in 4 grades, from 0 to 3, depending on the severity, where 0 refers to the absence of synovitis (or normal) and 3 is the most severe grade [106]. Fig. 8.1 represents the joint capsule morphology for each synovitis grade in the ultrasound images of the MCPJ dorsal view.

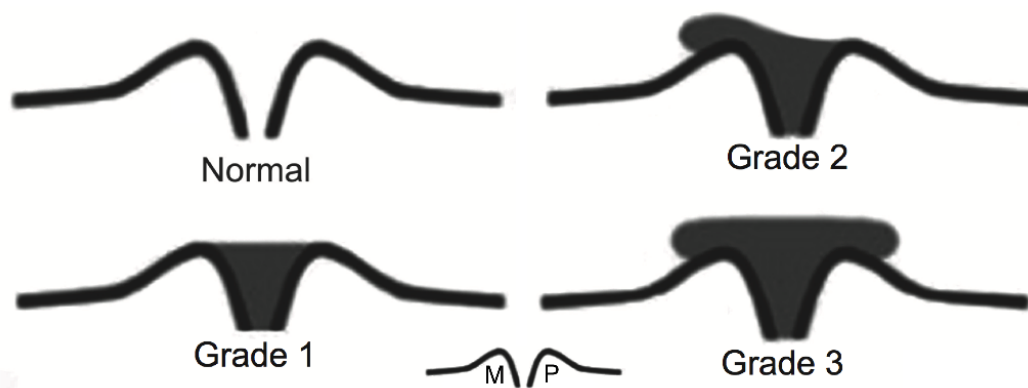


FIGURE 8.1: Representation of the synovitis grades convention. This convention is also used in the metatarsophalangeal joints and interphalangeal joints [107]. In gray is the joint capsule; M - Metacarpus; P - Phalange.

Fig. 8.1 indicates that in a healthy patient, classified with grade 0, the joint capsule is almost imperceptible. When synovitis is present, the normal synovial fluid circulation is compromised. This results in joint capsule swelling due to the accumulation of synovial fluid, which occurs as a darker region in the joint capsule area. As the severity increases the joint capsule grows, first near the joint region, then over the metacarpus, and after that over the phalange, representing respectively grades 1, 2 and 3. In Fig. 8.2, it is possible to visualize some examples of real images with different synovitis grades.

In Fig. 8.2 - A, a normal joint capsule is visible, with normal tissues between the bones extremities. In Fig. 8.2 - B, a small darker zone is visible between the bones extremities and this image is considered as grade 1. In the third image (Fig. 8.2 - C), it is possible to see that the darker zone has reached the metacarpus, making it a grade 2. In the last image (Fig. 8.2 - D), a larger swelling is visible, with the joint capsule extending over the phalange, which is considered as grade 3.

Typically, the distinction between grade 0 and 3 is fairly consensual, however, the differences between successive grades are much harder to determine. This is due to the image quality, artifacts and the inherent limitations of discretely quantifying a progressive condition. This is the main motivation for this chapter, since the automatic quantification of the synovitis grade could be of great help to the rheumatologist by reducing the diagnosis subjectivity.

As previously stated, the automatic quantification of synovitis in ultrasound images of

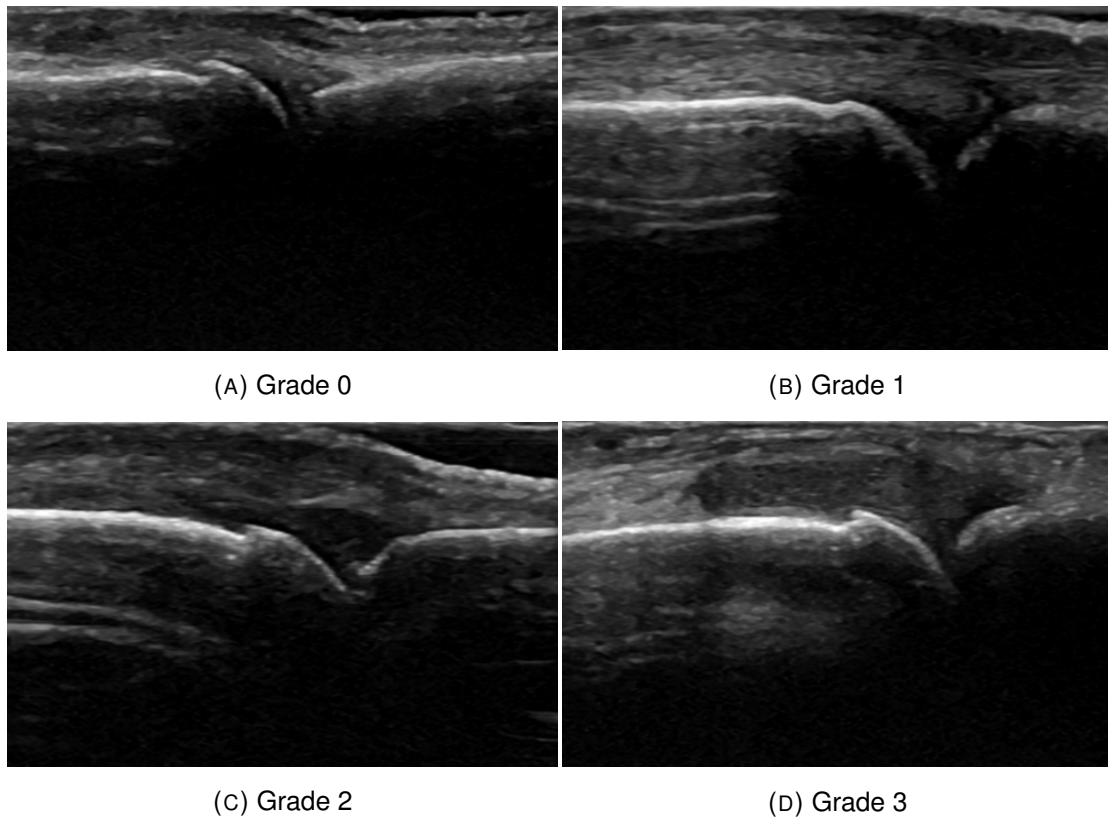


FIGURE 8.2: Example of images with different synovitis grades.

the MCPJ is still an open problem. In [108], a machine learning approach was proposed to identify the joint in ultrasound images of the MCPJ. The authors used a set of convolutional masks to capture local and contextual information and, in that way, find the joint location. Several descriptors were extracted from these masks (such as SIFT, FAST, ORB, BRISK, FREAK and SURF) and tested with different classifiers: Decisions Trees, Neural Networks and SVM. The authors concluded that the best results were obtained using a SVM classifier and SURF descriptors, achieving sensitivities above 98%, for a false positive rate of 10%. In a posterior work, [109], a clinical validation was presented, and the authors concluded that the algorithm has fair agreement with three expert rheumatologists, but it is not ready for clinical usage.

8.2 Proposed Work

Based on the literature review presented in Chapter 4 and the work of [108], the present study proposes the creation of a set of masks to capture image patterns caused by synovitis. These masks are obtained from the bones and extensor tendon locations

(Fig. 8.3), a set of metrics is extracted from each mask and two SVM models are applied to classify the image. One SVM model for the synovitis detection, and the other to infer the synovitis grading.

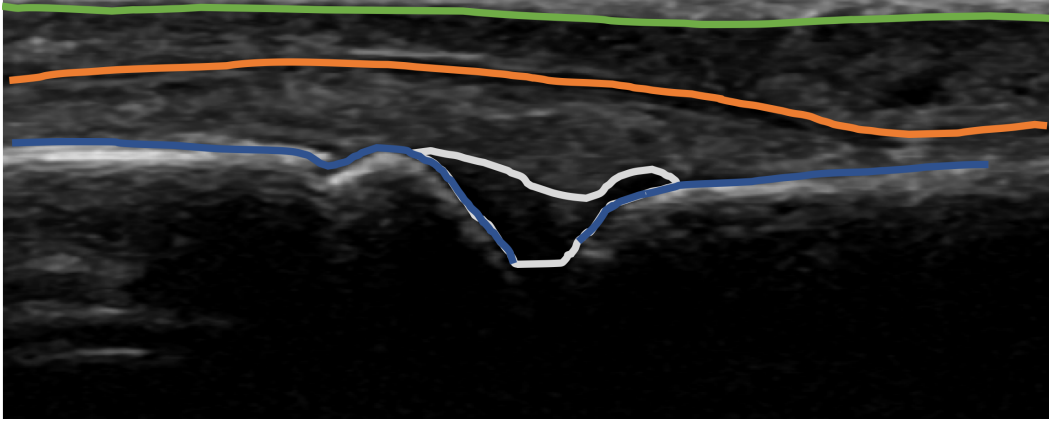


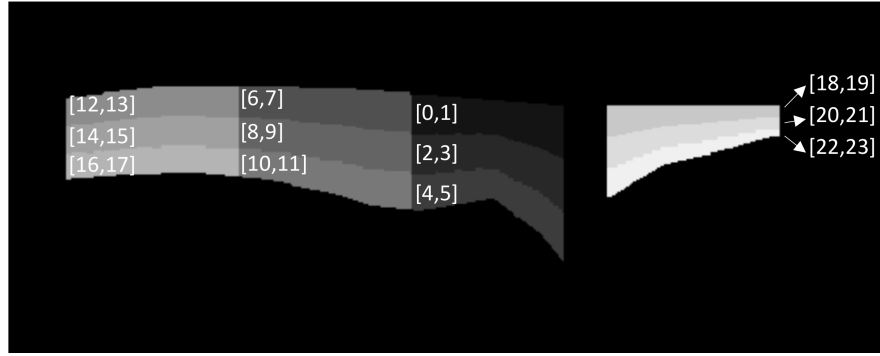
FIGURE 8.3: Manual annotations used in the feature extraction. Blue - metacarpus and phalange; Orange - extensor tendon; Green - skin line; White - joint capsule.

