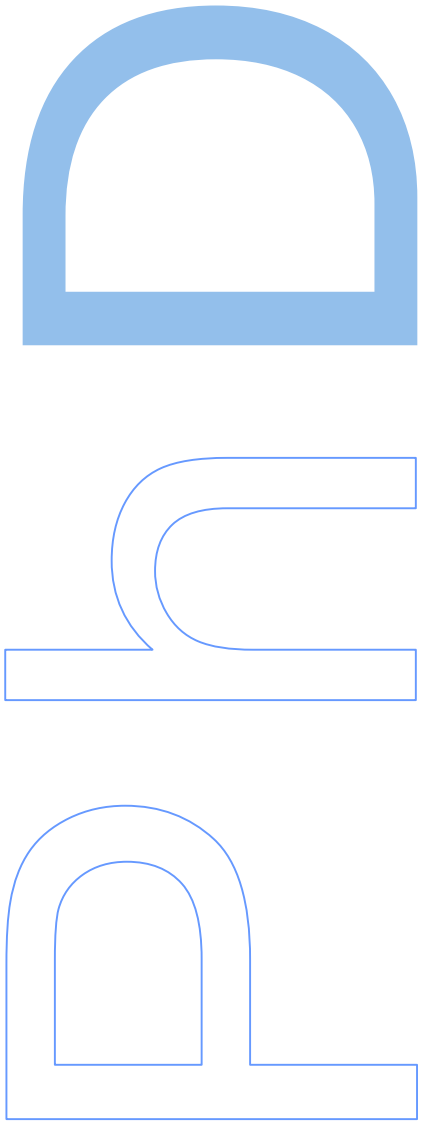
Computer Vision for Blood Vessel Segmentation and Related Applications

Ricardo Jorge Terroso de Araújo
Programa Doutoral em Informática das Universidades do Minho, Aveiro e Porto (MAPi)
Departamento de Ciência de Computadores
2021

Orientador
Hélder P. Oliveira, Professor Auxiliar Convidado, Faculdade de Ciências

Coorientador
Jaime S. Cardoso, Professor Catedrático, Faculdade de Engenharia



Resumo

A segmentação e a análise de vasos sanguíneos em imagens médicas têm vindo a permitir que os clínicos consigam avaliar inúmeras condições médicas que têm impacto direto nestas estruturas (tais como aneurismas e estenoses) e também obter informação sobre algumas doenças sistémicas (por exemplo, diabetes e hipertensão). Infelizmente, o estudo de redes vasculares é um procedimento moroso e repetitivo. A visão computacional tem vindo a desempenhar um papel importante no apoio à segmentação e análise de vasos sanguíneos, aumentando a eficiência dos cuidados de saúde. Apesar do enorme progresso que tem sido alcançado desde os primeiros trabalhos, a segmentação de vasos sanguíneos está ainda aquém do raciocínio utilizado pelos especialistas humanos e apresenta ainda algumas limitações. Esta tese descreve os desafios existentes e como a investigação realizada se propõe a lidar com os mesmos.

Uma grande porção da literatura relacionada com este tópico inclui algoritmos que focam um cenário específico de segmentação de vasos sanguíneos (uma certa combinação de rede vascular e protocolo de aquisição de imagem). Ter um caso de uso em mente permite utilizar mais conhecimento prévio e desenvolver metodologias mais otimizadas. Dois algoritmos para aplicações específicas são propostos. O primeiro consiste numa metodologia desenvolvida para extrair e analisar as perfurantes das artérias epigástricas inferiores profundas, vasos que se encontram na região anterior da parede abdominal. A sua identificação e caracterização são essenciais para o planeamento pré-cirúrgico dos retalhos estado-da-arte que usam tecido da barriga para reconstruir a mama. De forma a garantir que a nova mama é irrigada de forma adequada após a re-anastomose, os cirurgiões precisam de incluir perfurantes com características apropriadas. O segundo caso de uso visa a segmentação dos vasos sanguíneos em imagens da retina, atingindo um desempenho competitivo com os métodos estado-da-arte mas utilizando uma rede neuronal profunda mais eficiente, o que é relevante em programas de triagem, onde uma grande quantidade de dados é gerada num pequeno espaço de tempo.

O desenvolvimento de algoritmos com bom desempenho, e que simultaneamente generalizam apropriadamente para dados com uma distribuição diferente (por exemplo, outros vasos sanguíneos, técnicas de imagiologia diferentes), é um dos desafios-chave atuais na área de segmentação de vasos sanguíneos, e de facto, na visão computacional de forma geral. Isto é especialmente verdade quando se recorre a técnicas de aprendizagem máquina supervisionada, que são normalmente as que atingem melhor desempenho mas tendem a funcionar pobremente em dados com diferente distribuição. De forma a lidar com este problema, a versão profunda de uma metodologia que é conhecida pela sua intuição e boa generalização, a análise dos valores próprios da matriz Hessiana em diferentes escalas, é considerada. As experiências mostram que substituir as funções baseadas em conhecimento prévio por uma rede neuronal profunda permite melhorar o desempenho desta metodologia tradicional e generalizar melhor do que as metodologias típicas de aprendizagem profunda.

As redes vasculares podem ser interpretadas como um grafo, uma vez que, exceto o calibre local, elas podem ser aproximadas por um conjunto de arestas (segmentos de vasos sanguí-

neos) e vértices (bifurcações e terminações). Manter a sua estrutura de grafo correta durante a segmentação é relevante sempre que uma etapa de caracterização se segue, especialmente se for automatizada. No entanto, esta propriedade é tipicamente ignorada na literatura, e a falta de métricas adequadas para a sua avaliação é provavelmente um fator. Um índice de similaridade que é capaz de quantificar de forma objetiva o quão similar são duas redes vasculares do ponto de vista topológico é proposto. Além disso, investigação sobre técnicas que apoiam a aprendizagem de modelos neuronais profundos mais robustos topologicamente é também contemplada. Numa primeira abordagem, uma rede neuronal de duas etapas consistindo numa rede de segmentação típica seguida de um modelo probabilístico de refinamento é utilizada. Uma segunda metodologia considera uma nova função de custo, baseada no operador morfológico de fecho, de forma a penalizar erros que induzem árvores quebradas e aqueles que conectam árvores distintas.

Finalmente, as capacidades dos modelos de contornos ativos estado-da-arte são estendidas, com o objetivo de permitir que possuam um conjunto finito de regiões com diferentes propriedades de flexão. Esta rigidez heterogénea ao longo do contorno permite modelar algumas dinâmicas do contorno do objeto e evitar outras, o que é útil quando algumas das dinâmicas se devem a ruído ou outros artefatos. Um caso de uso que beneficia deste modelo de contornos ativos é a segmentação do pulmão, uma vez que se torna possível modelar de forma precisa o seu contorno medial e, simultaneamente, incluir os nódulos existentes na sua região periférica lateral. Um maior número de nódulos poderá assim ser identificado e caracterizado. Esta última etapa pode envolver a análise dos vasos sanguíneos na vizinhança, uma vez que as suas propriedades poderão fornecer informações relevantes quanto ao tipo de nódulo.

Abstract

The segmentation and the analysis of blood vessels in medical images have been allowing clinicians to assess several health conditions that directly impact these structures (such as aneurysms and stenosis) and also to obtain clues on some systemic diseases (for example, diabetes and hypertension). Unfortunately, the study of vascular networks is typically a time consuming and repetitive procedure. Computer vision has been playing an important role in supporting the segmentation and analysis of blood vessels, increasing the efficiency of medical care. Despite the huge progress that has been achieved since the first works, blood vessel segmentation is still far from matching the rationale used by human experts and is still lacking in some aspects. This thesis describes the existing challenges and how our conducted research proposes to address them.

A large portion of the related literature comprises algorithms focusing a given blood vessel segmentation scenario (a combination of a particular vascular tree and imaging acquisition protocol). Targeting a specific use case allows to more heavily rely on prior knowledge and design more optimised methodologies. Two application-specific algorithms are proposed. The first one concerns a framework tailored for the extraction and analysis of the Deep Inferior Epigastric Perforators, blood vessels lying in the anterior region of the abdominal wall. Their identification and characterisation are essential for the preoperative planning of the state-of-the-art flaps using tissue from the belly to reconstruct the breast. In order to guarantee that the new breast is adequately vascularised after re-anastomosis, the surgeons need to include blood vessels with proper characteristics. The second one targets the segmentation of blood vessels in retinal images, where competitive performance is achieved using a more efficient deep neural network design, which is relevant for screening programs, where large amounts of data are generated in a short amount of time.

The development of algorithms which perform well, and simultaneously generalise properly to differently distributed data (for example, other blood vessels or different image acquisition procedures), is one of the current key challenges of blood vessel segmentation, and in fact, computer vision overall. This is especially true when using supervised machine learning approaches, which are typically the best performing ones but tend to work poorly in differently distributed data. To tackle this issue, a deep version of a framework which is known for being intuitive and generalising well, the eigenvalue analysis of the Hessian matrix at a scale space, is considered. The experiments show how replacing the state-of-the-art prior knowledge based vesselness functions by a deep neural network allows to improve the performance of this traditional framework to the level of deep learning, while generalising better than typical deep learning methodologies.

Blood vessel networks can be interpreted as graph-like structures, since, apart from local calibre, they can be approximated by a set of edges (blood vessel segments) and vertices (bifurcations and vessel endings). Correctly capturing their graph structure during segmentation is relevant whenever a characterisation step follows, especially if it is an automated one. Nonetheless, this property is often overlooked in the literature, and the lack of proper metrics for this evaluation is probably one factor. A similarity index which can objectively quantify the similarity between

the topological structure of two blood vessel trees is proposed. Moreover, research on techniques which can help promoting the learning of deep neural networks which are more robust topological-wise is also contemplated. In a first approach, a two-stage neural network comprising a typical segmentation network followed by a probabilistic refinement model is considered. A second framework includes a novel loss function based on the morphological closing operator in order to increase the weight of errors inducing disjoint trees and of those merging distinct sub-trees.

Finally, the capabilities of state-of-the-art parametric active contour models are extended, with the goal of allowing them to have a finite set of contiguous regions displaying different bending properties. This heterogeneous rigidity along the contour allows to fit some dynamics of the boundaries of the object while disregarding others, which is useful whenever some of the dynamics are due to noise or other artefacts. One use case benefiting from this novel active contour model is the segmentation of the lung, since it enables the simultaneous precise fitting of the medial lung boundary while including in the segmentation the juxta-pleural nodules that occasionally exist in the lateral lung boundary. A larger number of nodules will then be candidate for the identification and characterisation steps. The latter may involve the assessment of the surrounding blood vessels, as their properties may encode relevant information regarding the type of nodule.

Agradecimentos

Quero começar por agradecer aos meus orientadores, Hélder Oliveira e Jaime Cardoso, por além de serem excelentes mentores, terem sido um ombro amigo nos momentos mais desafiantes e difíceis. Obrigado por toda a vossa disponibilidade e uma grande parte do meu crescimento deve-se sem dúvida a vocês. Foi a vossa recepção que me permitiu também ser recebido pela família VCMI, onde além de pessoas geniais, tive a oportunidade de fazer verdadeiros amigos que conto levar para a vida. Obrigado a todos vocês por fazerem do CTM uma segunda casa para mim, onde pudemos partilhar todo o tipo de momentos, desde conversas filosóficas até às discussões mais hilariantes. Quero ainda agradecer a toda a estrutura do CTM e do INESC TEC, por terem sempre possibilitado que pudesse realizar o meu trabalho de forma adequada.

Um agradecimento especial à Unidade da Mama da Fundação Champalimaud pela colaboração de longa data, e que me possibilitou explorar uma área muito desafiante e interessante. É longo o caminho que percorremos até hoje e espero que possamos atingir mais objetivos no futuro.

Ao João, pela amizade de longa data, por teres lá estado em praticamente todos os momentos da minha vida. Já é difícil imaginar quem seria eu sem a tua presença. Ao Bruno, Daniel e Tiago por serem também os meus pilares, por todos os momentos de descontração que tornam a vida mais leve.

Aos meus avós pelo amor incondicional, por me fazerem sentir que o tempo é breve demais e que viver o dia de hoje já é tarde de mais. Espero um dia poder contar aos meus netos as nossas histórias tal como vocês me contam as dos vossos avós. Aos meus tios e aos meus primos, por todos os exemplos que me passaram, por combinarem a experiência de vida e a vontade de viver o momento. Ao meu primo João, quero deixar um agradecimento especial, por ser para mim como um irmão mais velho, um exemplo e alguém com quem sempre anseio estar.

À minha mãe por todos os esforços que sempre fez por mim, que sempre me faz querer dar a melhor versão de mim para a deixar orgulhosa. Sei que nunca te conseguirei agradecer o suficiente. A ti pai, podes não estar cá fisicamente, mas a tua presença em mim é omnipresente e continuas a influenciar-me todos os dias. Não é só quem cá está que faz falta, mas também aqueles que queriam estar e não podem. Um obrigado também ao Nuno por ser um bom amigo e trazer alegria ao meu quotidiano.

Por fim, Cláudia, não há palavras que possam expressar tudo o que significas para mim. Sem ti, esta jornada seria sem dúvida mais difícil. Obrigado por estares sempre lá para mim, por toda a paciência que tiveste que ter comigo em diversos momentos, e por conseguires cativar um sorriso mesmo nos momentos mais difíceis. Agora que esta jornada começa a aproximar-se do fim, prometo que tentarei retribuir tudo o que fizeste por mim.

A todos vocês, obrigado por fazerem de mim quem sou hoje.

Ricardo J. Araújo

"Tudo vale a pena,
Se a alma não é pequena."