Additionally, a feature selection phase was conducted to extract the most informative features and exclude those that do not provide relevant information. The joint capsule area will also be taken into account to verify if the proposed masks are able to accurately classify the images, without identify the joint capsule. The advantage of this approach is that it might reduce the complexity of the system by removing the joint capsule segmentation, which is a challenging procedure.

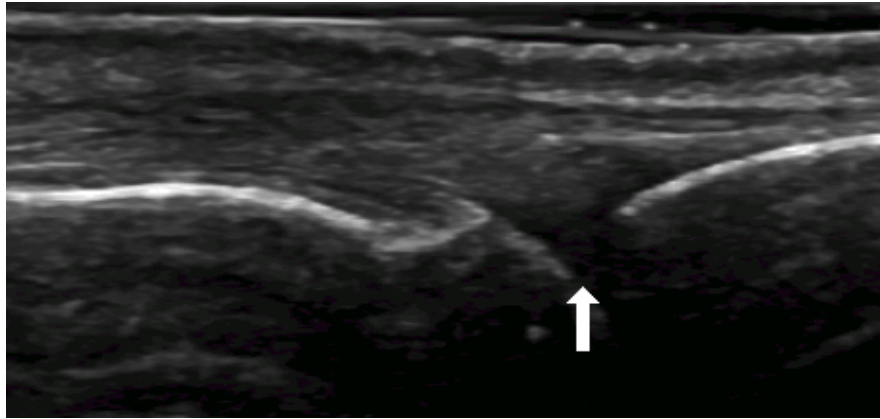
8.2.1 Feature Extraction

As presented in Fig.8.1, the synovitis appears as a darker structure growing from the joint capsule region, which stands between the metacarpus and phalange bones. It was also shown that the synovitis grade is dependent on the size of this dark region and that it follows some specific patterns. Attending to this information, a set of masks is proposed, as depicted in Fig. 8.4.

Starting with the joint masks (Fig. 8.4 - C), they were created having in mind the distinction between image with and without synovitis. They consist on elliptical and concentric regions, centered in P_0 , where the regions belonging to the bone are excluded. Their purpose is to capture the gray-scale signature of the joint region when the image is



(A) Metacarpus and phalange masks.



(B) Input image



(C) Joint masks

FIGURE 8.4: Proposed masks for the synovitis identification and grading. The arrow indicates the point P_0 , which stands between the metacarpus and the phalange. The numbers in each mask are the nomenclature used to facilitate the reference to each feature. The even numbers refer to the means and the odd numbers to the standard deviations of the respective regions.

classified as grade 0 or 1, 2 and 3. Attending to the size and shape of the finger joints, 4 regions with 20 pixels (1 *mm*) thickness were empirically defined.

The masks of the metacarpus and phalange (Fig. 8.4 - A) were defined to identify the synovitis grades 2 and 3, respectively. Vertically, they are limited by the metacarpus,

phalange and extensor tendon and are split into 3 regions with equal thickness. Horizontally, they start at the tip of the respective bone and grow outwards. The metacarpus masks were further divided into three horizontal regions with 100 pixels of width (5 mm), while for the phalange, a single horizontal region was defined with the same 100 pixels width (5 mm). The number of regions in the metacarpus and phalange is different because the morphological changes, caused by synovitis, are more significant above the metacarpus than above the phalange.

From each mask region, the pixels intensity mean and standard deviation was calculated using the input image after a Z-normalization. The Z-normalization parameters were computed for each image from the pixels of the merged regions of the metacarpus and phalange masks. Therefore, the variability between the images acquired with different settings is reduced. The features extracted from each mask region capture the intensity and variation of the gray-scale, which are two of the main features that the ultrasonographers take into consideration when interpreting the images.

In addition to these intensity-based features, it was also included the mean Euclidean distance between the bones and extensor tendon and the joint capsule area. These two metrics were correlated with the synovitis grade. The mean distance can be easily obtained from the same annotations used to create the previously presented masks (bones and extensor tendon annotations). The joint capsule area requires the segmentation of the joint capsule, which in a fully automatic system, would require an additional segmentation step. Therefore, a separated test was created, where the classifier will be tested with and without the inclusion of the joint capsule area.

8.2.2 Feature Selection

A feature selection phase was included in this study. By excluding features that do not have relevant information and selecting the most informative ones, it is possible to create a simpler model with better generalization capabilities. There are several feature selection techniques; in this work, a greedy forward approach was used, due to its simplicity. This strategy starts with an empty model and iteratively adds the feature that increased more significantly the classifier performance. The limitation of this approach is that it does not take into account the relationship between features. An SVM classifier was used and the classification performance was calculated using the mean

Youden's index [110] after a Monte Carlo cross-validation (MCCV) [111]. This metric and procedure were used because the dataset is highly unbalanced. The Youden's index is defined as:

$$J = \text{Sensitivity} + \text{Specificity} - 1 = \frac{TP}{TP + FN} + \frac{TN}{TN + FP} - 1, \quad (8.1)$$

where J is the Youden's index and TP , FP , TN and FN are the number of true positives, false positives, true negatives and false negatives, respectively. This metric merges the sensitivity and the specificity in a single value, thus avoiding biases between classes. This is achieved because the detection performance of each class is calculated independently, i.e., the capacity to detect the positive outcome (sensitivity) and the capacity to detect the negative outcome (specificity). Using this formulation, the mean Youden's index is expected to be low at the beginning, and then increases for several iterations until a point where it starts to decrease due to the inclusion of non-informative features. The point where the Youden's index is maximum will represent the optimal set of features.

8.3 Results and Discussion

The proposed features were implemented and tested on a database with 240 ultrasound images of the MCPJ acquired with a high frequency ultrasound device (12-15 MHz), and saved in anonymized DICOM format. The images have a spatial resolution between $40 \mu m$ and $55 \mu m$ and were cropped to 448×760 (rows, columns) pixels. In all images, the phalange, metacarpus, skin line, extensor tendon, and joint capsule were manually identified as well as the synovitis grade using a self-owned dedicated software as shown in Fig. 8.3. The tests were firstly divided in two distinct configurations, one for synovitis identification (grade 0 vs grade 1, 2 and 3) and other for synovitis grading (grade 1 vs grade 2 vs grade 3). These two configurations were then divided in two, where the effect of adding the joint capsule area was measured. Given the reduced number of images with grade 3, it was decided to merge them with the images of grade 2. Table 8.1 presents the distribution of the images as well as the number of images used in the training and validation phases.

TABLE 8.1: Database organization and number of images used in the Monte Carlo cross-validation train and validation (Train/Validation) attending to the synovitis grade.

Grade	0	1	2	3
Total	183	25	24	8
1 Vs 2+3	-	(17/8)	(17/15)	
0 Vs All	(39/144)		(39/18)	

Since the database is unbalanced, and a balanced training set is preferable, the separation attended to the less represented class (70% to train and 30% to validate). From the other classes, the same number of training examples was randomly extracted for training, and the remaining were used in the validation. The splitting was performed on every MCCV iteration, which was empirically set to 2000 repetitions.

To help understand the results, the numeration presented in Fig. 8.4 was used. The mean bone to extensor tendon distance and the joint capsule area (when added) were also added to the feature vector and were represented with the numbers 32 and 33, respectively.

Table 8.2 describes the results obtained in the previously refereed configurations using the greedy forward MCCV feature selection.

The results in Table 8.2 indicate that the number of features can be reduced using the proposed strategy. Moreover, the results were significantly improved after the feature selection phase when compared with the case where all features were used. In Table 8.3 a summary of Table 8.2 is presented.

Furthermore, the inclusion of the joint capsule area leads to a considerable improvement of the results, from a mean Youden's index of 0.812 to 0.937 in the synovitis grading and from 0.706 to 0.838 in the synovitis detection. Moreover, the use of the joint capsule area also leads to a reduction in the total number of features necessary, from 6 to 5 in the synovitis detection and from 18 to 9 in the synovitis grading. When included, the joint capsule area was always selected first, which highlights its importance to the synovitis identification and grading. The mean distance between the bones and extensor tendon was selected in two cases, the synovitis identification with the joint capsule area and the synovitis grading without the joint capsule area, which does not point to a clear conclusion, reducing the importance of this metric.