Fernando Pessoa

Contents

List of Figures

List of Tables

Acronyms

1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
ACM	Active Contour Model
AUC	Area Under the receiver operating characteristic Curve
B-COSFIRE	Bar-Combination Of Shifted FIlter REsponses
BCE	Binary Cross-Entropy
CAD	Computer Aided Diagnosis
CNN	Convolutional Neural Network
CRF	Conditional Random Field
CT	Computed Tomography
CTA	Computed Tomography Angiography
DIEA	Deep Inferior Epigastric Artery
DIEP	Deep Inferior Epigastric Perforator
DL	Deep Learning
DNN	Deep Neural Network
DRLSE	Distance Regularized Level Set Evolution
DVAE	Denoising Variational Auto-Encoder
FC	Fully Connected
FCM	Fuzzy C-Means
FCN	Fully Convolutional Network
FL	Focal Loss
GMM	Gaussian Mixture Model
GVF	Gradient Vector Flow
kNN	k-Nearest Neighbours
LSACM	Locally Statistical Active Contour Model
ML	Machine Learning

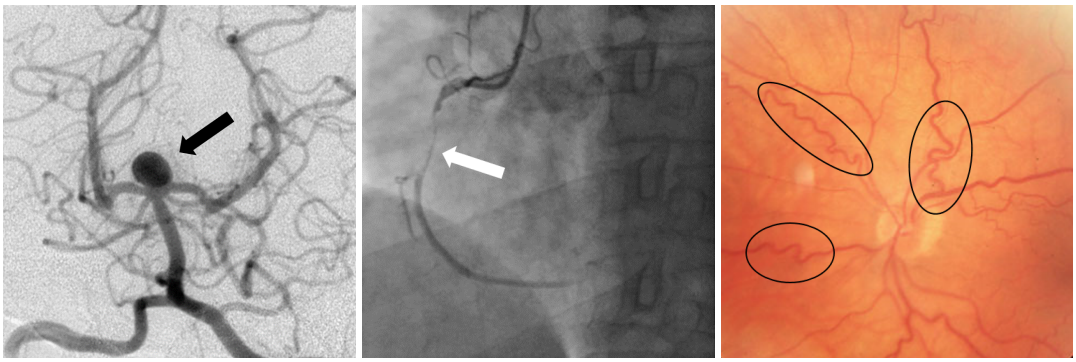
OCTA	Optical Coherence Tomography Angiography
ReLU	Rectified Linear Unit
ROC	Receiver Operating Characteristic
SBGFRLS	Selective Binary and Gaussian Filtering Regularized Level Set
SE	Structuring Element
SMB	Sparse Multi-Bending
SNR	Signal-to-Noise Ratio
SVM	Support Vector Machine
tACM	traditional parametric Active Countour Model
VAE	Variational Auto-Encoder

Chapter 1

Introduction

Blood vessels are the structures responsible for carrying blood throughout the different tissues of the body. They form such an intricate network that, considering the average human adult, lining them up would cover a distance of almost 100,000 km [1]. Arteries, the blood vessels taking blood away from the heart, eventually ramify into arterioles which in turn will occasionally give origin to very narrow capillaries, micro-vessels whose wall is composed by a single layer of endothelial cells. Here, substance exchange occurs with the surrounding interstitial fluid, which mediates exchanges with the neighbouring cells. Oxygen and glucose are some of the molecules leaving the lumen of the capillaries while carbon dioxide and lactic acid are some of the waste products of cellular metabolism that are collected. The blood then flows through venules and returns to the heart via veins. There are two different circulatory circuits, the pulmonary and the systemic circuits. The first is responsible for taking de-oxygenated blood to the pulmonary alveoli, where oxygen perfuses red blood cells again. The latter takes oxygenated blood to the different tissues of the body, being crucial for haemostasis.

An adequate blood flow is only possible when the properties of the vessel walls remain normal. Any stiffening, narrowing or enlargement may lead to serious medical conditions. Local distensions of the blood vessel wall, usually in the form of an outward bulge, are known as aneurysms (Figure 1.1a) and occur due to the weakening of the wall. The causes are generally uncertain, but most likely result from the combination of genetic factors and risk behaviours such as smoking and high blood pressure [2]. Most aneurysms are asymptomatic and do not pose a threat, however, those that keep increasing are prone to rupture, leading to fatal haemorrhages in most cases. They occur mostly in arteries, especially the aorta and the ones in the brain. Vascular stenosis (Figure 1.1b) is the abnormal narrowing of a vessel, usually caused by the progression of atherosclerosis, a process where lipid plaque is accumulated on the wall. This leads to the reduction of blood flow along the vascular network, and severe cases may cause ischaemia, when insufficient blood flow reaches the tissues. In addition, in case of plaque rupture, the formation of a blood clot, also known as thrombus, will be induced and may completely block the bloodstream, possibly causing a heart attack or a stroke, according to the location where the event occurs [3; 4]. If the thrombus detaches from the vessel wall and enters the bloodstream, it gets the designation of embolus,



(a) (b) (c)

Figure 1.1: Blood vessel imaging and examples of different medical conditions: (a) angiography of the brain vasculature showing an aneurysm (black arrow); (b) angiography of the coronary arteries showing a vascular stenosis (white arrow); and (c) fundus photo displaying retinal venules with abnormal excessive tortuosity (highlighted regions).

and may clog a different blood vessel site. One of the medical conditions that arises like this is pulmonary embolism, where a pulmonary artery is obstructed by an embolus that commonly has origin in the femoral veins [5]. The build-up of plaque is also not easily predictable, nonetheless there is a correlation with family history and risk factors such as bad cholesterol levels, smoking, and diabetes, to name a few [6]. Cardiovascular diseases are the major cause of death worldwide. In 2017, from the 56 million deaths registered, almost 18 million were due to cardiovascular compromise, followed by 9.5 million deaths due to cancer [7]. Ischaemic heart disease and stroke are the most problematic, nevertheless, according to the World Health Organization, 80% of the cases could be prevented [8], highlighting the relevance of proper diagnosis and follow up.

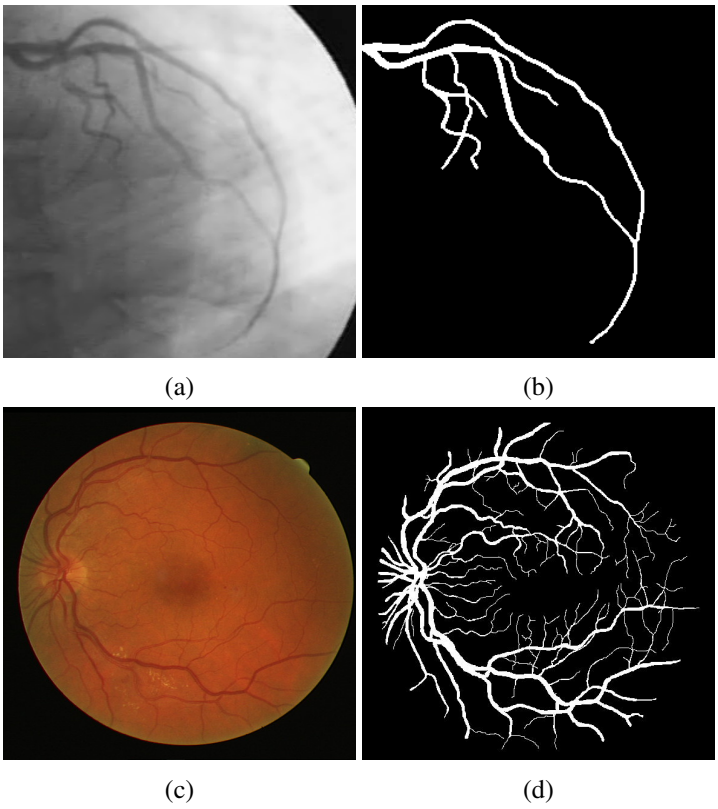
The retina, the inner-most layer of the eye, is an extension of our brain and the only portion of the central nervous system that can be analysed in a non-intrusive manner. The assessment of the retinal blood vessels allows to obtain insight on several systemic conditions such as hypertension, diabetes, and atherosclerosis [9; 10]. Some studies have shown that there is also a connection between structural changes on the retinal vasculature and acute events, such as strokes, which occurred in the past or are about to unfold [11; 12; 13]. Figure 1.1c shows a retinal fundus photo where retinal venules have calibre and tortuosity larger than normal, a sign of vascular occlusion [10]. Almost 80% of the people having diabetes for more than 15 years suffer from diabetic retinopathy, one of the major causes of vision loss worldwide [14]. This becomes even more frightening due to the number of people living with diabetes, which were 463 million adults in 2019 [15]. Diabetic retinopathy impacts the entire retina, as a result of progressive damage to the small retinal blood vessels caused by uncontrolled levels of sugar in the blood. Non-proliferative diabetic retinopathy, the first phase of the disease, generally does not include significant symptoms, as only micro-level structural changes occur, such as the emergence of micro-aneurysms. However, with the accumulation of these modifications, the blood flow may become compromised, and the second phase of the disease - proliferative diabetic retinopathy - is initiated. New blood ves-

sels emerge as the body response to recover from the reduced blood flow. However, these are very fragile and rupture easily, bleeding to the vitreous. Additionally, the risk of glaucoma (damage to the optic nerve) and retina detachment is increased [16]. Vision loss follows this stage of the pathology, however, fortunately it can be prevented in most cases when they are detected early, raising the importance of monitoring and screening programmes.

The analysis of blood vessels is also essential in other scenarios. In cancer therapeutics, it is known that angiogenesis, the process of formation of new blood vessels in a tissue, is promoted by tumour growth, such that it can be used as an indicator for assessing the response to therapeutic interventions [17; 18]. Concerning surgery eligibility and planning, some meticulous procedures can only be conducted after a careful assessment of the local blood vessels. For instance, the proximity of a tumour to major blood vessels may prevent tumour resection [19], and the lack of suitable perforating arteries may discourage the use of a free flap technique for reconstruction [20]. Flap is the designation given to the lifting of tissue from a donor site having intact blood supply and moved to the recipient location. A free flap is completely detached from the donor site, such that the blood vessels are cut and have to be reattached to those at the recipient site. When a surgery is deemed adequate, the planning step aims to gather local landmarks which will be helpful to guide the surgeons during the procedure, and also to select the most proper way of achieving the goal [21]. A particular scenario where this is important and which will be focused in this thesis is the Deep Inferior Epigastric Perforator (DIEP) flap, the state-of-the-art technique for autologous-based breast reconstruction [22]. In this procedure, a portion of belly tissue is extracted without significantly disturbing the abdominal muscle, in order to rebuild the breast. The assessment of the DIEPs (the vessels vascularising that region of the belly) before surgery is essential to guarantee that appropriate blood vessels are included in the extracted flap. This is vital for a good vascularisation of the reconstructed breast after reconnecting the blood vessels to those of the chest.

Given the importance of assessing the different vascular networks of our body, many imaging techniques and protocols have been proposed throughout the years for that particular need. The continuous evolution of technology allowed to obtain clearer images, directly leading to better diagnoses, and to create improved care routines which were not possible in the past. Chapter 2 discusses some of the imaging techniques that are commonly used to visualise blood vessel networks. Due to the huge amount of people suffering from cardiovascular disorders and also the suspicious cases, a large volume of data is acquired in daily clinical practice. Some of the vascular trees of the body are especially complex, such as the retinal, brain, and lung ones. Therefore, blood vessel analysis is a very time consuming and repetitive task. This becomes even more relevant in scenarios like screening programmes, where data from many people is collected at a given time. The computer vision researchers naturally started dedicating their time to design methodologies that could help reducing the burden faced by radiologists. The first methodologies were proposed in the 1980s decade, and since then a lot of researchers have been dedicating their efforts to this problem, following new trends, applications, and requirements. Nowadays, this computer vision topic already comprises a vast literature, and even though great progress has

been made, there are still challenges that require attention. Computer Aided Diagnosis (CAD) systems targeting blood vessel related applications usually divide them into two different steps, blood vessel segmentation, and local characterisation of the segments, yet some methodologies tackle them at once. Nonetheless, the characterisation of the vascular tree is usually simple to solve after a good segmentation has been achieved. In this thesis, we will mostly focus in blood vessel segmentation, the computer vision task seeking an automated or semi-automated extraction of vascular trees in medical images, as shown in Figure 1.2. It separates each unit of the data (pixel or voxel whether the data is Two-dimensional (2D) or Three-Dimensional (3D), respectively) into one of two classes: the foreground, which ideally should only contain blood vessels, and the background, which should include all the remaining body tissues. The characterisation step will be further discussed in Chapter 4.



(a) (b)

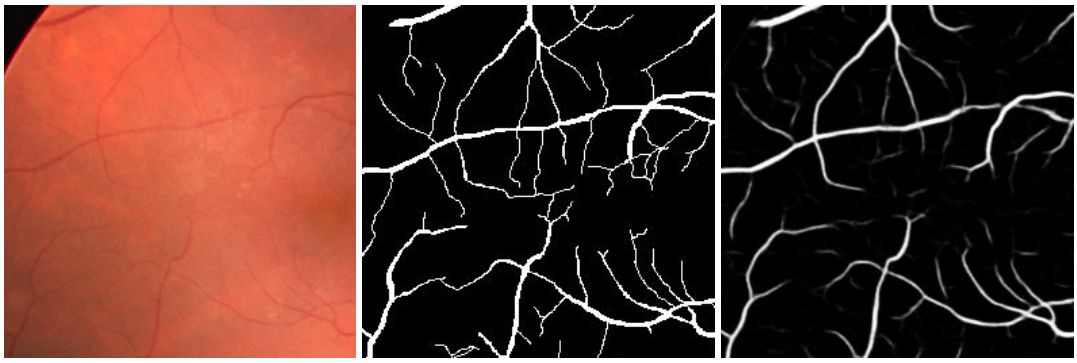
(c) (d)

Figure 1.2: The blood vessel segmentation task, exemplified in a coronary angiogram (a) and a retinal fundus photo (c). The manual annotations of the blood vessels (foreground) in these images by an expert are shown, respectively, in (b) and (d).

1.1 Motivation

Despite all the research that has been performed throughout the years on the topic of blood vessel segmentation, there are still open challenges. Some of them are due to the difficulty of encoding some typical properties of blood vessels in computers, others are related with the continuous

emergence of new applications and imaging techniques, which naturally bring new requirements. Vascular trees possess a very characteristic structure, a graph-like one, similar in many ways to roads and rivers as seen in satellite images. All of these basically consist of piecewise linear segments having occasionally some bifurcations. Moreover, blood vessels have varying sizes and, despite the major ones being more relevant in most clinical practices, the narrow segments also have clinical importance. Even though the state-of-the-art has achieved a very good performance in the larger vessels, there is still margin for improvement regarding narrow ones, as usually it is very challenging to address the compromise between detecting the narrow vessels and the emergence of false positives. As an example, Figure 1.3 shows the predictions obtained with a U-Net [23], a state-of-the-art architecture for biomedical image segmentation, on a challenging region of a retinal fundus photo.

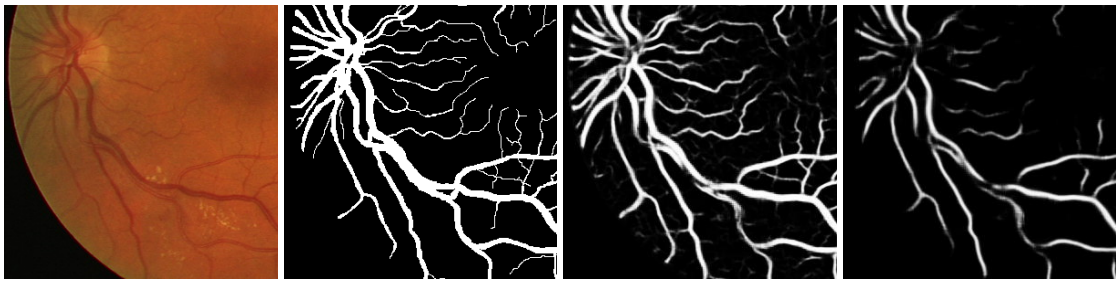


(a) (b) (c)

Figure 1.3: A U-Net model [23] fails at detecting narrow vessels in challenging conditions: (a) retinal fundus photo; (b) expert annotation; and (c) output of a U-Net.

One particular problem where the detection of very small vessels is crucial is in the preoperative planning of breast reconstruction using the DIEP flap technique, as mentioned before. The DIEPs have calibre in the range 1-3 mm, such that it is extremely challenging, even for experts, to characterise them in an accurate and repeatable manner. As reported by the clinical community, it is common to spend around 2 hours in the preoperative planning phase, which could be improved by employing computer vision routines. A CAD methodology could also make the analysis more objective, increasing the repeatability of characterisations. In the past, we proposed the first computer vision methodology for extracting the DIEPs [24], nonetheless clinical validation is required and more advanced techniques were deemed necessary after further testing.