TABLE 8.2: Feature selection results in the four tested configurations. μ_J is the mean Youden's index from the Monte Carlo cross-validation and in bold are the best results.
JCA - Joint Capsule Area

Iteration	Synovitis (Y/N)				Synovitis Grading			
	No JCA		With JCA		No JCA		With JCA	
	μ_J	Added	μ_J	Added	μ_J	Added	μ_J	Added
1	0.445	30	0.772	33	0.603	12	0.871	33
2	0.644	17	0.812	31	0.671	13	0.903	6
3	0.689	7	0.825	4	0.697	23	0.907	4
4	0.687	6	0.829	30	0.757	17	0.912	13
5	0.685	31	0.838	32	0.743	27	0.910	21
6	0.706	9	0.824	18	0.744	29	0.910	12
7	0.691	18	0.821	13	0.714	31	0.920	29
8	0.677	12	0.819	23	0.709	0	0.935	9
9	0.667	5	0.816	15	0.795	32	0.937	7
10	0.657	4	0.822	2	0.773	2	0.930	15
11	0.672	20	0.818	5	0.769	28	0.925	16
12	0.673	23	0.820	28	0.784	14	0.908	1
13	0.673	32	0.823	17	0.800	10	0.909	26
14	0.679	11	0.824	16	0.782	8	0.870	23
15	0.685	8	0.819	7	0.780	7	0.867	32
16	0.673	24	0.828	27	0.783	19	0.851	19
17	0.664	10	0.823	29	0.797	16	0.861	27
18	0.675	28	0.819	14	0.812	30	0.842	0
19	0.681	26	0.815	10	0.779	11	0.835	28
20	0.696	19	0.810	21	0.766	24	0.808	18
21	0.687	27	0.810	3	0.757	22	0.808	22
22	0.671	29	0.803	19	0.769	4	0.807	14
23	0.672	2	0.802	8	0.764	15	0.816	10
24	0.662	16	0.790	9	0.736	9	0.804	8
25	0.662	14	0.794	26	0.720	5	0.782	5
26	0.660	22	0.788	22	0.696	1	0.769	17
27	0.660	13	0.788	11	0.658	26	0.781	24
28	0.656	0	0.787	0	0.674	3	0.786	11
29	0.649	15	0.775	6	0.635	25	0.748	2
30	0.643	21	0.768	25	0.620	20	0.732	31
31	0.651	3	0.758	12	0.588	21	0.760	30
32	0.638	25	0.750	24	0.576	6	0.721	20
33	0.607	1	0.742	1	0.568	18	0.726	25
34	-	-	0.729	20	-	-	0.732	3

To help visualize the features selected in each configuration, an example with the masks and the respective selected features is presented in Fig. 8.5.

Attending to Fig. 8.5 - A and B, where the synovitis detection results of the feature selection are presented, it is observable that the phalange features were never used.

TABLE 8.3: Summary of the results obtained before and after the feature selection phase for each configuration. JCA - Joint Capsule Area

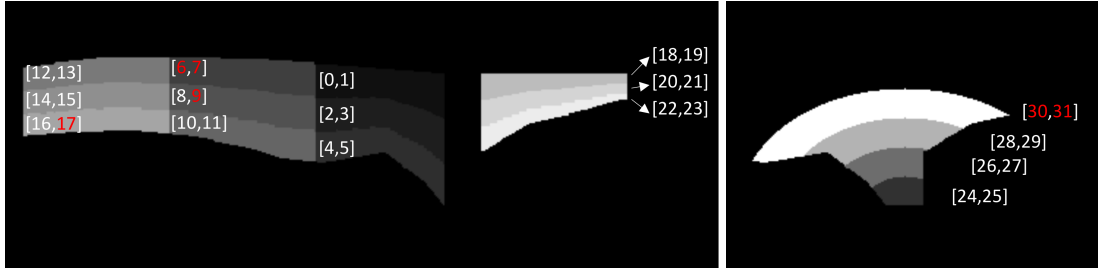
	Synovitis (Y/N)		Synovitis Grading	
	No JCA	With JCA	No JCA	With JCA
All Features	0.607	0.729	0.568	0.732
Selected Features	0.706	0.838	0.812	0.937
# of Features	6	5	18	9
Bone-Tendon Distance	<i>N</i>	<i>Y</i>	<i>Y</i>	<i>N</i>

This result is not surprising, since the difference between the grade 0 and the other grades is mostly in the joint capsule and metacarpus regions, while the phalange is mostly used to distinguish between grade 2 and 3. This fact also justifies the selection of the features in the outer ring of the joint capsule masks and the ones in the metacarpus area. In the synovitis grading, Fig. 8.5 - C and D, features from all regions were used. Once again, the phalange features were not selected as much as those of the metacarpus or joint capsule. The merge of the grades 2 and 3 may have decreased the importance of the phalange, since they were expected to help in that particular distinction.

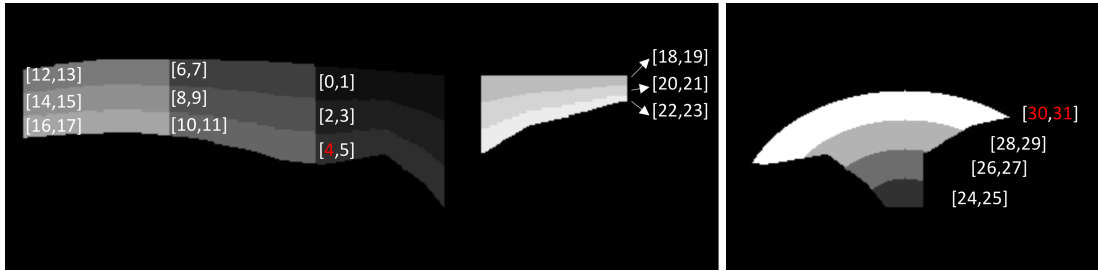
It is also interesting to note that the features extracted from the joint capsule mask, near the bones meeting point (features 24, 25 and 26) were never selected (and the 27 only once). In this region, the gray scale pattern is often darker than the surroundings, due to the metacarpus and phalange cartilage and due to the normal synovial fluid. Even when synovitis is present, this region is expected to maintain its appearance (i.e., dark), which explains why these features are not useful for detecting or grading synovitis.

8.4 Conclusions

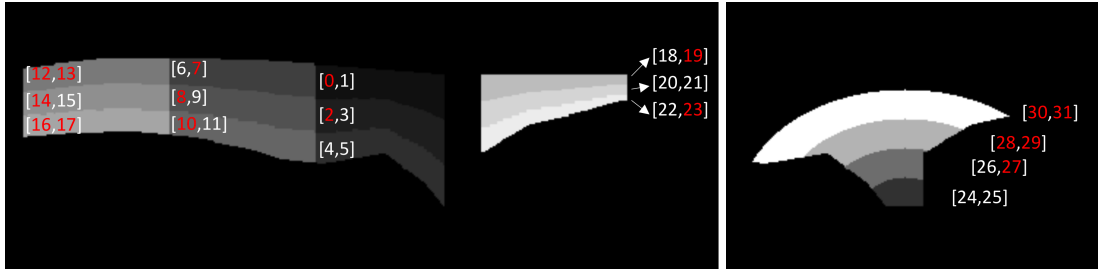
In this chapter, a new set of masks was proposed to extract information from the ultrasound images of the MCPJ, and to classify the joints in terms of synovitis presence and grading. The best configurations achieved a Youden's index of 0.838 and 0.937 for the synovitis identification and grading, respectively. These results are mainly indicative because the database size is small and potentially biased. In fact, this is the main



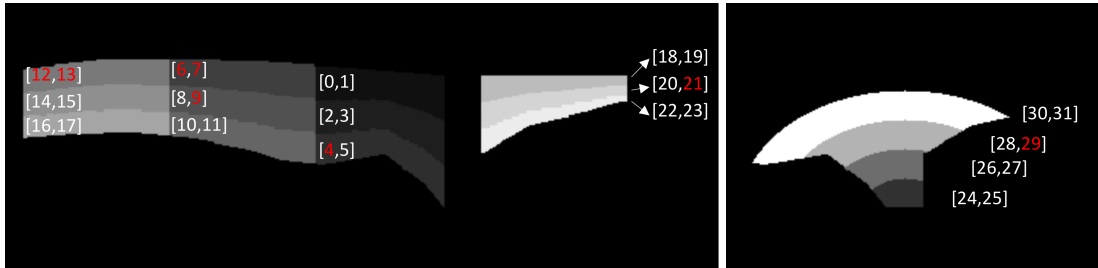
(A) Synovitis (Y/N) test without joint capsule area.



(B) Synovitis (Y/N) test with joint capsule area (Feature 33).



(C) Synovitis grade test without joint capsule area.



(D) Synovitis grade test with joint capsule area (Feature 33).

FIGURE 8.5: Selected features for the different synovitis detection and grading tests. The features included in each test are highlighted in red. In B) and C), the area between the bones and the extensor tendon was also included (Feature 32).

limitation of this approach, especially in the number of images with grade 3. A large-scale database is advised to properly validate these masks and extend the grading to discriminate between grade 2 and 3.

The proposed work could be further improved by studying parameters that were empirically defined when computing the features, i.e., the number of mask and their sizes.

The width of the metacarpus and phalange masks could be obtained automatically for each case by taking into account the anatomical structures. For example, the metacarpus head could be used to limit the first set of regions (features 0 to 5 in Fig. 8.4 - A), which would cover the metacarpus cartilage and would be more precise in the distinction between grade 1 and 2. The same could be done for the joint capsule masks, attending to the size of the joint. Finally, additional features could be extracted for the mask regions, and therefore to capture additional important image characteristics.

In the future, the anatomical structures should be identified automatically, thus providing a fully automated system. The database should be improved by acquiring more images particularly from pathological cases, and results on a test set reported. Follow-up exams should also be included to conduct a longitudinal study to evaluate the diseases progression, and correlated it with the proposed metrics. A continuous grading system, obtained automatically, could also be a useful addition for the ultrasonographers and it is present here as a possible future work.

Chapter 9

Conclusions and Future work

The main objective of this PhD thesis was the proposal of automatic image processing strategies to extract information from the ultrasound images of the metacarpophalangeal joint. Because the state of the art was short in this theme, a focus on the identification of structures was also carried out. This work creates the basis for possible future works and new automatic systems for the detection and quantification of rheumatic diseases using ultrasound imaging. A total of four specific contributions were produced during this work, that contribute to this objective.

The first contribution was the automatic identification of the metacarpus and phalange bones. These structures are of major importance, because they are the most visible ones in the image, providing spatial references for the other structures. Additionally, it is possible to find signs of erosions, osteophytes and subluxation, three important clinical findings. It should be noted that the ultrasound waves do not penetrate bones due to high acoustic impedance and so only noise and artifacts are found beneath them. The identification of both bones used a localizing active contours approach, which achieved a *confidence* of 90% for an *MHD* error below 0.3 *mm*. This method could be further improved by adding shape constraints to the model. This could reduce the burden on the post-processing and create a more robust model, especially in cases where the contours diverge from the bone location due to artifacts or low contrast.

The second contribution was the identification of the extensor tendon. This structure can be used to detect tendon related diseases, such as tenosynovitis. Together with the metacarpus and phalange, it is also useful to create a region of interest around the

joint capsule. The identification was achieved with a new proposed open-ended-active-contours framework with phase symmetry pre-processing and contextual information constrains based on the expected shape and relative position to the bones and skin line. A *confidence* of 95% for an *MHD* below 0.5 *mm* was achieved, proving that the proposed framework is suitable for this task. The phase symmetry was proved to be better than the same formulation without pre-processing. The energy formulation was also added to a fully automatic segmentation pipeline, resulting in a loss of 7% in the overall *confidence* for the same *MHD* of 0.5 *mm* mainly due to the segmentation errors from the previous steps (bone and skin line segmentations). The main drawbacks of this proposal are the dependency on the other segmentation methods to create the search limits and the low scalability due to the exponential complexity growth with the number points to be fit. The main advantages are that the inclusion of new constrains is simple and no precise initialization is needed.

The third contribution was the identification of the joint capsule. This is the main structure of study when addressing rheumatoid arthritis, and so its identification can be used as a direct measure of swelling. Two segmentation methods were proposed for this problem. One uses a split and merge approach, $SM - R$, where the simple linear iterative clustering algorithm was used to split the image in small clusters and then a special region growing with shape restrictions was used to merge them into a coherent shape. The other approach uses the state of the art *UNet* model with data augmentation to directly identify the joint capsule. In the $SM - R$ method, the inclusion of a shape constraint metric helped the region growing process to stop when the joint capsule limits were not visible. The use of the *DICE* coefficient between the object and its own convex hull was proved to enforce the expected balloon shape of the joint capsule. The results obtained with the *UNet* model showed that it outperforms the $SM - R$ method by a considerable margin. It identified 95% of the images with a *DICE* higher than 0.6 in contrast with the $SM - R$ that identified 90% of the image with a *DICE* higher than 0.34. The *UNet* method is also expected to scale better with more data, while the $SM - R$ would require the inclusion of additional constrains or a deeper rework to catch the *UNet* performance.

The last contribution was the identification and quantification of the synovitis grading. For this task, the segmentation of the bones and extensor tendon was used to create a set of masks from where features were extracted (mean and standard deviation).

These masks capture local information from the images, which is then used on two SVM models. One SVM model for detection and the other to classify the synovitis grading. A feature selection phase was conducted and a Youden's index of 0.838 and 0.937 was achieved for the synovitis identification and grading, respectively. The feature selection phase highlighted the importance of specific regions of the image, namely, intensity changes over the metacarpus region for the synovitis identification and over the phalange area for the synovitis grading. Given the results obtained with deep learning techniques in the joint capsule segmentation, it would be interesting to test those approaches for the direct identification and/or quantification of synovitis and other possible image findings.

In the future, it is advised the creation of an improved database with all fingers included and with metatarsophalangeal joint and carpal bones images. The inclusion of other non-imaging parameters such as age, sex and blood tests would also be interesting for a complete diagnosis tool. This database would be helpful to validate and potentially improve the proposed algorithms.

Bibliography

- [1] Claudia Schueller-Weidekamm. Quantification of synovial and erosive changes in rheumatoid arthritis with ultrasound–revisited. *Eur J Radiol*, 71(2):225–31, Aug 2009. doi: 10.1016/j.ejrad.2009.02.008.
- [2] A Marques, JC Branco, JT Costa, LC Miranda, M Almeida, P Reis, RA Santos, V Tavares, A Diniz, and VM Queiroz. Programa nacional contra as doenças reumáticas. *Direcção Geral da Saúde, Lisboa*, pages 1–92, 2005.
- [3] Daniel Aletaha, Tuhina Neogi, Alan J Silman, Julia Funovits, David T Felson, Clifton O Bingham, Neal S Birnbaum, Gerd R Burmester, Vivian P Bykerk, Marc D Cohen, et al. 2010 rheumatoid arthritis classification criteria: an american college of rheumatology/european league against rheumatism collaborative initiative. *Arthritis & Rheumatism*, 62(9):2569–2581, 2010.
- [4] Esperanza Naredo, María Montoro, and Iustina Janta. *Musculoskeletal Ultrasonography in Rheumatic Diseases*, chapter 3 - Rheumatoid Arthritis. Springer, 2015.
- [5] Richard J Wakefield, Wayne W Gibbon, Philip G Conaghan, PHILIP O'Connor, Dennis McGonagle, Colin Pease, Michael J Green, Douglas J Veale, John D Isaacs, and Paul Emery. The value of sonography in the detection of bone erosions in patients with rheumatoid arthritis. *Arthritis Rheum*, 43(12):2762–70, 2000.
- [6] Pravin Patil and Bhaskar Dasgupta. Role of diagnostic ultrasound in the assessment of musculoskeletal diseases. *Therapeutic advances in musculoskeletal disease*, 4(5):341–355, 2012.

- [7] Blausen.com staff. Medical gallery of blausen medical 2014. *WikiJournal of Medicine* 1 (2), 2014.
- [8] N. Martins, S. Sultan, D. Veiga, M. Ferreira, F. Teixeira, and M. Coimbra. A new active contours approach for finger extensor tendon segmentation in ultrasound images using prior knowledge and phase symmetry. *IEEE Journal of Biomedical and Health Informatics*, pages 1–1, 2018. ISSN 2168-2194. doi: 10.1109/JBHI.2017.2723819.
- [9] N. Martins, M. S. Sultan, D. Veiga, M. Ferreira, and M. Coimbra. Segmentation of the metacarpus and phalange in musculoskeletal ultrasound images using local active contours. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4097–4100, Aug 2016. doi: 10.1109/EMBC.2016.7591627.
- [10] N. Martins, M. S. Sultan, D. Veiga, M. Ferreira, and M. Coimbra. Joint capsule segmentation in ultrasound images of the metacarpophalangeal joint using a split and merge approach. In *2018 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, pages 243–246, March 2018. doi: 10.1109/BHI.2018.8333414.
- [11] M. S. Sultan, N. Martins, M. J. Ferreira, and M. T. Coimbra. Segmentation of bones mcp joint region of the hand from ultrasound images. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3001–3004, Aug 2015. doi: 10.1109/EMBC.2015.7319023.
- [12] Malik Saad Sultan, Nelson Martins, Diana Veiga, Manuel Ferreira, and Miguel Coimbra. Automatic segmentation of extensor tendon of the mcp joint in ultrasound images. In *Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies*, BIOSTEC 2016, pages 71–76, Portugal, 2016. SCITEPRESS - Science and Technology Publications, Lda. ISBN 978-989-758-170-0. doi: 10.5220/0005692500710076.
- [13] M. S. Sultan, N. Martins, D. Veiga, M. J. Ferreira, and M. T. Coimbra. Tracking of the anterior mitral leaflet in echocardiographic sequences using active contours. In *2016 38th Annual International Conference of the IEEE Engineering*

- in Medicine and Biology Society (EMBC)*, pages 1074–1077, Aug 2016. doi: 10.1109/EMBC.2016.7590889.
- [14] Malik Saad Sultan, Nelson Martins, Eva Costa, Diana Veiga, Manuel João Ferreira, Sandra da Silva Mattos, and Miguel Tavares Coimbra. Real-time anterior mitral leaflet tracking using morphological operators and active contours. In *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2017) - Volume 2: BIOIMAGING, Porto, Portugal, February 21-23, 2017.*, pages 39–46, 2017. doi: 10.5220/0006244700390046.
- [15] Eva Costa, Nelson Martins, Malik Saad Sultan, Diana Veiga, Manuel João Ferreira, Sandra da Silva Mattos, and Miguel Tavares Coimbra. A preliminary approach for the segmentation of mitral valve regurgitation jet in doppler ecocardiography images. In *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2017) - Volume 2: BIOIMAGING, Porto, Portugal, February 21-23, 2017.*, pages 47–54, 2017. doi: 10.5220/0006248900470054.
- [16] M. S. Sultan, N. Martins, E. Costa, D. Veiga, M. J. Ferreira, S. Mattos, and M. T. Coimbra. Tracking large anterior mitral leaflet displacements by incorporating optical flow in an active contours framework. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3244–3247, July 2017. doi: 10.1109/EMBC.2017.8037548.
- [17] M. S. Sultan, N. C. Martins, E. Costa, D. Veiga, M. J. Ferreira, S. Mattos, and M. T. Coimbra. Virtual m-mode for echocardiography: A new approach for the segmentation of the anterior mitral leaflet. *IEEE Journal of Biomedical and Health Informatics*, pages 1–1, 2018. ISSN 2168-2194. doi: 10.1109/JBHI.2018.2799738.
- [18] Anahi Perlas Vincent Chan. *Atlas of ultrasound-guided procedures in interventional pain management*, chapter Chapter 2 - Basics of Ultrasound Imaging. Springer Science; Business Media, 2010.
- [19] Fabio Martino, Enzo Silvestri, Walter Grassi, and Giacomo Garlaschi. *Musculoskeletal sonography: technique, anatomy, semeiotics and pathological findings in rheumatic diseases*, chapter 1. Springer Science and Business Media, 2007.