It is well known that deep learning models excel at the tasks where they were trained, given sufficient data and similarity between the train and test distributions. However, when the test data does not have the same distribution, the trained models tend to fail (see Figure 1.4). Thus, it comes as no surprise that one of the trending topics in the machine learning community consists of finding new network designs and training procedures that allow to learn parametrisations that generalise well to different distributions, sometimes even without using any labelled data from the latter. Almost all the literature of blood vessel segmentation focuses on the simpler scenario where the



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

Figure 1.4: Models learned on a given dataset in a normal supervised fashion fail at generalising to sufficiently different data: (a) retinal fundus photo; (b) manual annotation by an expert; (c) segmentation by a model that has been trained on a similar data distribution; and (d) segmentation by a model that has been trained on differently distributed data.

training and test distributions are similar. Interestingly, humans are very good at detecting blood vessels in medical images, even if they are not experts on the topic. In fact, humans excel at finding any graph-like structures in images, such as blood vessels, roads, rivers, and others alike, since they know their underlying properties: they are locally linear, they are connected, their cross-section diameter varies smoothly, and so on. Thus, it seems plausible that advances in the Domain Adaptation field may lead to a state where algorithms generalise properly, performing adequately on significantly different data without requiring training again. This is also relevant in blood vessel segmentation as there is a huge heterogeneity in the amount of data available for the different vascular networks. A general algorithm would allow deep learning known capabilities in terms of performance to become a reality in a broader range of vascular related scenarios.

Finally, the problem of blood vessel segmentation may be interpreted as a multi-pixel classification one, where we output a score for each pixel encoding how confident we are that it belongs to a vessel, thus summing up to a binary classification problem. In such scenario, it is common to consider the minimisation of the Binary Cross-Entropy (BCE) loss, or even the smooth Dice loss, when learning models for solving this task. However, these losses only account for pixel-wise errors, failing to induce properties that are known to be relevant in structures like vascular trees, such as their connectivity. As can be seen in Figure 1.5, these losses do not promote the models to effectively capture topological properties of the objects of interest.

1.2 Research Aims

After introducing the relevance of the segmentation of blood vessels in medical images and discussing some aspects that still require attention, the research aims of this thesis are the following:

- Designing segmentation algorithms that are capable of detecting very small vessels while having a good compromise in terms of false positives;

(a) (b) (c)

Figure 1.5: A U-Net [23] trained with BCE loss easily produces broken segments in challenging conditions: (a) retinal fundus photo; (b) expert annotation; and (c) the obtained segmentation.

- Validating the methodology proposed for the extraction of the DIEPs, and improving the framework in order to better fit the clinical needs;
- Developing deep learning approaches that generalise well to data distributions other than those used during training, in order to better leverage the knowledge extracted from the latter to different vascular networks;
- Exploring novel neural network architectures and/or losses to induce topological coherence in the final segmentations.

1.3 Main contributions

The main contributions of this thesis are the following:

1. The improvement and clinical validation of our preliminary work targeting the DIEPs, where a semi-automatic approach robust to the particular challenges of this use case extracts the subcutaneous and intramuscular course of the DIEPs, and relevant descriptors, such as the calibre and localisation where they perforate the anterior fascia of the abdominal muscle. The outcome of this work shows promising results for achieving a faster and more objective pre-operative planning of DIEP flaps. This research line has lead to several publications and a patent:
    - (Conference) R.J. Araújo, and H.P. Oliveira. Segmentation of the rectus abdominis muscle anterior fascia for the analysis of deep inferior epigastric perforators. In *Iberian Conference on Pattern Recognition and Image Analysis*, pp. 537–545, Springer, Cham, 2017.
    - (Journal abstract) C. Mavioso, J.C. Anacleto, M.A. Vasconcelos, R.J. Araújo, H.P. Oliveira, D. Pinto, P. Gouveia, C. Alves, F. Cardoso, J. Cardoso, and M.J. Cardoso.

The development of an automatic tool to improve perforators detection in angio ct in dieap flap breast reconstruction. *European Journal of Cancer*, 92, 2018.

- (Journal) R.J. Araújo, V. Garrido, C.A. Baraças, M.A. Vasconcelos, C. Mavioso, J. Anacleto, M.J. Cardoso, and H.P. Oliveira. Computer aided detection of deep inferior epigastric perforators in computed tomography angiography scans. *Computerized Medical Imaging and Graphics*, 77, 2019.
- (Journal) C. Mavioso, R.J. Araújo, H.P. Oliveira, J.C. Anacleto, M.A. Vasconcelos, D. Pinto, P. Gouveia, C. Alves, F. Cardoso, J. Cardoso, and M.J. Cardoso. Automatic detetion of perforators for microsurgical reconstruction. *The Breast*, 50, pp. 19–24, 2020.
- (Journal abstract) D. Pinto, C. Mavioso, R.J. Araújo, H.P. Oliveira, J. Anacleto, M.A. Vasconcelos, P. Gouveia, N. Abreu, C. Alves, J.S. Cardoso, M.J. Cardoso, and F. Cardoso. Automatic detection of perforators for microsurgical reconstruction and correlation with patient's body mass index. *European Journal of Cancer*, 138, 2020.
- (Patent) R.J. Araújo, and H.P. Oliveira. Method and apparatus for segmentation of blood vessels, granted in Europe (3352135), China (108324300) and Japan (6776283).

2. A fully convolutional neural network learning single resolution feature maps for the fast extraction of retinal vessels in color fundus photos, enabling competitive performance regarding the state-of-the-art approaches at a faster inference speed. This becomes very relevant at population screening scenarios. This research was presented in a conference:

   - (Conference) R.J. Araújo, J.S. Cardoso, and H.P. Oliveira. A single-resolution fully convolutional network for retinal vessel segmentation in raw fundus images. In *International Conference on Image Processing and Analysis*, pp. 59–69, Springer, Cham, 2019.

3. Learning a data-driven deep metric from the multi-scale Hessian analysis, instead of relying in vesselness functions based on prior knowledge and which are likely sub-optimal. This approach surpasses the enhancement of traditional vesselness filters, both in same and different data distributions, and also generalises better than more complex deep learning models. This work was published in:

   - (Conference) R.J. Araújo, J.S. Cardoso, H.P. Oliveira. Deep vesselness measure from scale-space analysis of hessian matrix eigenvalues. In *Iberian Conference on Pattern Recognition and Image Analysis*, pp. 473–484, Springer, Cham, 2019.

4. Raising awareness and designing mechanisms to deal with the literature limitations regarding the topological properties of blood vessel masks. A similarity index is proposed to enable the objective and efficient benchmarking of these properties. Moreover, two approaches are designed to promote the learning of deep neural network models that produce

better segmentations topological-wise. One of them tackles the issue from an architecture design point of view, through a cascade of segmentation and probabilistic refinement models. The topological errors are interpreted as a noise process and the probabilistic refiner operates as a denoising model. The second approach addresses the challenge via a more adequate loss function. A framework based on the morphological closing operator is designed to penalise centreline errors that induce broken trees or merge distinct ones. This research line has originated the following publications:

- (Conference) R.J. Araújo, J.S. Cardoso, and H.P. Oliveira. A deep learning design for improving topology coherence in blood vessel segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 93–101, Springer, Cham, 2019.
- (Pre-print) R.J. Araújo, J.S. Cardoso, and H.P. Oliveira. Topological blood vessel segmentation, arXiv, 2021.

5. The design of a novel parametric active contour model which extends the state-of-the-art capabilities, by allowing to have a set of regions with different bending properties, such that some dynamics of an object may be fitted while others disregarded. This is a wanted behavior when some of the dynamics are caused by noise or other processes. A real scenario where this is relevant concerns the lung segmentation, where we wish to include as many peripheral nodules as possible, and simultaneously, fit accurately the natural dynamics of the lung boundary. This work was published in:

   - (Journal) R.J. Araújo, K. Fernandes, J.S. Cardoso. Sparse multi-bending snakes. *IEEE Transactions on Image Processing*, 28(8), pp. 3898–3909, 2019.

1.4 Document Structure

Chapter 1 introduced the problem of blood vessel segmentation, its relevance for the clinical practitioners and how it captured the interest of the computer vision community. Our motivation, research aims and main contributions were stated. The rest of this thesis is structured as follows:

- Chapter 2 presents a non-exhaustive description of the blood vessel trees that are related to the experiments we discuss along the thesis. The used imaging techniques and the respective databases are also described in this Chapter;
- Chapter 3 reviews the literature concerning the research that has been made in blood vessel segmentation and analysis;
- Chapter 4 describes the methodologies we designed for two particular scenarios of blood vessel segmentation, the DIEPs and the retinal vessels, with a focus on the narrower vessels;

- Chapter 5 contains our experiments towards general blood vessel segmentation, by introducing regularisation based on strong prior knowledge coming from the eigenvalue analysis of the Hessian matrix at different scales;
- Chapter 6 presents our two contributions promoting topological coherence of the segmented blood vessels, one by means of architecture design, and the other by introducing a loss function which helps achieving that goal; Moreover, it also describes the novel similarity index for benchmarking the topological coherence of blood vessel masks;
- Chapter 7 discusses a novel extension of parametric active contour models, allowing them to have a finite set of regions behaving differently in terms of bending resistance;
- Chapter 8 presents the main conclusions of the thesis and discusses future work that is expected to bring important contributions to this field.

Chapter 2

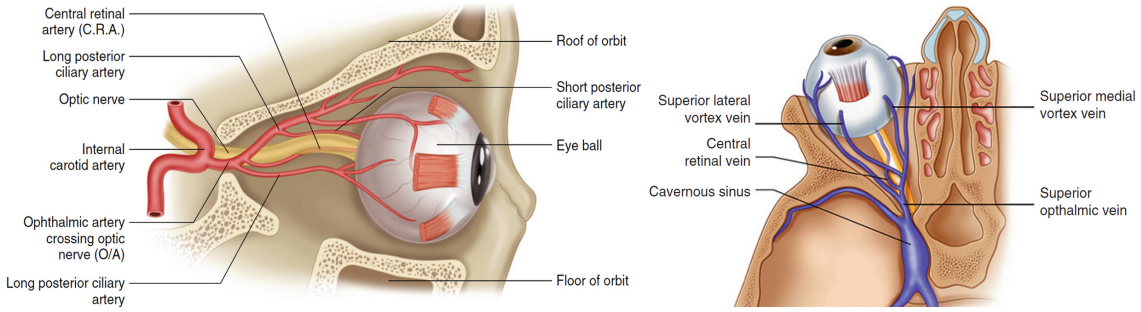
Anatomy, Imaging, and Datasets of Blood Vessels

Blood vessels share common properties independently of their location in the human body. They are elongated piecewise linear structures whose diameter decreases as they evolve from their root. Nonetheless, the complexity of vascular networks varies between distinct anatomical regions. The structure of a given blood vessel tree directly impacts the type of methodologies that are suitable for its segmentation. This Chapter briefly describes, according to [25; 26], the anatomy of the vascular networks which are related with the experiments performed in this thesis: the retinal, coronary, deep inferior epigastric, and pulmonary networks. The imaging techniques which are commonly used for assessing each of these trees and the benchmarks for blood vessel segmentation which we consider for the evaluation of the proposed algorithms will be presented.

2.1 Retinal blood vessels

The retina provides the opportunity of obtaining insight on many medical conditions in a non-invasive manner. It is vascularised by two different arterial blood vessel trees, both having origin in the ophthalmic artery, which in turn is a bifurcation of the internal carotid artery, as represented in Figure 2.1a. The posterior ciliary arteries irrigate the outer and middle layers of the retina, while the central retinal artery supplies the inner retina. The latter enters the optic nerve and divides into four main branches spreading radially away from the optic disc and curving around the macula. These branches may be assessed when obtaining images of the retinal fundus, which is represented in Figure 2.2. Veins are usually darker and slightly wider than their arterial counterpart. The central retinal vein drains blood from the capillaries of the retina into the superior ophthalmic vein or into the cavernous sinus. Finally, the blood returns to the heart via the superior vena cava.

Direct ophthalmoscopy is the main technique for visualising the retinal fundus in primary and emergency care, where the physician uses a hand-held device with lenses to inspect the back of the eye and search for abnormalities, determining if the patient should be referred to an appointment with an ophthalmologist [28; 29]. Dilation of the pupil is generally required for a clear



(a) (b)

Figure 2.1: Vascularisation of the retina: (a) arterial supply, and (b) venous drainage. Adapted from [27].

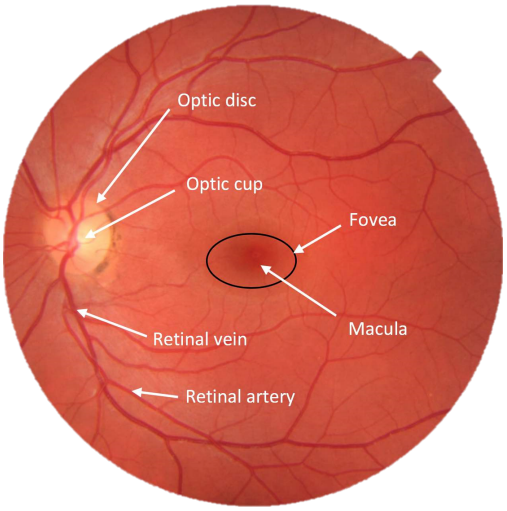
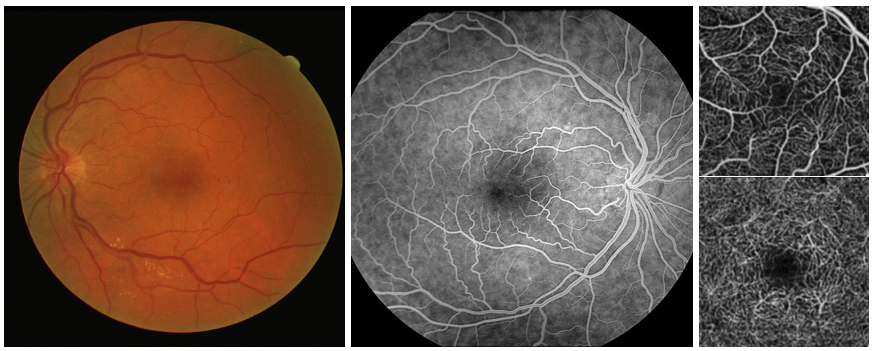


Figure 2.2: Representation of the retinal fundus[1].

assessment [30]. Furthermore, this examination requires extensive practice and confidence, and the acquisition of such skill seems to be often overlooked [31; 32]. Fundus photography uses specialised cameras integrating a microscope to obtain pictures of the retinal fundus. This technique is frequently used since it is very simple and allows to document the findings, which is crucial to assess the progress of any abnormality or disease in future medical appointments. It is less invasive, as no pupil dilation is required and it enables telemedicine, which is becoming essential nowadays due to the progression of diabetes, especially in low and middle-income countries, where the number of experts may not be sufficient to analyse all the data [32; 33]. The COVID-19 pandemic has further shown the relevance of tele-ophthalmology [34]. In a face-to-face consultation, dyes such as fluorescein and indocyanine green may be injected in the blood flow to better visualise, respectively, retinal and choroidal blood vessels [35; 36]. The main disadvantage of fundus photography is the lack of depth information and portability. Optical Coherence Tomog-

[1] Adapted from https://simpleosce.com/img/mainpics/fundoscopy/fundus.jpg

raphy Angiography (OCTA) has been recently proposed as an additional non-invasive technique for assessing the fundus. It identifies blood vessels due to the varying laser light reflectance at the surface of red blood cells, being capable of distinguishing vessels at different tissue layers and providing more depth information. It is useful for diagnosing many retinal conditions, even though it is extremely motion-sensitive, thus very dependent on patient collaboration, and has a relatively small field of view [37]. Figure 2.3 shows images of the retinal fundus acquired with the described techniques.