- [20] URL http://www3.gehealthcare.com.sg/~media/images/product/product-categories/ultrasound/voluson/voluson{%}20transducers/voluson_transducers_lg.jpg.
- [21] Zhong Tao, Hemant D. Tagare, and J. D. Beaty. Evaluation of four probability distribution models for speckle in clinical cardiac ultrasound images. *IEEE Trans. Med. Imaging*, 25(11):1483–1491, 2006. doi: 10.1109/TMI.2006.881376. URL <http://doi.ieeecomputersociety.org/10.1109/TMI.2006.881376>.
- [22] S Kalaivani Narayanan and RSD Wahidabanu. A view on despeckling in ultrasound imaging. 2009.
- [23] Margeret Callan, editor. *The Rheumatology Handbook*. Imperial College Press, 2012.
- [24] Ade Adebajo. *ABC of Rheumatology*, volume 189, chapter 26. John Wiley & Sons, 2011.
- [25] Charles G Helmick, David T Felson, Reva C Lawrence, Sherine Gabriel, Rosemarie Hirsch, C Kent Kwok, Matthew H Liang, Hilal Maradit Kremers, Maureen D Mayes, Peter A Merkel, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the united states: Part i. *Arthritis & Rheumatism*, 58(1): 15–25, 2008.
- [26] Helena Canhao, Augusto Faustino, Fernando Martins, and Joao Eurico Fonseca. Reuma. pt-the rheumatic diseases portuguese register. *Acta reumatologica portuguesa*, 36(1):45–56, 2010.
- [27] Jaime C Branco, Ana M Rodrigues, Nélia Gouveia, Mónica Eusébio, Sofia Ramiro, Pedro M Machado, Leonor Pereira da Costa, Ana Filipa Mourão, Inês Silva, Pedro Laires, et al. Prevalence of rheumatic and musculoskeletal diseases and their impact on health-related quality of life, physical function and mental health in portugal: results from epireumapt—a national health survey. *RMD open*, 2(1):e000166, 2016.
- [28] Inês Silva, Ana M Rodrigues, Helena Canhão, Nélia Gouveia, Jaime C Branco, Ana Filipa Mourão, Pedro Laires, Sofia Ramiro, Pedro Machado, Filipe Araújo,

- et al. Epireumapt—the study of rheumatic and musculoskeletal diseases in portugal: a detailed view of the methodology. *Acta Reumatológica Portuguesa*, 40(2), 2015.
- [29] Teresa Monjardino, Raquel Lucas, and Henrique Barros. Frequency of rheumatic diseases in portugal: a systematic review. *Acta reumatológica portuguesa*, 36 (4):336–363, 2010.
- [30] O Sangha. Epidemiology of rheumatic diseases. *Rheumatology*, 39(suppl_2): 3–12, 2000.
- [31] Anthony S Fauci et al. *Harrison's principles of internal medicine*, volume 2. McGraw-Hill, Medical Publishing Division New York, 2008.
- [32] Ade Adebajo. *ABC of Rheumatology*, volume 189, chapter 12. John Wiley & Sons, 2011.
- [33] C. Kent Kwok, Larry G. Anderson, Jerry M. Greene, Dorothy A. Johnson, James R. O'Dell, Mark L. Robbins, W. Neal Roberts, Robert W. Simms, and Robert A. Yood. Guidelines for the management of rheumatoid arthritis: 2002 update - american college of rheumatology subcommittee on rheumatoid arthritis guidelines. *Arthritis and Rheumatism; research*, 46(2):328–346, 2002. ISSN 0004-3591. doi: 10.1002/art.10148.
- [34] Walter Grassi and Claudio Cervini. Ultrasonography in rheumatology: an evolving technique. *Annals of the rheumatic diseases*, 57(5):268–271, 1998.
- [35] G Meenagh, E Filippucci, D Kane, A Taggart, and W Grassi. Ultrasonography in rheumatology: developing its potential in clinical practice and research. *Rheumatology (Oxford)*, 46(1):3–5, Jan 2007. doi: 10.1093/rheumatology/kel317.
- [36] Yasser El Miedany. *Musculoskeletal Ultrasonography in Rheumatic Diseases*. Springer, 2015.
- [37] Kenneth G Saag, Gim Gee Teng, Nivedita M Patkar, Jeremy Anuntiyo, Catherine Finney, Jeffrey R Curtis, Harold E Paulus, Amy Mudano, Maria Pisu, Mary Elkins-Melton, et al. American college of rheumatology 2008 recommendations for the use of nonbiologic and biologic disease-modifying antirheumatic drugs in rheumatoid arthritis. *Arthritis Care & Research*, 59(6):762–784, 2008.

- [38] Margeret Callan, editor. *The Rheumatology Handbook*, chapter 3. Imperial College Press, 2012.
- [39] J Alison Noble, Nassir Navab, and H Becher. Ultrasonic image analysis and image-guided interventions. *Interface Focus*, 1(4):673–85, Aug 2011. doi: 10.1098/rsfs.2011.0025.
- [40] Alexander K Scheel, Kay-Geert A Hermann, Elke Kahler, Daniel Pasewaldt, Jacqueline Fritz, Bernd Hamm, Edgar Brunner, Gerhard A Müller, Gerd R Burmester, and Marina Backhaus. A novel ultrasonographic synovitis scoring system suitable for analyzing finger joint inflammation in rheumatoid arthritis. *Arthritis & Rheumatism*, 52(3):733–743, 2005.
- [41] Kei Ikeda, Yohei Seto, Akihiro Narita, Atsushi Kawakami, Yutaka Kawahito, Hiromu Ito, Isao Matsushita, Shigeru Ohno, Keiichiro Nishida, Takeshi Suzuki, et al. Ultrasound assessment of synovial pathologic features in rheumatoid arthritis using comprehensive multiplane images of the second metacarpophalangeal joint: identification of the components that are reliable and influential on the global assessment of the whole joint. *Arthritis & Rheumatology*, 66(3):523–532, 2014.
- [42] Ahmed S Zayat, Karen Ellegaard, Philip G Conaghan, Lene Terslev, Elizabeth MA Hensor, Jane E Freeston, Paul Emery, and Richard J Wakefield. The specificity of ultrasound-detected bone erosions for rheumatoid arthritis. *Annals of the rheumatic diseases*, 74(5):897–903, 2015.
- [43] February 2016. URL https://www.niams.nih.gov/Health_Info/Rheumatic_Disease/default.asp.
- [44] J. Alison Noble and Djamal Boukerroui. Ultrasound image segmentation: a survey. *IEEE Trans. Med. Imaging*, 25(8):987–1010, 2006. doi: 10.1109/TMI.2006.877092. URL <http://doi.ieeecomputersociety.org/10.1109/TMI.2006.877092>.
- [45] Oriol Pujol, Misael Rosales, Petia Radeva, and Eduard Fernández-Nofrerías. Intravascular ultrasound images vessel characterization using adaboost. In *Functional Imaging and Modeling of the Heart, Second International Workshop, FIMH 2003, Lyon, France, June 5-6, 2003 Proceedings*, pages 242–251, 2003. doi: 10.1007/3-540-44883-7_25. URL http://dx.doi.org/10.1007/3-540-44883-7_25.

- [46] D Boukerroui, J A Noble, M C Robini, and M Brady. Enhancement of contrast regions in suboptimal ultrasound images with application to echocardiography. *Ultrasound Med Biol*, 27(12):1583–94, Dec 2001.
- [47] Y Zimmer and S Akselrod. Image segmentation in obstetrics and gynecology. *Ultrasound Med Biol*, 26 Suppl 1:S39–40, May 2000.
- [48] Kalpana Saini, ML Dewal, and Manojkumar Rohit. Ultrasound imaging and image segmentation in the area of ultrasound: a review. *International Journal of advanced science and technology*, 24, 2010.
- [49] José C Seabra, Francesco Ciompi, Oriol Pujol, Josepa Mauri, Petia Radeva, and Joao Sanches. Rayleigh mixture model for plaque characterization in intravascular ultrasound. *IEEE Transactions on Biomedical Engineering*, 58(5):1314–1324, 2011.
- [50] Gregory G. Slabaugh, Gozde B. Unal, Tong Fang, and Michael Wels. Ultrasound-specific segmentation via decorrelation and statistical region-based active contours. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2006), 17-22 June 2006, New York, NY, USA*, pages 45–53. IEEE Computer Society, 2006. ISBN 0-7695-2597-0. doi: 10.1109/CVPR.2006.318. URL <http://dx.doi.org/10.1109/CVPR.2006.318>.
- [51] Gonzalo Vegas-Sanchez-Ferrero, Jose Seabra, Oriol Rodriguez-Leor, Angel Serrano-Vida, Santiago Aja-Fernandez, Cesar Palencia, Marcos Martin-Fernandez, and Joao Sanches. Gamma mixture classifier for plaque detection in intravascular ultrasonic images. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 61(1):44–61, 2014.
- [52] T Loupas, WN McDicken, and PL Allan. An adaptive weighted median filter for speckle suppression in medical ultrasonic images. *Circuits and Systems, IEEE Transactions on*, 36(1):129–135, 1989.
- [53] A Achim, A Bezerianos, and P Tsakalides. Novel bayesian multiscale method for speckle removal in medical ultrasound images. *IEEE Trans Med Imaging*, 20(8): 772–83, Aug 2001. doi: 10.1109/42.938245.