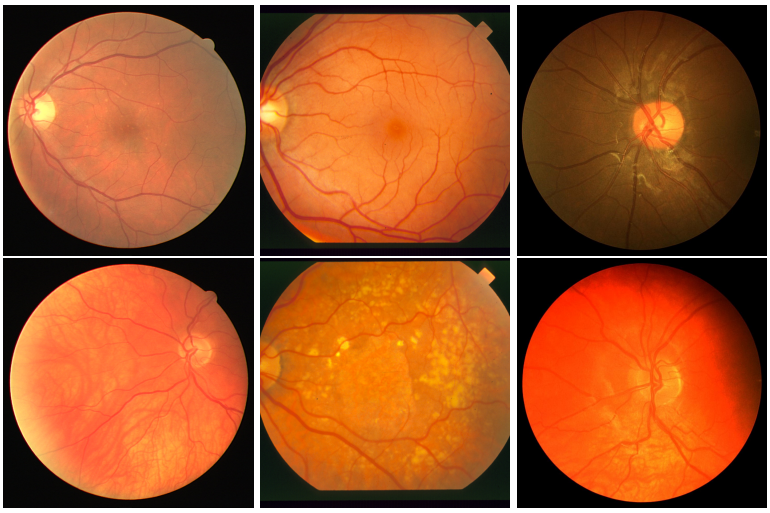


(a) (b) (c)

Figure 2.3: Images of the retinal fundus acquired with different imaging techniques: (a) Color Fundus Photography from DRIVE [38], (b) Fluorescein Angiography[2], and (c) OCTA adapted from [37], where two different layers are shown, the superficial vessels (top) and deep vessels (bottom).

Several color fundus photography benchmarks are publicly available, even though only some of them have manual annotations of the blood vessels. The retinal vascular network is the most targeted one by the computer vision community, possibly due to the wide availability of public data. In this thesis, we resort to the three mostly used benchmarks in the literature: DRIVE [38], STARE [39], and CHASEDB1 [40]. DRIVE is a result of a diabetic retinopathy screening program conducted in The Netherlands comprising 40 images, 7 of which showing pathology, divided into training and test sets, each with 20 images. There is available one manual annotation for each of the training images, while there exist two for the test ones, such that it is possible to compare the performance of an algorithm with that of a human. STARE is a dataset containing 400 images with expert annotations of 39 different pathologies. From all the images, 80 have available the ground truth of optic nerve detection, 10 artery/veins labeling and 20 of them the manual annotation of blood vessels. Two experts have provided their annotations of the images, and a large variability exists regarding the segmentation of the narrow vessels. Half of these images have pathology. CHASEDB1 was compiled after the Child Heart and Health Study in England, containing images of the left and right retinas of 14 different individuals. Central reflex is particularly abundant in this dataset. The images are centred on the optic disc with full field illumination. Figure 2.4 shows

[2]Adapted from https://www.laretinacenter.com/services/fluorescein-angiography



(a) (b) (c)

Figure 2.4: Example images contained in the benchmarks of retinal vessel segmentation: (a) DRIVE [38]; (b) STARE [39]; and (c) CHASEDB1 [40]. Healthy (top) and pathological cases (bottom) are shown.

healthy and pathological images from each of these benchmarks. Table 2.1 contains a summary of the main characteristics of each dataset.

Table 2.1: Summary of the characteristics of the databases used in this thesis for retinal blood vessel segmentation.

Dataset	Camera	No. images	Resolution	FOV (°)
DRIVE	Canon Cr5 non-mydriatic 3CCD	40	584×565	45
STARE	TopCon TRV-50	20	605×700	35
CHASEDB1	Nidek non-mydriatic handheld camera	28	960×999	30

2.2 Coronary vessel tree

The heart tissue, as any other tissue in the human body, requires oxygen to keep its activities. The coronary vessel network is responsible for the vascularisation of the heart muscle (see Figure 2.5). The coronary arterial tree (Figure 2.5a) takes oxygenated blood to the heart muscle. The two main coronary arteries have origin in the aorta. One of them, the left main coronary artery, splits into the circumflex artery and the left anterior descending artery. The first vascularises the left atrium and the lateral and posterior regions of the left ventricle. The latter supplies the anterior and bottom regions of the left ventricle and the anterior portion of the septum. The other main coronary artery, the right coronary artery, takes blood to the right atrium, right ventricle, bottom portion of both ventricles, and to the posterior region of the septum. The coronary venous tree (Figure 2.5b) takes deoxygenated blood back to the right atrium of the heart, along with some

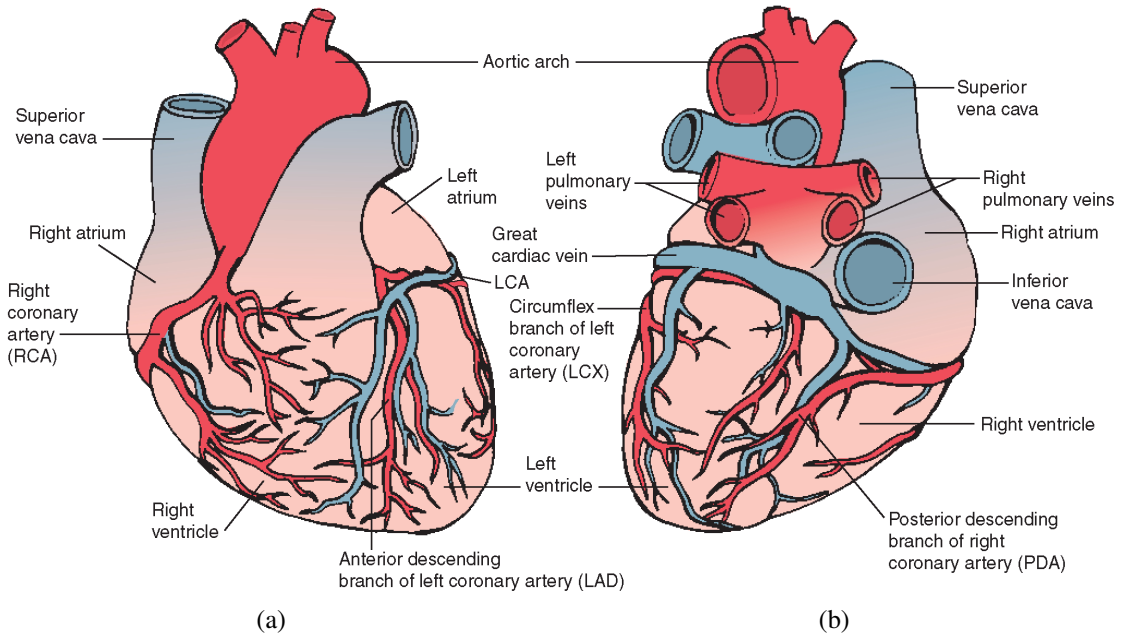


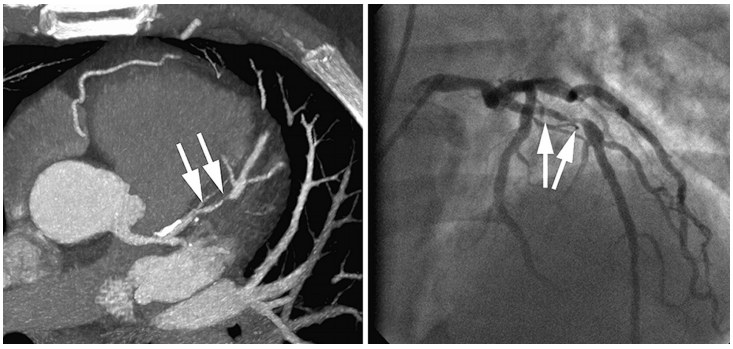
Figure 2.5: Coronary (a) arterial and (b) venous trees[3].

metabolic waste compounds. From the numerous capillaries lying in the heart tissue, the blood flows to venules which join into veins. Most of them return the blood to the right atrium through the coronary sinus.

Practitioners request imaging of coronaries when there are symptoms of coronary artery disease (such as chest pain), when they suspect there are restrictions of blood flow to the heart, or when abnormal findings were obtained in a non-invasive test, for instance an electrocardiogram, an ecocardiogram, or a stress test [41; 42]. Coronary angiogram is the state-of-the-art approach for obtaining images of the coronaries [42]. It comprises catheterisation to inject a radio-opaque contrast agent that allows to obtain clear images of the vessels using X-ray beams. The most commonly used technique is the Digital Subtraction Angiography, where an image is obtained before injecting the contrast and allows to remove the background signal from the frames acquired during the study, enhancing the visualisation of blood vessels. This procedure, besides diagnosis, can be combined with other catheterisations for repairing some conditions, such as removing clots and placing stents in narrowed arteries [43]. Due to the invasive behaviour, it is a relatively long and expensive procedure, and comprises some risks, even though they are rare (infection, vessel rupture, allergy, among others). Due to the advances of the Computed Tomography (CT) scanning technology (where a computer combines several X-ray images taken at different angles to produce cross-sectional images of the tissues - volumetric data), the CT Angiography (CTA) is nowadays being suggested as the first imaging protocol to consider in low to intermediate risk scenarios [44]. It requires the use of a iodine-based dye but no catheterisation is needed. Additionally, it allows

[3] Adapted from http://what-when-how.com/nursing/the-cardiovascular-system-structure-and-function-nursing-part-2/

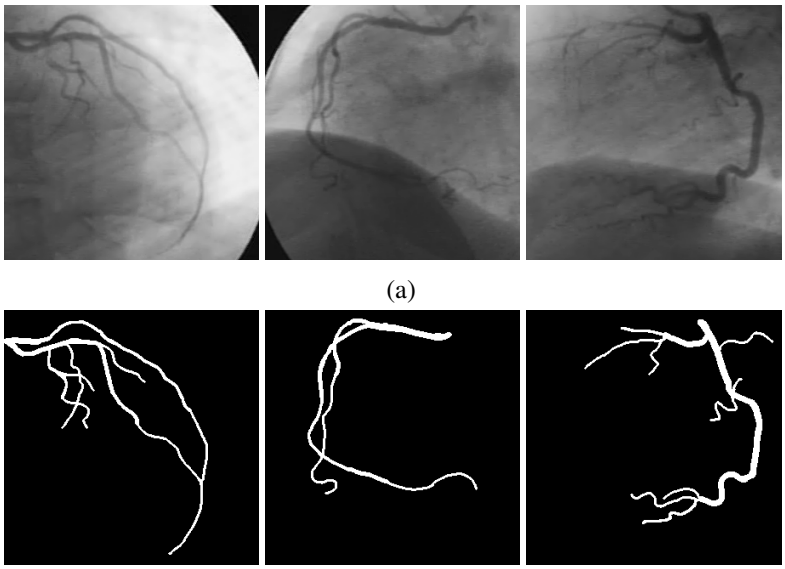
to analyse other tissues which may be the real sources of discomfort, such as those belonging to the lungs or nearby bones. Figure 2.6 provides examples of coronary images acquired with each of these techniques.



(a) (b)

Figure 2.6: Coronary images obtained with different techniques: (a) a slice of a coronary CTA, and (b) a coronary angiogram. White arrows highlight a stenosis. Adapted from [45].

Recently, a benchmark for the segmentation of coronary blood vessels in coronary angiograms has become publicly available [46]. The dataset contains 134 angiograms and the respective annotations of the blood vessels outlined by a cardiologist. It was provided by the Cardiology Department of the Mexican Social Security Institute. The images have a size of 300×300 pixels and are grey-scale. Figure 2.7 shows some examples of the data contained in this dataset.



(a)

(b)

Figure 2.7: Example data from a publicly available coronary dataset [46]: (a) coronary angiograms, and (b) respective manual annotations of the blood vessels.

2.3 Inferior epigastric vasculature

The Deep Inferior Epigastric Artery (DIEA) provides blood to the anterior region of the abdominal wall, more precisely the inferior portion. It has origin in the external iliac artery above the inguinal ligament and ascends along the medial margin of the abdominal inguinal ring. Afterwards, it pierces the posterior lamella of the rectus abdominis muscle sheath, and ascends between the lamella and the muscle, until it divides in branches that anastomose with the superior epigastric artery (see Figure 2.8).

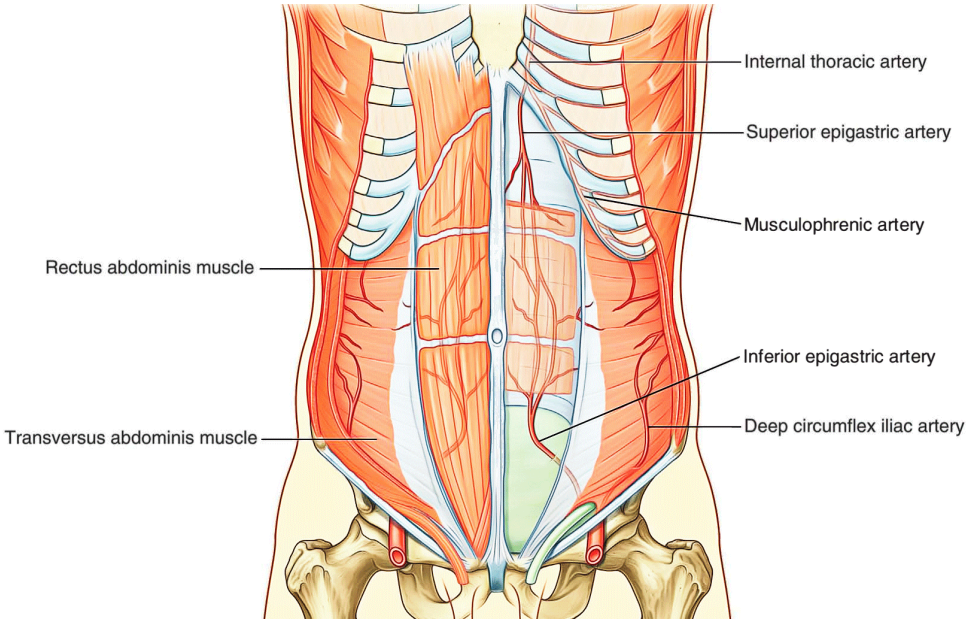


Figure 2.8: The abdominal wall vasculature[4]. The left side shows the superficial blood vessels, while the deep vascular tree (including the DIEA) is present on the right.