- [54] S Gupta, R C Chauhan, and S C Sexana. Wavelet-based statistical approach for speckle reduction in medical ultrasound images. *Med Biol Eng Comput*, 42(2): 189–92, Mar 2004.
- [55] X Zong, A F Laine, and E A Geiser. Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing. *IEEE Trans Med Imaging*, 17(4):532–40, Aug 1998. doi: 10.1109/42.730398.
- [56] Pedro Silva. Sistema de análise de imagens de ecografia para reumatologia, 2014.
- [57] Yongjian Yu and Scott T Acton. Speckle reducing anisotropic diffusion. *IEEE Trans Image Process*, 11(11):1260–70, 2002. doi: 10.1109/TIP.2002.804276.
- [58] Khaled Z Abd-Elmoniem, Abou-Bakr M Youssef, and Yasser M Kadah. Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion. *IEEE Trans Biomed Eng*, 49(9):997–1014, Sep 2002. doi: 10.1109/TBME.2002.1028423.
- [59] Rohit Kamal Chatterjee and Avijit Kar. Robust speckle reducing anisotropic diffusion. In *Synthetic Aperture Radar, 2012. EUSAR. 9th European Conference on*, pages 800–803. VDE, 2012.
- [60] Adib Keikhosravi and SeyyedHadi Hashemi-Berenjabad. Ultrasound medical image speckle reduction using fourth-order partial differential equation. In *Machine Vision and Image Processing (MVIP), 2011 7th Iranian*, pages 1–5. IEEE, 2011.
- [61] Mariam Wagih Attia, FEZ Abou-Chadi, Hossam El-Din Moustafa, and Nagham Mekky. Classification of ultrasound kidney images using pca and neural networks. *IJACSA) International Journal of Advanced Computer Science and Applications*, 6(4), 2015.
- [62] José Seabra and Joao Sanches. Modeling log-compressed ultrasound images for radio frequency signal recovery. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 426–429. IEEE, 2008.

- [63] José C Seabra and João M Sanches. On estimating de-speckled and speckle components from b-mode ultrasound images. In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, pages 284–287. IEEE, 2010.
- [64] Ilker Hacihaliloglu, Rafeef Abugharbieh, Antony J Hodgson, and Robert N Rohling. Bone surface localization in ultrasound using image phase-based features. *Ultrasound Med Biol*, 35(9):1475–87, Sep 2009. doi: 10.1016/j.ultrasmedbio.2009.04.015.
- [65] Pezhman Foroughi, Emad Bector, Michael J Swartz, Russell H Taylor, and Gabor Fichtinger. P6d-2 ultrasound bone segmentation using dynamic programming. In *Ultrasonics Symposium, 2007. IEEE*, pages 2523–2526. IEEE, 2007.
- [66] Lixin Gong, Sayan D Pathak, David R Haynor, Paul S Cho, and Yongmin Kim. Parametric shape modeling using deformable superellipses for prostate segmentation. *IEEE Trans Med Imaging*, 23(3):340–9, Mar 2004. doi: 10.1109/TMI.2004.824237.
- [67] I Mikić, S Krucinski, and J D Thomas. Segmentation and tracking in echocardiographic sequences: active contours guided by optical flow estimates. *IEEE Trans Med Imaging*, 17(2):274–84, Apr 1998. doi: 10.1109/42.700739.
- [68] Xin Yang, Jiaoying Jin, Mengling Xu, Huihui Wu, Wanji He, Ming Yuchi, and Mingyue Ding. Ultrasound common carotid artery segmentation based on active shape model. *Comput Math Methods Med*, 2013:345968, 2013. doi: 10.1155/2013/345968.
- [69] Ning Lin, Weichuan Yu, and James S Duncan. Combinative multi-scale level set framework for echocardiographic image segmentation. *Med Image Anal*, 7(4): 529–37, Dec 2003.
- [70] Victor S. Lempitsky, Michael Verhoeck, J. Alison Noble, and Andrew Blake. Random forest classification for automatic delineation of myocardium in real-time 3d echocardiography. In *Functional Imaging and Modeling of the Heart, 5th International Conference, FIMH 2009, Nice, France, June 3-5, 2009. Proceedings*, pages 447–456, 2009. doi: 10.1007/978-3-642-01932-6_48. URL http://dx.doi.org/10.1007/978-3-642-01932-6_48.

- [71] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, and Clara I. Sánchez. A survey on deep learning in medical image analysis. 02 2017. URL <https://arxiv.org/abs/1702.05747>.
- [72] David Golan, Yoni Donner, Chris Mansi, Jacob Jaremko, and Manoj Ramachandran. *Fully Automating Graf's Method for DDH Diagnosis Using Deep Convolutional Neural Networks*, pages 130–141. Springer International Publishing, Cham, 2016. ISBN 978-3-319-46976-8. doi: 10.1007/978-3-319-46976-8_14. URL https://doi.org/10.1007/978-3-319-46976-8_14.
- [73] G. Carneiro and J. C. Nascimento. Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11): 2592–2607, Nov 2013. ISSN 0162-8828. doi: 10.1109/TPAMI.2013.96.
- [74] G. Carneiro, J. C. Nascimento, and A. Freitas. The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. *IEEE Transactions on Image Processing*, 21(3):968–982, March 2012. ISSN 1057-7149. doi: 10.1109/TIP.2011.2169273.
- [75] Hao Chen, Yefeng Zheng, Jin-Hyeong Park, Pheng-Ann Heng, and S. Kevin Zhou. *Iterative Multi-domain Regularized Deep Learning for Anatomical Structure Detection and Segmentation from Ultrasound Images*, pages 487–495. Springer International Publishing, Cham, 2016. ISBN 978-3-319-46723-8. doi: 10.1007/978-3-319-46723-8_56. URL https://doi.org/10.1007/978-3-319-46723-8_56.
- [76] Shaima I Jabbar, Charles R Day, Nicholas Heinz, and Edward K Chadwick. Using convolutional neural network for edge detection in musculoskeletal ultrasound images. In *Neural Networks (IJCNN), 2016 International Joint Conference on*, pages 4619–4626. IEEE, 2016.
- [77] Yoo Na Hwang, Ju Hwan Lee, Ga Young Kim, Yuan Yuan Jiang, and Sung Min Kim. Classification of focal liver lesions on ultrasound images by extracting hybrid textural features and using an artificial neural network. *Bio-medical materials and engineering*, 26(s1):S1599–S1611, 2015.

- [78] K Bommanna Raja, M Madheswaran, and K Thyagarajah. Analysis of ultrasound kidney images using content descriptive multiple features for disorder identification and ann based classification. In *Computing: Theory and Applications, 2007. ICCTA'07. International Conference on*, pages 382–388. IEEE, 2007.
- [79] Komal Sharma and Jitendra Virmani. A decision support system for classification of normal and medical renal disease using ultrasound images: a decision support system for medical renal diseases. *International Journal of Ambient Computing and Intelligence (IJACI)*, 8(2):52–69, 2017.
- [80] Wan Mahani Hafizah, Eko Supriyanto, and Jasmy Yunus. Feature extraction of kidney ultrasound images based on intensity histogram and gray level co-occurrence matrix. In *Modelling Symposium (AMS), 2012 Sixth Asia*, pages 115–120. IEEE, 2012.
- [81] Ricardo Ribeiro and João Sanches. Fatty liver characterization and classification by ultrasound. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 354–361. Springer, 2009.
- [82] Robert M Haralick, Karthikeyan Shanmugam, et al. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621, 1973.
- [83] Karim Lekadir, Alfia Galimzianova, Àngels Betriu, Maria del Mar Vila, Laura Igual, Daniel L Rubin, Elvira Fernández, Petia Radeva, and Sandy Napel. A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound. *IEEE journal of biomedical and health informatics*, 21(1):48–55, 2017.
- [84] Maurizio Maravalle, Federica Ricca, Bruno Simeone, and Vincenzo Spinelli. Carpal tunnel syndrome automatic classification: electromyography vs. ultrasound imaging. *TOP*, 23(1):100–123, 2015.
- [85] Nikolas Petteri Tiilikainen. A comparative study of active contour snakes. *Copenhagen University, Denmark*, pages 21–26, 2007.
- [86] Gareth Jones. Genetic and evolutionary algorithms. *Encyclopedia of Computational Chemistry*, 2:1127–1136, 1998.