After the DIEA pierces the muscle sheath, it occasionally gives origin to small vessels, the DIEPs (or just epigastric perforators), as represented in Figure 2.9. Their designation comes from the fact that they perforate the abdominal muscle to vascularise the subcutaneous region. The number of perforators is variable, but usually there are between 6 and 8 of these blood vessels. Their intramuscular course is highly variable both in terms of length and tortuosity. After leaving the muscle, the perforators may also present several courses and branching configurations, and usually there are anastomoses with the superficial inferior epigastric system. In the subcutaneous region, deoxygenated blood moves to venules that drain into the inferior epigastric vein, which is the venous counterpart of the DIEA. The blood then flows to the external iliac vein, in direction to the inferior vena cava.

As described before, clinicians study these blood vessels during the preoperative planning of DIEP flap based breast reconstruction. A precise description of each perforator is required to design a flap that is properly vascularised and not excessively difficult to extract. The calibre of

[4]Adapted from https://www.earthslab.com/anatomy/abdominal-wall/

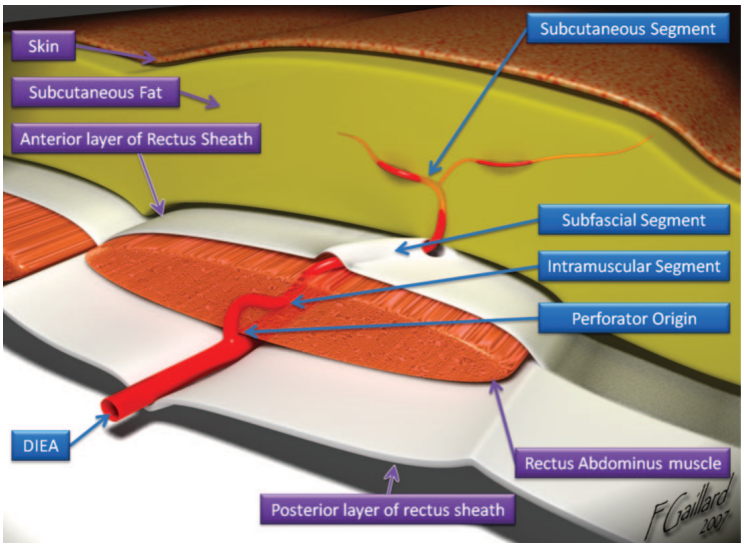
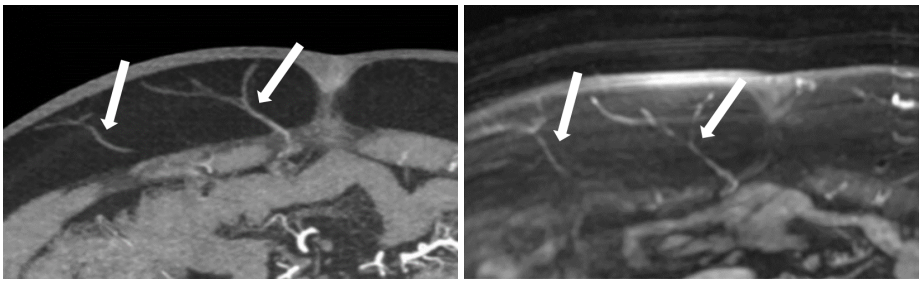


Figure 2.9: 3D representation of the course of a DIEP vessel. Adapted from [26].

perforators is usually taken into account to assess how adequate a perforator is for vascularising the tissue and linear and short intramuscular courses are preferred for easier extraction [22]. The standard technique to visualise the epigastric perforators is CTA [47; 48; 22], even though Magnetic Resonance Angiography (MRA) is also a viable option [49; 50]. It uses gadolinium-based contrast agents instead of iodine-based ones, and avoids radiation by relying on the dynamics shown by protons in the presence of a strong magnetic field and after a radio-frequency current traverses the patient. Images acquired with both techniques are presented in Figure 2.10. CTA has

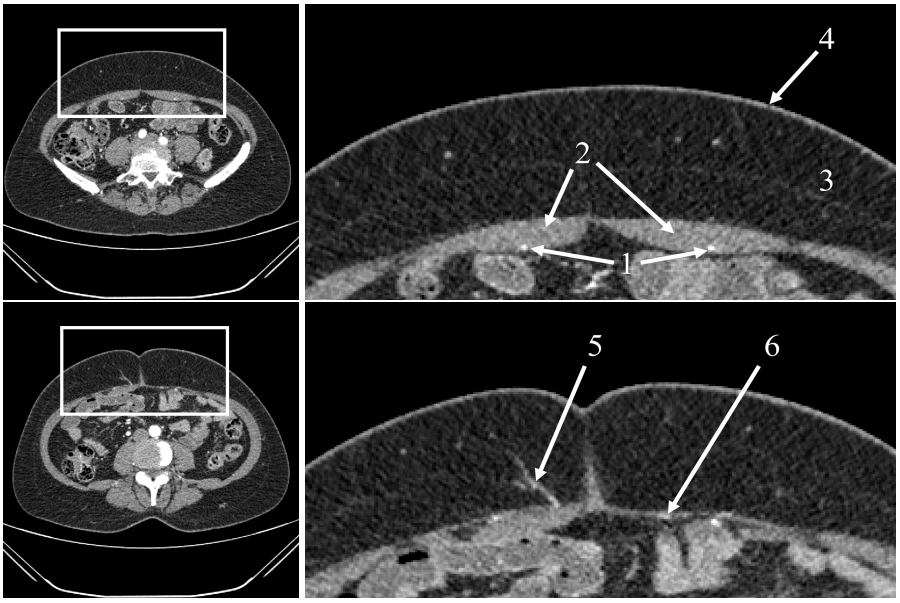


(a) (b)

Figure 2.10: Maximum intensity projection reconstruction of axial subvolume in (a) CTA and (b) MRA scans, for the visualisation of inferior epigastric perforators (highlighted by white arrows). Adapted from [49].

a lower cost but involves radiation, which may become problematic in patients that have already been subjected to radiation a significant number of times in the past. The contrast agent used in CTA is also more prone for inducing allergic reaction. Nonetheless, MRA has a larger acquisition time, is prone to include noise related with respiratory movement, and the overall acquisition process is more arduous to the patient [50].

To the best of our knowledge, no public databases exist for assessing the segmentation or analysis of the epigastric perforators. During this thesis, in a collaboration with the Breast Unit of the Champalimaud Foundation, a proprietary database containing CTAs from 40 patients was assembled. After obtaining the consensus of the patients to participate in the study, the images were acquired with a CT Spectral Scanner with a 64-detector row (Philips Igon), a 140 kV tube voltage and a 120-160 mA fixed tube current. Further settings are 0.75 rotation time, collimation of 128, pitch of 0.609 and a 512 matrix. 80-100 ml of contrast were injected and images acquired in Bolus Tracking. Images were collected starting from 5 cm above the umbilicus and until the lesser trochanter (range of 32 cm) with a thickness of 0.8 mm and an increment of 0.4 mm. Figure 2.11 shows examples of how DIEAs and DIEPs appear in the collected volumes of data. It is possible to see that the vessels are very small when comparing with the neighbouring structures.



(a) (b)

Figure 2.11: Example data comprised in the proprietary database concerning the DIEPs: (a) axial slices of a CTA scan, and (b) the respective region of interest for the analysis of DIEPs. Some relevant structures are labelled: 1) DIEAs, 2) rectus abdominis muscle, 3) subcutaneous region, 4) skin, 5) subcutaneous course of a DIEP, 6) intramuscular course of a DIEP.

2.4 Pulmonary tree

The lungs are the organs responsible for the respiration process. Millions of alveoli are involved in this task, each of them surrounded by blood capillaries, such that the pulmonary vessel network is very complex. A representation of the lung vasculature is depicted in Figure 2.12.

Deoxygenated blood leaves the heart from the right ventricle to the pulmonary trunk, which gives origin to the right and left pulmonary arteries ahead. Each of them conducts the blood

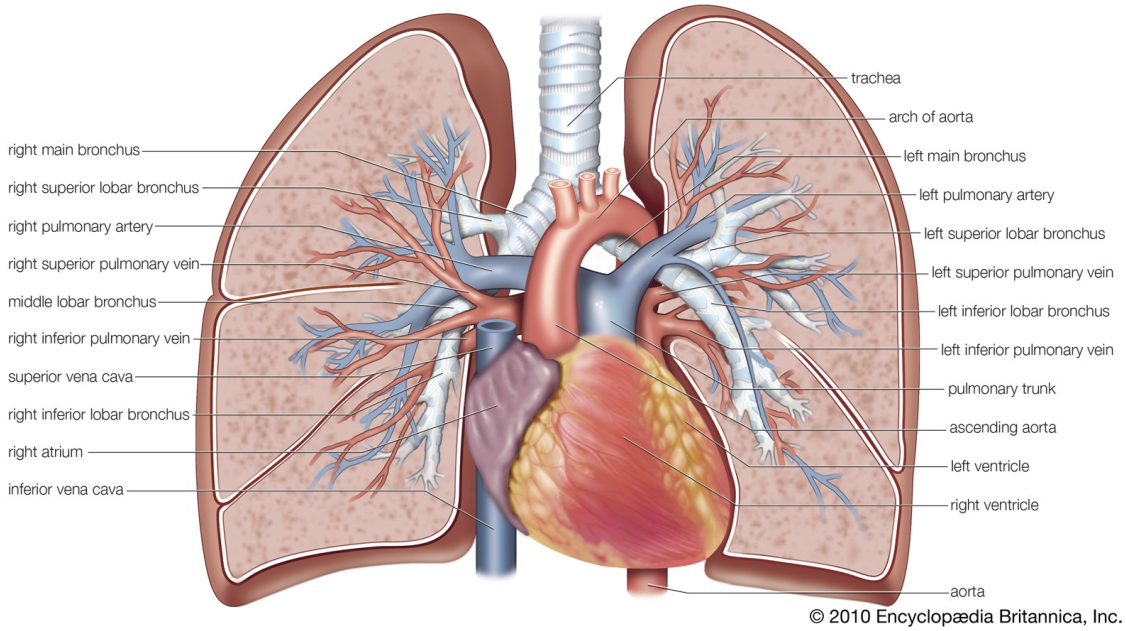


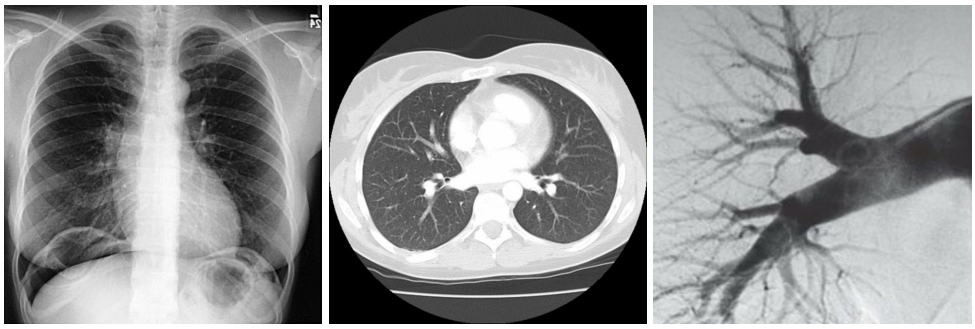
Figure 2.12: Representation of the pulmonary vessel tree[5].

into the respective lung, where the oxygenation process occurs. The arborisation pattern of each pulmonary artery is different, but it closely follows the branching pattern of the bronchial three. It leads to the truncal, lobar, segmental, and subsegmental arterial segments, which take blood into the different lobes of the lung. In the alveolar capillaries, the oxygenation of the blood occurs and it is then returned via pulmonary veins to the left atrial chamber of the heart.

Chest X-rays will be the first request of most practitioners who suspect the patient has a lung disease. This technique is capable of detecting some lung nodules and conditions such as pneumonia, infection, and lung emphysema [51]. When the findings are not conclusive or a condition has to be better characterised, a CT scan is usually conducted [52]. The 3D view of the data allows to have more detail when assessing the structures of the lung, especially blood vessels and nodules. The use of contrast (CTA) may be requested by the clinician, in order to further enhance the signal of blood vessels. A good visualisation of the vascular tree is essential to characterise the surroundings of a nodule, which is relevant for its characterisation, and also to detect pulmonary embolism [53]. Pulmonary angiogram is uncommon and is usually only performed when there is a suspicion of a pulmonary embolism requiring treatment, which can be conducted during catheterisation [54; 55]. Figure 2.13 shows example images acquired with the described techniques.

Regarding datasets dedicated to the segmentation of lung blood vessels, a competition named VESSEL12 was held in the past and the data is still available[6]. It comprises 20 different CT scans, 9 of them with contrast, 5 of them low-dose in radiation, and 1 is a high-resolution CT scan. CTAs were acquired with a Siemens SOMATOM Sensation 64 or a Toshiba Acquilion

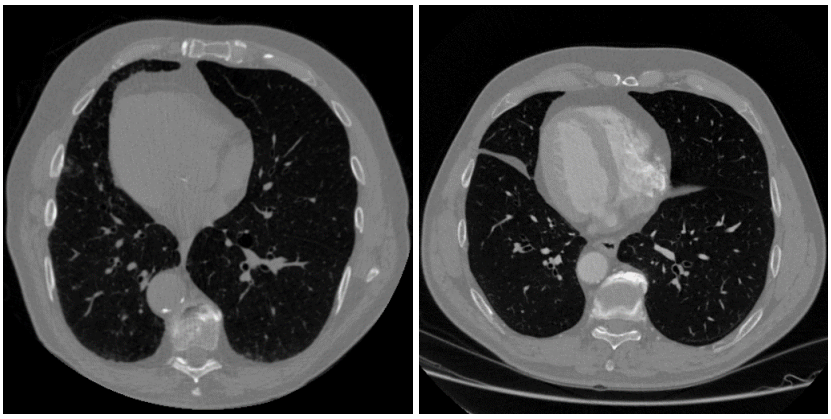
[5] Adapted from https://www.britannica.com/science/pulmonary-circulation
[6] https://vessel12.grand-challenge.org



(a) (b) (c)

Figure 2.13: Images of the lung acquired with a (a) chest X-ray[7], (b) chest CT scan[8], and (c) pulmonary digital subtraction angiography[9].

ONE. The latter was used to obtain all of the low-dose CT scans and also the high-resolution one. The regular CT scans were produced with either a Phillips Mx8000 IDT 16 or a Philips Brilliance 16P. Different reconstruction kernels, and axial and inter-slice resolutions have been considered. Further details are present in the overview article related to the competition [56]. Most of the scans contain relevant findings, such as interstitial lung disease, pulmonary thromboembolism, and lung nodules. A complete annotation of the blood vessels does not exist, as the effort required in delineating them in the 3D data would be extremely large. Nevertheless, labels for a set of voxels are provided, where classes such as blood vessel, nodule, lung parenchyma, airway wall, and lesion are considered. Figure 2.14 shows some images contained in the database.



(a) (b)

Figure 2.14: Example axial slices from volumes included in the VESSEL12 dataset: (a) a slice from a regular chest CT scan, and (b) one from a CTA scan, where the signal of blood vessels is further enhanced.