- [87] Shawn Lankton and Allen Tannenbaum. Localizing region-based active contours. *Image Processing, IEEE Transactions on*, 17(11):2029–2039, 2008.
- [88] Peter Kovesi et al. Symmetry and asymmetry from local phase. In *Tenth Australian joint conference on artificial intelligence*, volume 190. Citeseer, 1997.
- [89] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11):2274–2282, 2012.
- [90] Xiao-Diao Chen, Weiyin Ma, Gang Xu, and Jean-Claude Paul. Computing the hausdorff distance between two b-spline curves. *Computer-Aided Design*, 42(12):1197 – 1206, 2010. ISSN 0010-4485. doi: <https://doi.org/10.1016/j.cad.2010.06.009>. URL <http://www.sciencedirect.com/science/article/pii/S001044851000148X>.
- [91] Marie-Pierre Dubuisson and Anil K Jain. A modified hausdorff distance for object matching. In *Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on*, volume 1, pages 566–568. IEEE, 1994.
- [92] K Radlak, N Radlak, and B Smolka. Automatic detection of bones based on the confidence map for rheumatoid arthritis analysis. In *Computational Vision and Medical Image Processing V: Proceedings of the 5th Eccomas Thematic Conference on Computational Vision and Medical Image Processing (VipIMAGE 2015, Tenerife, Spain, October 19-21, 2015)*, page 215. CRC Press, 2015.
- [93] Nam-Deuk Kim, Larry Booth, Viren Amin, Jaein Lim, and Satish Udpa. Ultrasonic image processing for tendon injury evaluation. In *Time-Frequency and Time-Scale Analysis, 1998. Proceedings of the IEEE-SP International Symposium on*, pages 241–244. IEEE, 1998.
- [94] Bo-I Chuang, Yung-Nien Sun, Tai-Hwa Yang, Fong-Chin Su, Li-Chieh Kuo, and I-Ming Jou. Model-based tendon segmentation from ultrasound images. In *Bioengineering Conference (NEBEC), 2014 40th Annual Northeast*, pages 1–2. IEEE, 2014.

- [95] Rishu Gupta, Irraivan Elamvazuthi, Sarat Chandra Dass, Ibrahima Faye, Pandian Vasant, John George, and Faizatul Izza. Curvelet based automatic segmentation of supraspinatus tendon from ultrasound image: a focused assistive diagnostic method. *Biomedical engineering online*, 13(1):157, 2014.
- [96] David E Goldberg et al. *Genetic algorithms in search optimization and machine learning*, volume 412. Addison-wesley Reading Menlo Park, 1989.
- [97] Karolina Nurzynska and Bogdan Smolka. Automatic finger joint synovitis localization in ultrasound images. In *SPIE Photonics Europe*. International Society for Optics and Photonics, 2016.
- [98] Ivan Selesnick. Total variation denoising (an mm algorithm). *Connexions*, 12 2012.
- [99] Elisa Veronese, Roberto Stramare, Andrea Campion, Bernd Raffeiner, Valeria Beltrame, Elena Scagliori, Alessandro Coran, Luca Ciprian, Ugo Fiocco, and Enrico Grisan. Improved detection of synovial boundaries in ultrasound examination by using a cascade of active-contours. *Med Eng Phys*, 35(2):188–94, Feb 2013. doi: 10.1016/j.medengphy.2012.04.014.
- [100] Karolina Nurzynska and Bogdan Smolka. Segmentation of finger joint synovitis in ultrasound images. In *Communications and Electronics (ICCE), 2016 IEEE Sixth International Conference on*, pages 335–340. IEEE, 2016.
- [101] 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.
- [102] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.
- [103] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015. doi: 10.1007/s11263-015-0816-y.

- [104] 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.
- [105] 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.
- [106] Hilde Berner Hammer, Pernille Bolton-King, Vivi Bakkeheim, Torill Helene Berg, Elisabeth Sundt, Anne Katrine Kongtorp, and Espen A Haavardsholm. Examination of intra and interrater reliability with a new ultrasonographic reference atlas for scoring of synovitis in patients with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, 2011. ISSN 0003-4967. doi: 10.1136/ard.2011.152926. URL <http://ard.bmj.com/content/early/2011/07/21/ard.2011.152926>.
- [107] Carlos Frederico Arend. Ultrasonography in rheumatoid arthritis: what rheumatologists should know. *Revista Brasileira de Reumatologia (English Edition)*, 53 (1):8899–93100, 2013.
- [108] Kamil Wereszczyński, Jakub Segen, Marek Kulbacki, Paweł Mielnik, Marcin Fojcik, and Konrad Wojciechowski. Identifying a joint in medical ultrasound images using trained classifiers. In Leszek J. Chmielewski, Ryszard Kozera, Bok-Suk Shin, and Konrad Wojciechowski, editors, *Computer Vision and Graphics*, pages 626–635, Cham, 2014. Springer International Publishing. ISBN 978-3-319-11331-9.
- [109] Paweł Mielnik, Marcin Fojcik, Jakub Segen, and Marek Kulbacki. A novel method of synovitis stratification in ultrasound using machine learning algorithms: Results from clinical validation of the medusa project. *Ultrasound in medicine & biology*, 44(2):489–494, 2018.
- [110] William J Youden. Index for rating diagnostic tests. *Cancer*, 3(1):32–35, 1950.
- [111] Werner Dubitzky, Martin Granzow, and Daniel P Berrar. *Fundamentals of data mining in genomics and proteomics*. Springer Science & Business Media, 2007.

Appendix A

Database Information and Study

A.1 Database Characteristics

In order to implement and test the proposed algorithms, a new database with images of the second MCPJ was created. The images were acquired with the collaboration of Unidade Local de Saúde do Alto Minho using a high frequency ultrasound equipment (12 or 15 MHz), GE Healthcare LOGIQ-S8, and saved in anonymised DICOM format. With respect to image characteristics, their size was cropped to 448x760 (rows, columns) pixels with 8 bits/pixel and a spatial resolution between 40 μm and 55 μm . There was no special care in the balancing of the gender, age or other person related characteristic. In all images, the phalange, metacarpus, skin line, extensor tendon and joint capsule were manually identified using a self-owned dedicated software. In Fig. [A.1](#) it is possible to visualize the annotations used in this study.

Additionally, all images were graded according to the presence of synovitis attending to the 4 grade synovitis criteria.

In Table [A.1](#) it is possible to see the number of images acquired for each synovitis grade.

TABLE A.1: Final database size according to the synovitis grade.

Grade	0	1	2	3
Total	183	25	24	8

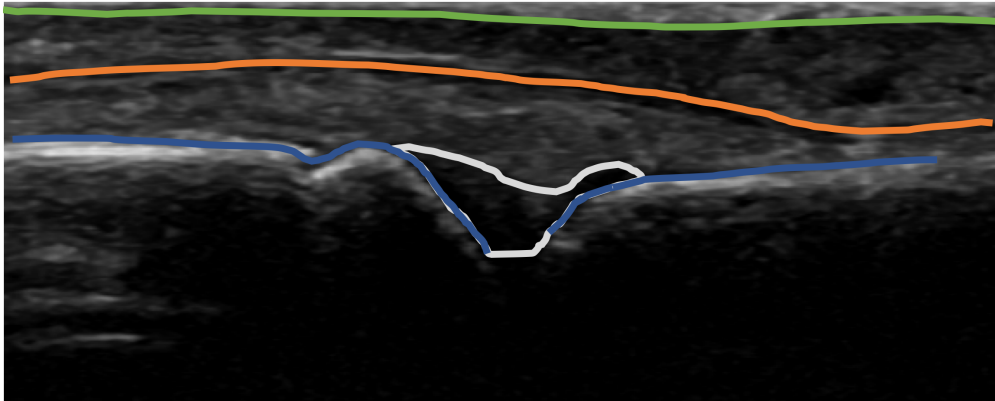


FIGURE A.1: Manual annotations used to validate the proposed algorithms. Blue - metacarpus and phalange; Orange - extensor tendon; Green - skin line; White - joint capsule.

A.2 Clinical Visual Features

In this section, a preliminary study of the images characteristics and annotations is presented in the context of synovitis quantification and detection. Since the state of the art is limited in this field, it was decided to correlate the images changes with the presence of synovitis. The anatomical implications of the synovitis were addressed in Chapter 3 and here, the aim is to confirm the theoretical background using image processing analysis. For that, the following set of features will be study:

- Mean distance between bones and extensor tendon;
- Mean distance between extensor tendon and skin line;
- Joint capsule area;
- Mean intensity inside and outside the joint capsule;
- Intensity standard deviation inside and outside the joint capsule.

These metrics were extracted from the database previously described in Section A.1.

A.2.1 Distances Between Bones, Extensor Tendon and Skin Line

The first two tests were performed using only the manual annotation. The mean vertical Euclidean distance between the bones and the extensor tendon, and between the extensor tendon and the skin line were computed.

These features are based on the idea that the joint capsule swelling should move the extensor tendon away from the bones. Additionally, since the swelling occurs in the upward direction, it is possible that the distance between the tendon and the skin line is decreased due to tissue compression. In Fig. A.2 and A.3, it is possible to visualize the results of these two metrics.

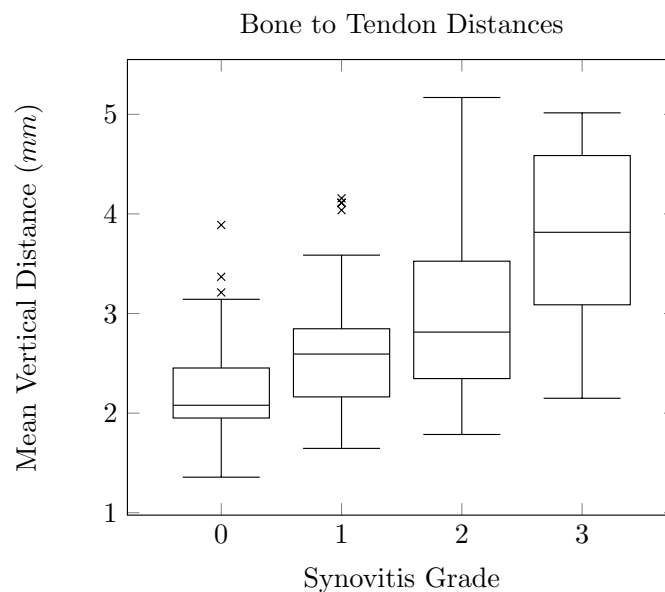


FIGURE A.2: Box plot of the mean distances between the bones and the extensor tendon.

From Fig. A.2, it is possible to verify that the mean distance between the bones and the extensor tendon carries relevant information about the progression of synovitis, because the distance increases as the synovitis grade increases. The mean distance between the extensor tendon and the skin line do not to provide useful information, since the distribution of the distances exhibit a considerable overlap in all grades, Fig. A.3. The Welch's T-test was used to verify if these differences were statistically significant and the results are presented in Table A.2.

Table A.2 corroborates the observations presented in Fig. A.2 and A.3. The mean distance between the bones and the extensor tendon is a discriminative measurement

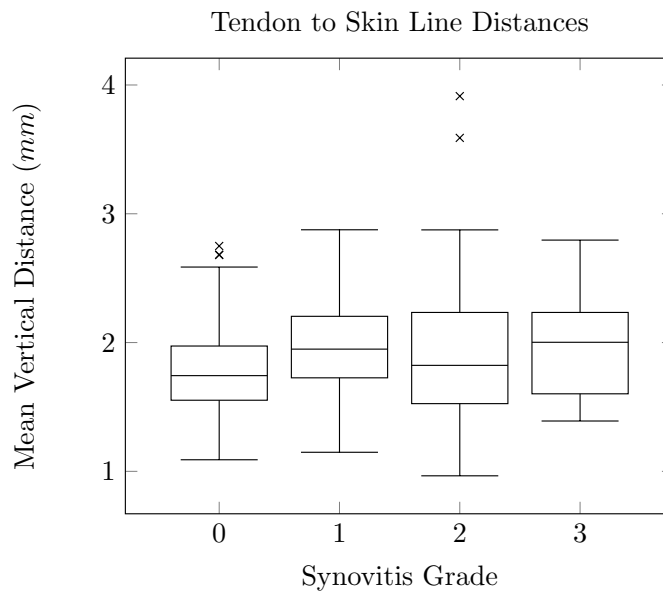


FIGURE A.3: Box plot of the mean distances between the extensor tendon and the skin line.

TABLE A.2: Statistical significance (p -values) between the mean distance measurement and the synovitis degree.

Grades	Bones-Tendon	Tendon-Skin Line
0 Vs 1	$1.61E^{-5}$	0.0002
1 Vs 2	0.0260	0.7584
2 Vs 3	0.0042	0.7440

in all synovitis grades. In turn, the mean distance between the extensor tendon and the skin line only has significant differences in the distinction between grade 0 and 1. This result indicates that the swelling does not cause a significant compression of tissues above the extensor tendon. Moreover, the mean distance between the extensor tendon and the skin line from grade 0 to grade 1 increased instead of decreasing, which contradicts the initial idea of tissue compression.

A.2.2 Joint Capsule Area

The third feature extracted from the manual annotations was the joint capsule area. In Fig. A.4, it is possible to see the results for each synovitis grade.

From Fig. A.4, it is clearly visible that the joint capsule area increases with the increase in the synovitis grade. This result is in accordance with the theoretical background,

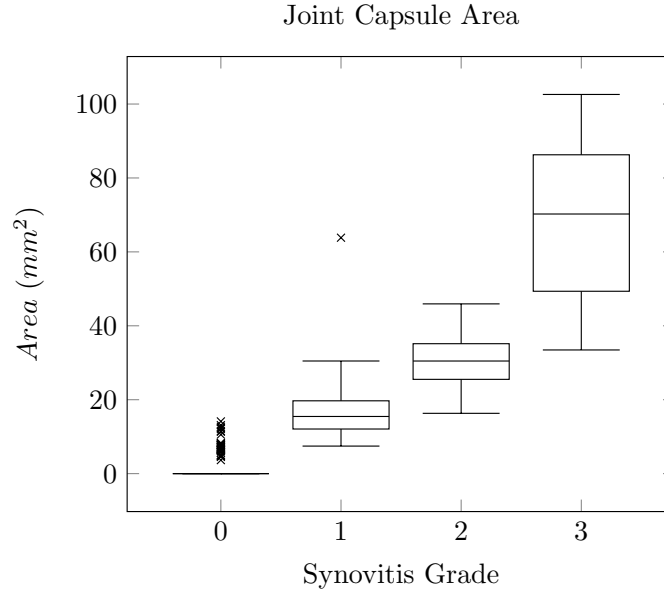


FIGURE A.4: Box plot of the joint capsule area extracted from the manual annotations for each synovitis grade.

because the synovitis progression is characterized by the swelling of the joint capsule. Nevertheless, this difference alone is not enough to classify the MCPJ images with high confidence. There is still a considerable overlap between the measured areas of each grade and, consequently, additional features are needed to accurately classify the images. In the grade 0, there is a big number of outliers, because most of the images have no visible joint capsule (area equal to zero).

The p-values of the Welch's T-test are presented in Table A.3, which confirm the observations made in Fig. A.4: the differences between the joint capsule areas of successive synovitis grades are statistical significant (p-values < 0.05).

TABLE A.3: Statistical significance of the difference in the joint capsule area in different synovitis grades.

Grades	p -value
0 V_s 1	$1.63E^{-18}$
1 V_s 2	$5.99E^{-13}$
2 V_s 3	$4.63E^{-5}$

The joint capsule area and the mean distance between the bones and the extensor tendon are influenced by the same anatomical alteration, i.e., the joint capsule swelling. To verify that these two features provide different types of information, it was computed

the correlation between the two for each synovitis grade. Table A.4 presents those correlation values.

TABLE A.4: Correlation between the joint capsule area and the mean distance between the bones and the extensor tendon for each synovitis grade.

Grade	Correlation
0	0.062
1	0.428
2	0.378
3	0.885

From Table A.4, it is possible to conclude that for the grade 0 the correlation is almost none. This is due to the fact that the mean distance is still possible to calculate, while the joint capsule area is not possible to measure for most grade 0 cases, which results in an area of 0. In the pathological cases there is some correlation, however only the most severe cases (grade 3) exhibits a high correlations. Given these results, it is concluded that these features have some similarity, but they provide different types of information.

A.2.3 Intensity Based

The last two features are the mean and standard deviations of the pixels inside and outside the joint capsule region. The joint capsule swelling is typically darker than the surrounding structures, because it is composed of synovial fluid. Here, the joint capsule region is defined by the manual joint capsule identification and the outside is defined as the region between the tendon and the bones without this joint capsule region. This way, the darker region of the extensor tendon and the brighter regions of the bones are not included. The proposed measurements are depicted in Fig. A.5 and Fig. A.6.

From Fig. A.5, it is possible to verify that the mean intensity is always higher outside the joint capsule region when compared with the inside. Regarding the standard deviation, it is also possible to verify that in most cases it is higher outside the joint capsule, Fig. A.6. These observations are in agreement with the theoretical background.

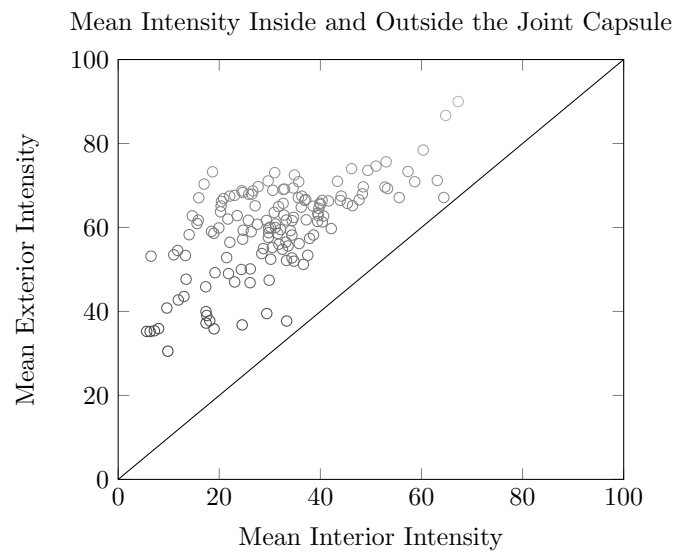


FIGURE A.5: Plot of the mean intensity extracted from inside and outside the joint capsule manual annotations.

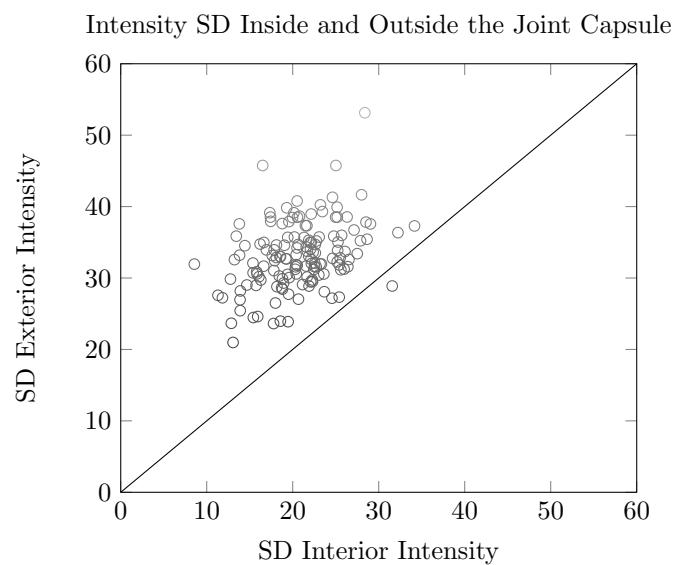


FIGURE A.6: Plot of the intensity standard deviation extracted from inside and outside the joint capsule manual annotations.