[7] Adapted from https://radiologykey.com/the-normal-chest-x-ray-reading-like-the-pros/
[8] Adapted from https://radiopaedia.org/cases/normal-chest-ct-lung-window-1
[9] Adapted from https://thoracickey.com/pulmonary-angiography/

Chapter 3

Literature Review

This Chapter provides a description of how the literature related to blood vessel segmentation has been evolving through the years. The algorithms are divided in two main categories: unsupervised and label-driven approaches. Unsupervised methods encode known properties of vascular trees into different enhancement or extraction frameworks, whereas label-driven methodologies rely on annotated data and machine learning (ML) techniques to produce segmentations, either from a set of hand-crafted features (traditional ML) or from images themselves (deep learning (DL)). Traditional ML also heavily relies on prior knowledge since the features used in the learning process have to embed useful properties of vascular trees. DL may also imbue prior knowledge in the form of regularisation of feature representations, nonetheless this is usually not required for obtaining well-performing models when an adequate architecture is used.

There are relevant surveys in the literature concerning the topic of vessel segmentation: Kirbas and Quek [57] reviewed in 2003 the main approaches at that time targeting the extraction of tubular-shaped structures, with a heavy focus on blood vessels; Lesage et al. [58] described algorithms for the extraction of blood vessels in 3D data in 2009; Fraz et al. [59] surveyed in 2012 the main approaches for retinal blood vessel segmentation, a largely focused vascular network due to the large number of available datasets; recently, in 2018, Moccia et al. [60] discussed the latest trends in the field, focusing ML, deformable models and tracking methods.

3.1 Unsupervised approaches

This Section covers algorithms making several assumptions on the structure of blood vessels (such as being piecewise linear structures and having smooth calibre variation) in order to design unsupervised enhancement functions or extraction frameworks. Even though these methodologies are unsupervised, they may be data-driven in the sense that clustering approaches and the statistics of particular datasets can be explored. Moreover, computer vision practitioners often analyse the data to understand the challenges involved in a given visual task. Regarding the particular scenario of blood vessel segmentation, prior knowledge about vessel structure often does not suffice. For instance, there may exist neighbouring structures displaying similar properties or even pathologies

distorting the expected patterns. Even then, unsupervised methods do not rely on annotations of the data.

Regarding the enhancement of blood vessels, matched filtering, Hessian analysis, and mathematical morphology are reviewed. Even though these methodologies do not provide a segmentation, a simple threshold can be used to obtain it. Concerning frameworks for blood vessel extraction, centreline tracking, region growing, active contour models, and graph-cuts approaches are considered.

3.1.1 Matched Filtering

Matched filtering for vessel enhancement was introduced by Chaudhuri et al. [61] and became a popular technique since then. It consists in the design of kernels that account for the grey-level profile of vessels, such that their convolution with the data produces high responses at vessel locations. Chaudhuri et al. [61] regarded vessels as piecewise linear structures with a different intensity than the background (lower in their use case) and whose diameter decreases very gradually. The authors analysed the profiles of the cross sections of different vessels and concluded that they resembled a Gaussian shape, albeit not ideal and showing variation among vessels. Based on this intuition, the authors designed a 2D kernel that is decomposable in two One-Dimensional (1D) linear filters: a zero-mean Gaussian (which produces higher responses when aligned with the cross section of the vessel) and a constant filter (to account for a smoothing effect along the vessel). Note that to deal with vessel orientation, the kernel must be convolved with the image at different orientations. The final response at a given pixel is given by the maximum response across the different orientations.

A supervised variant was introduced by Al-Rawi et al. [62], where labelled data was used to infer the best filter parameters for a specific dataset, namely the length of the kernel, the side trail truncation level, and the standard deviation of the Gaussian. However, the best approach to deal with vessels of varying widths in medical images is to use a bank of filters. Such is accomplished by varying the standard deviation and truncation of the zero-mean Gaussian. Figure 3.1 shows a bank of three different kernels.

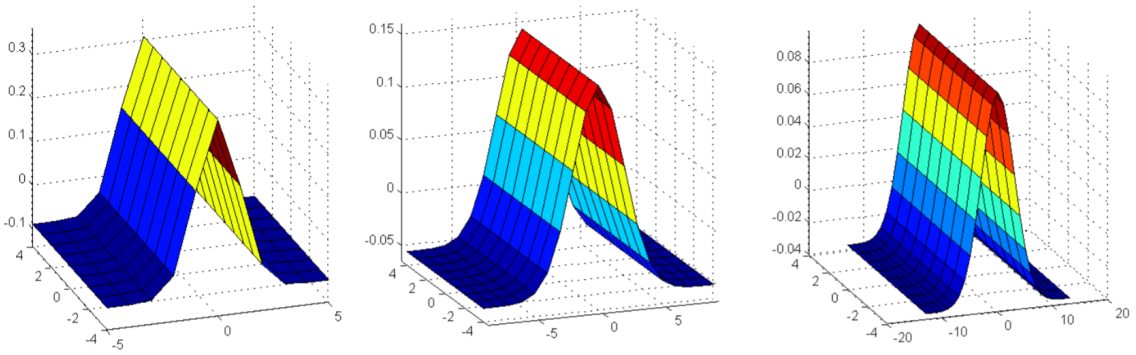


Figure 3.1: 2D kernels for enhancing vessels of increasing diameter, from left to right. Notice the increasing scale of the axis related to the zero-mean Gaussian.

Imagining a scenario where 4 kernels are used at 12 different rotations, a total of 48 convolutions are required. This has a relevant computational cost and explains why this methodology has been essentially applied to 2D images, as the 3D case has even more degrees of freedom. Poli and Valli [63] targeted this efficiency aspect and proposed a methodology for the decomposition of large kernels into small atomic masks, establishing a framework for real-time detection of vessels. The following research has mainly focused two aspects: the proper way of representing the cross section profiles of vessels, which is strongly dependent of the image acquisition step and the used protocol; and the design of filters for better enhancement of vessels, while decreasing the response to other structures showing step-like edges, such as the optic disk and bright lesions in retinal images. The cross section of vessels has been commonly modelled as a Gaussian [61; 63; 39; 64; 65; 66]. Even then, it has also been regarded as bar-like [67; 68] and more complex functions [69]. Regarding kernel design, besides the zero-mean Gaussian [61; 39; 66], the second-order derivative of the Gaussian [64; 65; 67] and multi-wavelet kernels [70] have been used for obtaining high responses at the cross section of vessels. Zhang et al. [66] addressed the high responses at non-vessel locations provided by Gaussian based kernels. The authors explored the fact that a vessel cross section is ideally symmetric, while a step-edge has a strong asymmetric behaviour. Thus, besides the typical second-order derivative of Gaussian kernel, a first-order derivative of the Gaussian was also considered. The analysis of the local average response to the latter filter allows to distinguish these two types of profiles, as shown in Figure 3.2. Wang et al. [70] used the multi-wavelet system of [71] and deviated an approach to distinguish step edges from vessel edges. Odstrcilik et al. [72] learned from a database five different cross section profiles, each for a specific vessel width.

Azzopardi et al. [68] enhanced piecewise linear structures by locally analysing the response to a set of shifted Difference of Gaussian filters. The authors designed two different filters from those responses, one which is symmetric for the enhancement of vessel segments, and another which is asymmetrical, in order to adequately enhance vessel endings. They designated these filters as, respectively, symmetric and asymmetric Bar-Combination Of Shifted FIlter REsponses (B-COSFIRE).

3.1.2 Hessian based filters

The second-order local intensity variations have been widely used in the past to enhance vascular structures in medical data [73; 74; 75; 76; 77; 78; 79]. As vessels are piecewise linear structures, they are expected to produce large responses when a second derivative of a 1D Gaussian is convolved along their cross section profile. In addition, such filter should not lead to significant responses when applied along their length. This intuition favoured the use of the Hessian matrix, H, which encodes the second order partial derivatives of the image. The Hessian of a scalar-valued function $f(\mathbf{x})$, where $\mathbf{x} = [x_1, ..., x_D]^T$ is a D-dimensional vector, is a $D \times D$ matrix with elements

$$H_{ij}(\mathbf{x}) = \frac{\delta^2 f(\mathbf{x})}{\delta x_i \delta x_j} \quad \text{for } i, j = 1, \ldots, D \tag{3.1}$$

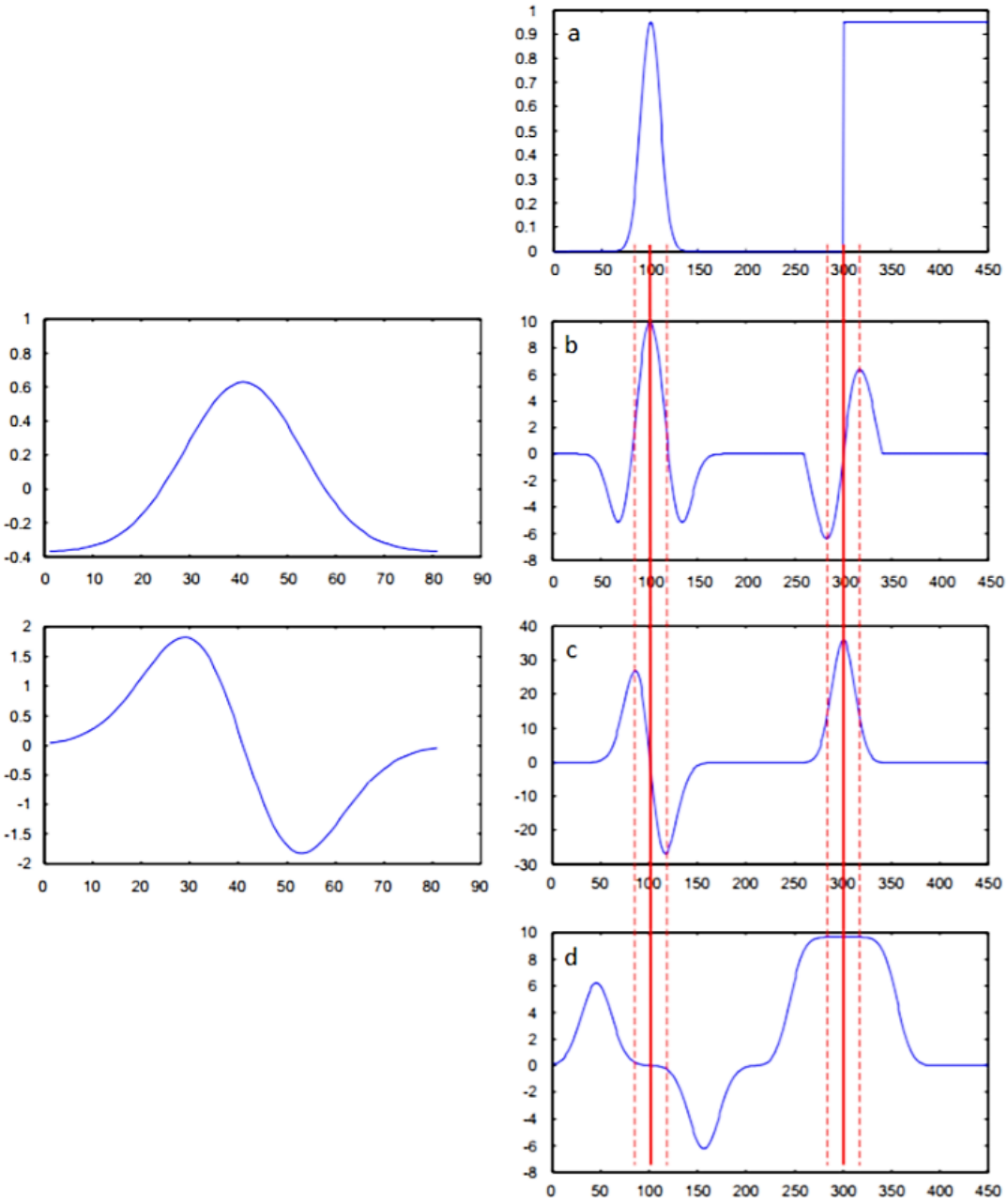


Figure 3.2: Responses of zero-mean Gaussian and first-order derivative of the Gaussian (top left and bottom left, respectively) to a Gaussian cross-section and an ideal step-edge: (a) a Gaussian cross-section and an ideal step-edge; (b) the response to the zero-mean Gaussian; (c) the response to the first-order derivative of the Gaussian; (d) the local mean of the response to the latter. Adapted from [66].

In the particular case where the goal is to enhance vascular structures of varying width, it is important to evaluate H at different scales. Let $\mathbf{I}$ be a D-dimensional image, such that $\mathbf{I}(\mathbf{x})$ denotes the intensity at position $\mathbf{x}$. If $f(\mathbf{x})$ is simply given by $\mathbf{I}(\mathbf{x})$, only a single scale is addressed. Thus, previous works have evaluated H at a scale-space by blurring the original data with

Gaussian kernels of varying σ. This has the additional advantage of reducing noise. Note that $\delta^2(\mathbf{I}(\mathbf{x}) * G(\mathbf{x},\sigma))/(\delta x_i \delta x_j)$ and $\mathbf{I}(\mathbf{x}) * \delta^2 G(\mathbf{x},\sigma)/(\delta x_i \delta x_j)$ are equivalent. Hence, the Hessian of I at position $\mathbf{x}$ and scale σ is a $D \times D$ matrix with elements

$$H_{ij}(\mathbf{x},\sigma) = \sigma^2 \mathbf{I}(\mathbf{x}) * \frac{\delta^2 G(\mathbf{x},\sigma)}{\delta x_i \delta x_j} \quad \text{for } i,j = 1,\ldots,D \tag{3.2}$$

where $G(\mathbf{x},\sigma)$ is a D-variate Gaussian and $*$ denotes the convolution operation. The factor σ^2 compensates the increasing blurring effect [80]. Let $\lambda_1,\ldots,\lambda_D$ denote the eigenvalues of H, such that $|\lambda_i| \leq |\lambda_{i+1}|$, and $\mathbf{e}_1,\ldots,\mathbf{e}_D$ be their associated eigenvectors. The eigenvector $\mathbf{e}_D$ points towards the direction in which the second derivative is maximum, and λ_D gives the value of the second derivative. The sign of λ_D is related to the contrast between the data element and the local neighbourhood, i.e. if it is brighter or darker than the neighbour regions. The eigen-analysis of H is capable of differentiating between tubular, rounded and plate-like shapes. Tables 3.1 and 3.2 show how local geometrical interpretations can be made according to the eigenvalues of H, for 2D and 3D data, respectively.

Table 3.1: 2D local shape inference based on the eigenvalue analysis of the Hessian matrix.

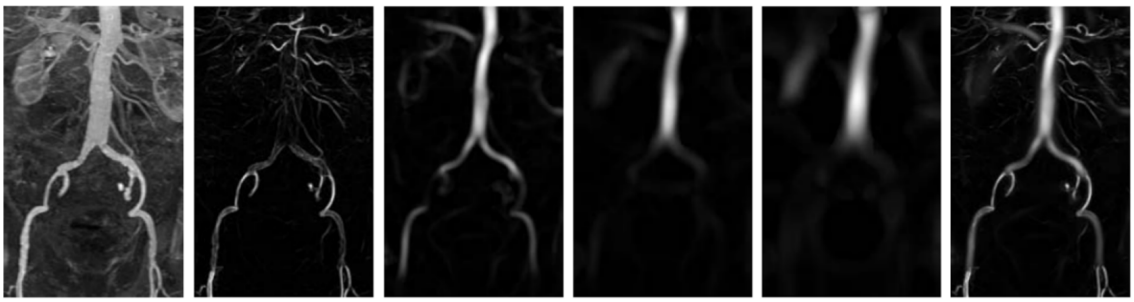
Shape		**Eigenvalues**
no shape		$\lvert\lambda_1,\lambda_2\rvert$ small
rounded	brighter	$\lvert\lambda_1,\lambda_2\rvert$ large; $\lambda_1,\lambda_2 < 0$
	darker	$\lvert\lambda_1,\lambda_2\rvert$ large; $\lambda_1,\lambda_2 > 0$
tubular	brighter	$\lvert\lambda_1\rvert$ small, $\lvert\lambda_2\rvert$ large; $\lambda_2 < 0$
	darker	$\lvert\lambda_1\rvert$ small, $\lvert\lambda_2\rvert$ large; $\lambda_2 > 0$

Table 3.2: 3D local shape inference based on the eigenvalue analysis of the Hessian matrix.

Shape		**Eigenvalues**
no shape		$\lvert\lambda_1,\lambda_2,\lambda_3\rvert$ small
rounded	brighter	$\lvert\lambda_1,\lambda_2,\lambda_3\rvert$ large; $\lambda_1,\lambda_2,\lambda_3 < 0$
	darker	$\lvert\lambda_1,\lambda_2,\lambda_3\rvert$ large; $\lambda_1,\lambda_2,\lambda_3 > 0$
tubular	brighter	$\lvert\lambda_1\rvert$ small, $\lvert\lambda_2,\lambda_3\rvert$ large; $\lambda_2,\lambda_3 < 0$
	darker	$\lvert\lambda_1\rvert$ small, $\lvert\lambda_2,\lambda_3\rvert$ large; $\lambda_2,\lambda_3 > 0$
planar	brighter	$\lvert\lambda_1,\lambda_2\rvert$ small, $\lvert\lambda_3\rvert$ large; $\lambda_3 < 0$
	darker	$\lvert\lambda_1,\lambda_2\rvert$ small, $\lvert\lambda_3\rvert$ large; $\lambda_3 > 0$

Given the possibility of inferring local shape through the analysis of the eigenvalues of H, it is possible to define metrics involving their values which aim at enhancing target structures, such as vessels. Lorenz et al. [73] and Sato et al. [74] were among the first to design vesselness filters using this approach. Nonetheless, only some of the eigenvalues were used in their metrics. Frangi et al. [75] proposed a vesselness measure comprising all the eigenvalues, such that terms responsible

for blob and plate-like structures suppression were included. The noise robustness of this metric made it very popular and it is still one of the most used techniques among Hessian-based filters. Figure 3.3 shows the enhancement of vessels at multiple scales achieved by this method.



(a) (b) (c) (d) (e) (f)

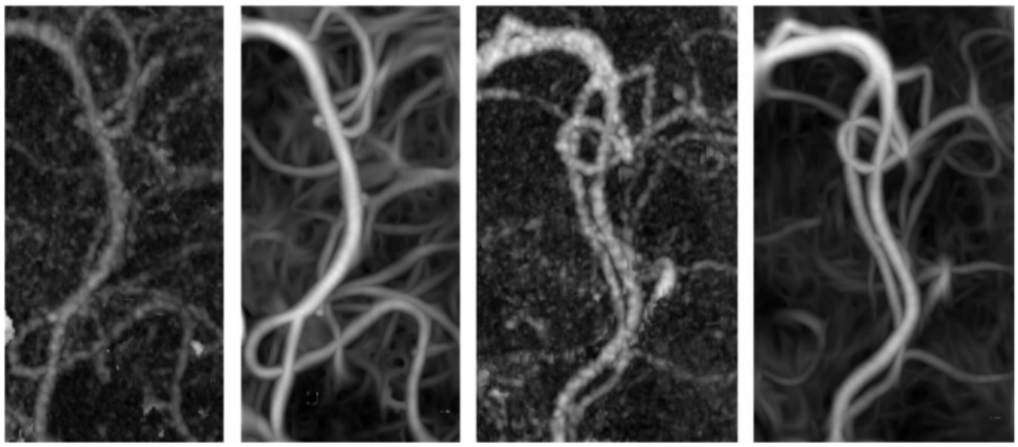
Figure 3.3: Vessel enhancement using Frangi's vesselness filter [75]: (a) maximum intensity projection of an MRA image; (b-e) vessel enhanced images at four increasing scales; and (f) the combined response.

Li et al. [77] developed a more user friendly filter, as it does not involve parameter tuning, as required in the approach of Frangi. Nonetheless, the lack of a structureness term does not give a strong noise robustness to this method. Zhou et al. [81] followed a different direction in terms of target structures to enhance. As bifurcations tend to locally resemble blob-like structures, the authors decided to also enhance this kind of shapes. This consequently enhances structures such as aneurysms and nodules. The applicability of such filter is very dependent on the application at hand, since enhancing blob-like structures may be strongly undesirable in scenarios where we do not want to include nodules and noisy regions which tend to look like small blobs. Recently, Jerman et al. [79] also shared this perspective, as the enhancement of aneurysms was intended.

Manniesing et al. [82] used a smoothed version of Frangi's vesselness to design a vesselness diffusion equation. The interesting property of such process lies in its anisotropy, such that diffusion is strong along vessels and strongly inhibited perpendicularly. This selective smoothing is able to strengthen the signal of vessels without significant distortion of edges, as can be seen in Figure 3.4.

3.1.3 Mathematical Morphology

Mathematical morphology theory analyses an image by means of non-linear order based operations. It relies on two basic building blocks, the dilation and erosion operators. They can be used to formulate more complex and interesting operators, such as closing, opening, top-hat and bottom-hat transforms. The geometrical properties embedded in mathematical morphology arise from the definition of a Structuring Element (SE) that dictates the influence range of the non-linearities. Thus, this theory is easily adapted to the assumptions about vessel structure. Even though initially mathematical morphology only addressed binary images, extension to grey-scale has been made.



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

Figure 3.4: Vessel diffusion in low dose CT scans of the cerebral vasculature: (a,c) regions of interest in the scans; and (b,d) the respective diffusion results after 40 iterations. Adapted from [82].

Low-level morphological operations were used by Thackray and Nelson [83] for the enhancement of vessel networks. The authors considered a morphological opening using an elongated SE. Varying vessel orientation was addressed by rotating the SE into eight different orientations, thus obtaining eight opened images. The enhanced result was obtained by the pixel-wise maximum response across the images. Zana and Klein [84] also considered morphological openings with an elongated SE, in order to obtain a sum of top-hat transforms along each direction for oriented vessel enhancement. However, such approach requires long SE in order to remove the wider vessels, leading to significant noise retrieval during the sum of top-hat transforms. The authors considered instead a geodesic reconstruction of the opened images into the original one. Likewise, Walter and Klein [85] calculated the top-hat transform from the supremum of openings using a large elongated SE in different directions. Mendonça and Campilho [86] also relied on mathematical morphology for vessel enhancement, however, a modified top-hat transformation was adopted to avoid the sensitivity to noise (see Figure 3.5). The authors further used mathematical morphology theory, given that binary morphological reconstruction was performed on images containing vessels enhanced at multiple scales, for application in a vessel filling method. Figueiredo and Leitão [87] estimated the boundaries of coronary vessel segments through a morphological edge operator. Smoothness constraints were considered for electing the most likely candidate estimates rather than directly smoothing them, in order to avoid an unwanted impact on the quantitative analysis of blood vessel stenosis.

Other complex operations were considered for vessel enhancement in the 3D setting, such as connected set filters [88; 89] and grey-level hit-or-miss transform [90; 91]. Dufour et al. [92] combined Hessian-based local shape description and spatially variant morphological closing to reduce noise and merge vessel segments affected by signal loss. The Hessian analysis provides a fast mechanism to evaluate local shape and obtain principal curvature direction, while the mor-

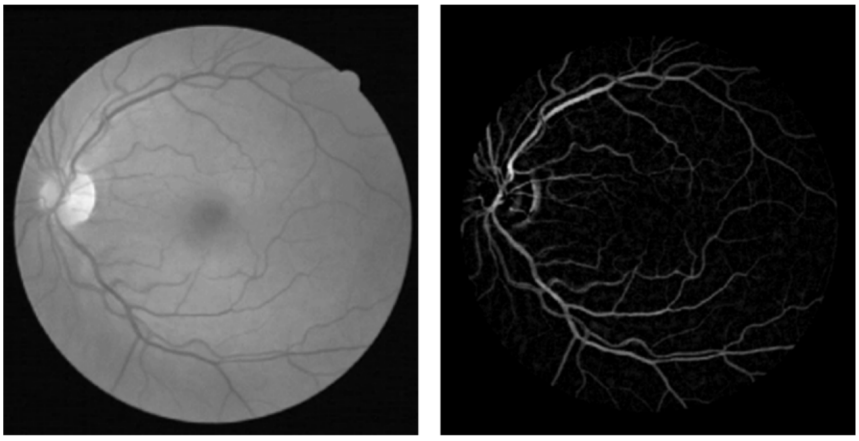


Figure 3.5: Mathematical morphology based enhancement of vessels applied by Mendonça and Campilho [86].

phological reconnection allows a fast local vessel signal recovery, as opposed to the approach of Manniesing et al. [82].

3.1.4 Centreline Tracking

Centreline tracking methodologies iteratively find new points along a vessel according to some local feature. Even though such approach only extracts the centrelines of vessels, this type of algorithm commonly incorporates a mechanism to estimate their local diameter, allowing to produce a complete segmentation of the lumen. The detection of bifurcations is a special case that has to be addressed in order to automatically track vessel trees.

In one of the pioneer works, Sun [93] developed a tracking method for finding the vessel contours in digital coronary arteriograms (see Figure 3.6). Each centreline point is represented by its position, $\mathbf{p}_k$, local vessel direction, $\hat{\mathbf{u}}_k$, and local width, w_k. Given such a centreline point, the next one is first estimated by $\tilde{\mathbf{p}}_{k+d} = \mathbf{p}_k + d \cdot \hat{\mathbf{u}}_k$, where d is proportional to w_k. A profile of length $2w_k + 1$ perpendicular to $\hat{\mathbf{u}}_k$ and centred at $\tilde{\mathbf{p}}_{k+d}$ is then taken. A matched filtering process is performed with a rectangular kernel, and the maximum response is regarded as a new estimate of the centreline point, $\mathbf{p}'_{k+d}$. A new profile centred at $\mathbf{p}'_{k+d}$ and perpendicular to the vector defined by locations $\mathbf{p}_k$ and $\mathbf{p}'_{k+d}$ is used to detect vessel edges through the analysis of roll off points. Finally, $\mathbf{p}_{k+d}$ is set to the midpoint between the edges and w_{k+d} is updated accordingly. This process is iteratively conducted until a stopping criterion involving the lack of local contrast is met.

Zhou et al. [94] modelled the vessel cross section profile as a Gaussian instead. Additionally, the authors measured local vessel diameter by taking the 95% confidence interval of the matched filter response. Chutatape et al. [67] dedicated their work to the detection of vessels in ocular fundus images. As the optic disk in these images is the source of vessels, an automated initialisation step was proposed. The second-order derivative of the Gaussian is convolved with the intensity profile of a circumference around that structure, such that local maxima of the response

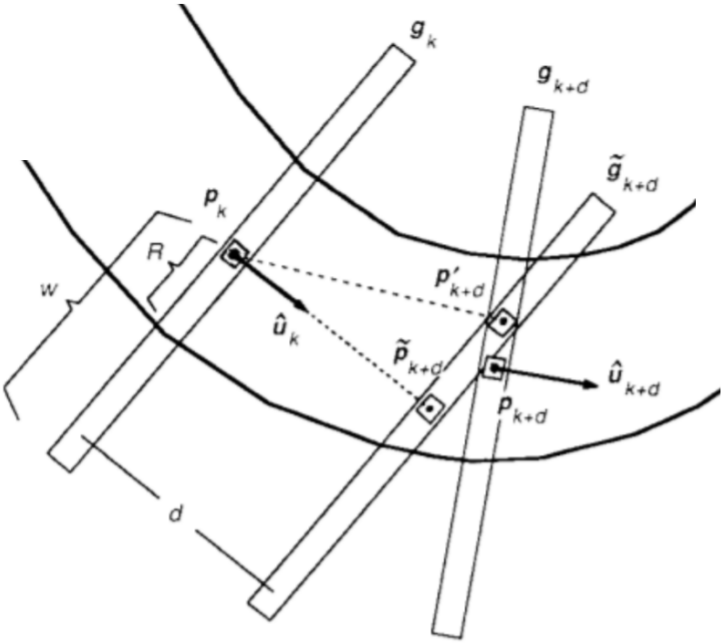
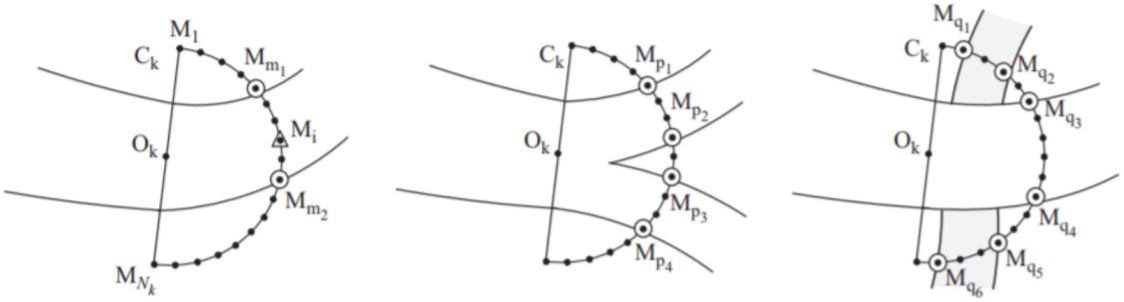


Figure 3.6: Representation of the tracking approach of Sun [93].

are potential sites for starting the vessel tracking procedure. For each tracked vessel, a Kalman filter [95] is responsible for estimating new centrelines, which are corrected by analysing the cross sectional profiles. While the previous described works were not able to detect bifurcations, here, at each centreline point, a scheme based on the convolution of a Gaussian filter to the locations lying in a front half of such point is proposed. Tolias and Panas [96] also addressed the automatic retrieval of locations to initialise the tracking procedure in ocular fundus images. The authors used a Fuzzy C-Means (FCM) clustering method on the intensity profile of the circumference around the optic disk. At the iterative tracking step, a vessel membership function is evaluated through the FCM algorithm applied to the cross section profile, and the edge points are set as the locations where such membership is 0.5. Bifurcation detection and tracking termination are accomplished by heuristic-based decisions. Wörz and Rohr [97] designed a tracking model for 3D data, where new centreline points were estimated through fitting a cylindrical model to the data.

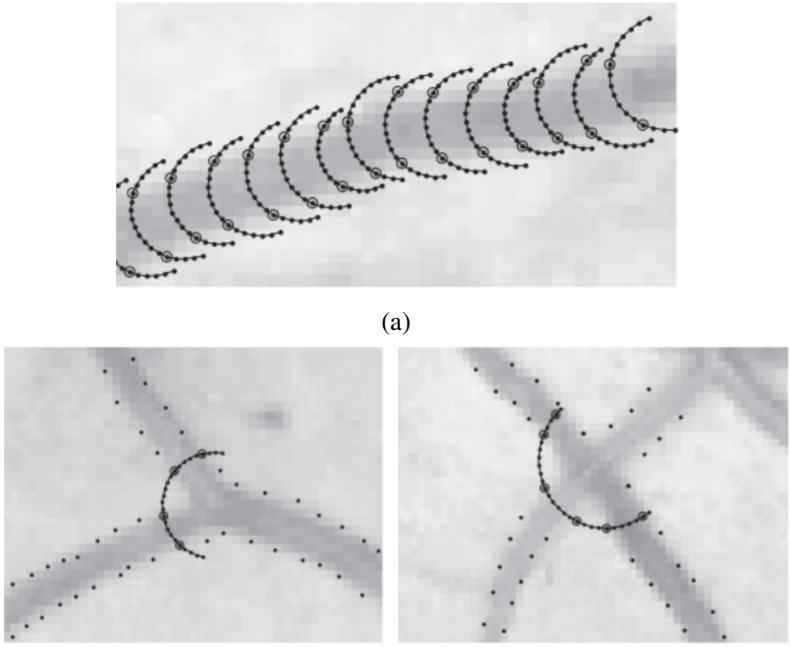
Friman et al. [69] target the tracking of small vessels, where frequently the signal is disguised in the background noise. The authors state that modelling the profile of the cross section of such vessels as a Gaussian is not proper, designing a steeper profile instead. Tracking vessels in low contrast regions becomes possible with the introduction of a multiple hypothetical vessel path approach. Given a centreline point, a set of pre-defined locations for the next one is considered and the vessel template is used to assess each solution. Based on the pattern of the scores, a new estimate is made.

Yin et al. [98] proposed a probabilistic framework for iteratively detecting edge points along vessel structures. At each iteration, a semi-ellipse is centred at the current centreline point according to the local vessel direction estimate, as represented in Figure 3.7. A set of points is drawn



(a) (b) (c)

Figure 3.7: The three different types of configurations considered in the tracking approach of Yin: (a) normal; (b) bifurcation; and (c) crossing. The points marked with circles are used to define a possible configuration. At each step, all 2, 4, and 6-combination of points are used. Adapted from [98].



(a)

(b) (c)

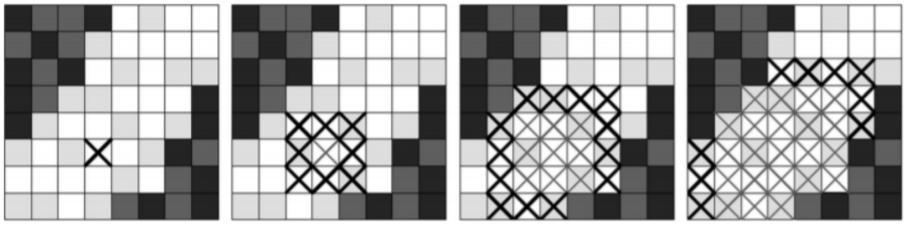
Figure 3.8: Edge points tracked by the algorithm of Yin [98] at different vessel patterns: (a) single vessel, (b) bifurcation, and (c) crossover.

from the semi-ellipse to obtain a vector of intensities which is used to infer about local structure, i.e. if a single segment, a bifurcation or a crossover lies ahead. The estimated edge points are the ones maximising the posterior probability of a configuration given the intensity vector. The likelihood models assume that the background has constant intensity and the cross section of vessels is Gaussian-like. The prior probability promotes edge continuity. Figure 3.8 presents results obtained at single vessel, bifurcation and crossover regions.

Bekkers et al. [99], inspired by the mechanisms of human visual cortex, developed a tracking algorithm based on orientation scores. One of the experiments considered the use of cake wavelets to obtain an invertible orientation score. This property allows to disentangle crossing structures, such that accurate tracking of blood vessels is possible in these locations, which are commonly problematic.

3.1.5 Region Growing

Region growing is an iterative scheme for image segmentation, requiring an initial set of mask pixels, known as seed regions. These regions may be provided by an user or automatically induced from image information. If automatic initialisation is considered, the algorithm must be robust and return adequate locations, or else region growing will fail. At each iteration of the procedure, seed-region neighbouring pixels are candidates to join that region. They are merged if they meet all the specified criteria, which are generally based on intensity information (see Figure 3.9). The simplicity of these methods makes them computationally efficient, although the obtained result depends on the initial seed point(s) location. To overcome this initialisation problem, Wan and Higgins [100] proposed a symmetric region growing algorithm capable of extracting the same vessel tree for different starting seed points.



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

Figure 3.9: Region growing algorithm using intensity related growth criterion: (a) the initial seed (dark cross); (b-d) candidate pixels (dark crosses) and expanded seed region (light crosses) as the algorithm progresses. Adapted from [101].

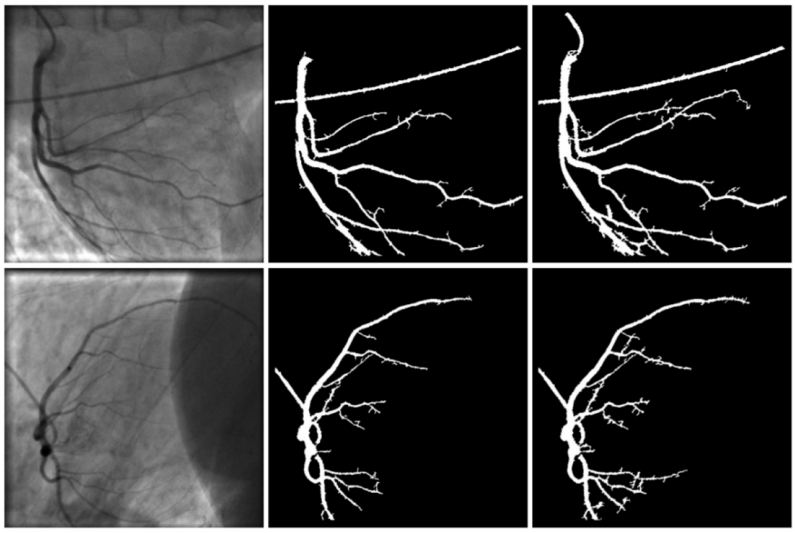
Martínez-Pérez et al. [102] employed a two-stage region growing algorithm based on features related to the scale-space analysis of the first and second order derivatives of the intensities. A set of seed regions was obtained by analysing the histograms of the extracted features. In the first stage, growing occurs in regions of low gradient magnitude, while the second stage relaxes such constraint in order to obtain more accurate vessel edges. The authors extended their work in [103] and a parallel implementation was later made [104].

Boskamp et al. [101] combined three intensity-based criteria for growing the seeds: a low threshold, such that no candidate with a lower intensity is included in the segmented region; an adaptive high threshold, combined with size limit, such that calcifications may be included; and a gradient-based threshold, enforcing that only neighbours with similar intensity are accepted. The authors state that setting the low threshold is not a trivial task. Hence, an iterative framework is

used, where the low threshold is decreased at each iteration, and the initial seed regions are the locations comprising the segmentation obtained at the previous iteration.

Metz et al. [105] developed a region growing methodology accounting for bifurcation and leakage detection, two essential steps for the extraction of vessel networks. The growing criterion is very simple, as it only regards if the intensity of a candidate pixel lies within a certain range. Bifurcation and leakage detection are achieved by analysing the behaviour of the grown regions in recent iterations. Concerning bifurcation detection, if the recently grown regions are not connected, then a bifurcation exists. Regarding leakage, if the number of voxels that were recently appended to the segmented region surpasses a specified threshold, then leakage is assumed to have occurred.

Li et al. [106] obtained two feature maps from the images, one of them used for extraction of wider vessels, while the other is tailored for addressing narrower ones. A novel vesselness measure is used by the authors to obtain a multi-scale vessel enhanced image. The seed is initialised at the location where the response to the enhancement is higher. A threshold is obtained using the Otsu's method [107] over the histogram of the enhanced image. A region growing approach follows where pixels are appended to the seed region if their response in the enhanced image is higher than the threshold. As this does not include narrower vessels in the segmentation, the multi-scale first-order derivative of the intensities is considered as a second feature map. Region growing using the previous segmentation as seed region and the new feature map response allows to obtain a more accurate result (see Figure 3.10).



(a) (b) (c)

Figure 3.10: Outputs of the region growing approach of Li et al. [106]: (a) original angiograms; (b) wide vessel extraction; and (c) refined segmentations.

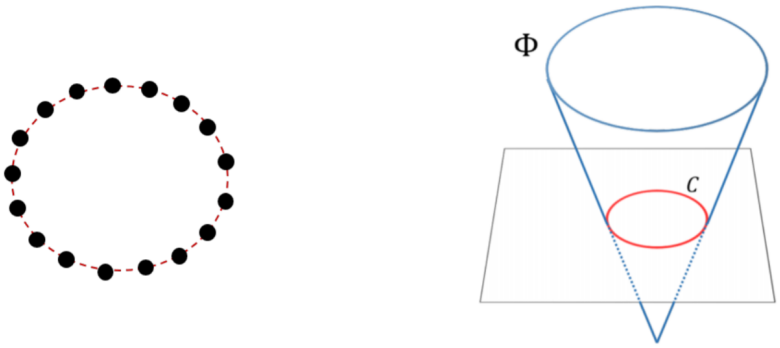
Zhao et al. [108] combined two segmentation procedures into a final output. One of them was obtained by applying region growing to an image enhanced by the 2D Gabor wavelet. Seed

initialisation is automated via analysis of the histogram of the enhanced image. A simple intensity similarity criterion is used in the growing scheme, however it is parametrised by a value induced from each specific image.

3.1.6 Active Contour Models

Active Contour Models (ACMs) were introduced by Kass et al. [109], and quickly became one of the most used methodologies for image segmentation. An ACM is a curve moving in an image (or a surface in a volume), in order to minimise an energy function that accounts for its topology, forces derived from the image, and possibly, user defined attractive and/or repulsive forces. The first accounts for forces that are inherent to the shape of the ACM, in order to resist stretching and bending efforts. The second is responsible for driving the contour into the features of interest, such as the edges of an object. The last is an user dependent mechanism that allows to induce attracting and/or repelling forces, guiding the contour to the desired configuration.

The curve may be represented explicitly [109] or implicitly [110; 111] (see Figure 3.11). An



(a) (b)

Figure 3.11: Possible representations of a curve: (a) explicit; and (b) implicit.

explicit representation regards the curve as a finite set of points moving in the image domain according to a field of forces. The implicit representation typically uses the level set method [112] to represent the curve as the zero-level set of a higher dimensional function. The advantages of this representation are the natural handling of topological changes and the trivial extension to higher dimensional problems. Nevertheless, level set based approaches require periodical re-initialisations of the level set function to a signed distance one, in order to maintain stability throughout the evolution of the contour. Besides being difficult to know when the re-initialisation should be made, a significant computational cost is also introduced. The need of re-initialisation was eliminated by imposing the level set function to be similar to the signed distance function, which was achieved with the introduction of a penalty term that weights the distance between those functions [113].

Regarding the nature of the image features guiding the ACM to the object boundary, edge-based snakes are attracted to strong gradient regions while region-based snakes move along the image according to region statistics, looking for the configuration that maximises the difference of such features in each side of the contour. The latter are able to detect boundaries of objects which are not defined by gradient, and tend to be more robust when considering the influence of contour initialisation in the evolution process.

Sum and Cheung [114], aware that the model of Chan and Vese [111] is inadequate for non-uniform illumination, proposed a two-step procedure for the extraction of vessels with a level set based active contour. In the first step, the Chan-Vese model is used to obtain an initial segmentation, prone to bad results in regions where vessels have intensities similar to the background. A refined result is produced by performing more iterations of the level set method using a locally normalised version of the image. Sun et al. [115] also considered the extension of the Chan-Vese model to the vessel segmentation problem, by modifying the region-based terms in the energy function. The authors replaced the mean region intensities by the maximum and minimum fuzzy opening operators.

Al-Diri et al. [116] designed the Ribbon of Twins parametric model, comprising two pairs of open ACMs. One of the pairs is intended to fit the interior edges of the vessel wall, while the second pair should fit the exterior ones. The coherence of the model is kept by imposing repulsive forces between the interior ACMs and attractive forces between ACMs modeling the same vessel wall. This is achieved with the introduction of an additional term in the model energy, penalising the difference between the expected vessel width and the distance between the interior ACMs.

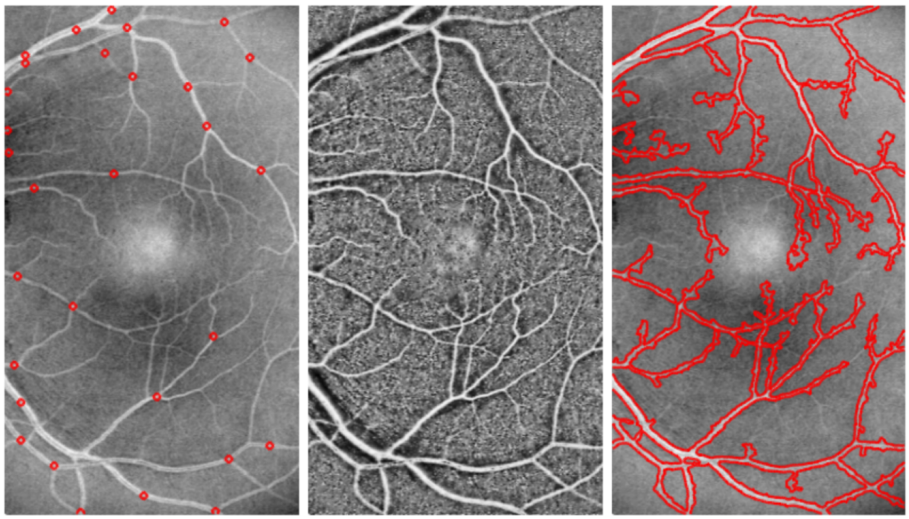
Läthén et al. [117] used a geodesic ACM in order to introduce regularisation into the segmentation of vessels in images enhanced by quadrature filtering. The curve length penalising term included in that model proved to be crucial for increased noise robustness. Example results are shown in Figure 3.12.

Shang et al. [118] proposed a region competition-based ACM making use of a Gaussian Mixture Model (GMM). As such design is not capable of extracting the narrower vessels, the authors added a term to the level set motion equation that is responsible for promoting the evolution of the contour into the thin vessel region from its central line. The term comprises a vector field that is derived in the local vessel direction and whose magnitude comes from a vesselness measure.

Zhao et al. [119] extended the infinite perimeter ACM [120] in order to combine different sources of information, namely intensity and vesselness information. The latter is obtained by analysing the local phase, as in [117]. Additionally, the regulariser is modified, aiming for improved detection of small oscillatory structures, such as bifurcations.

3.1.7 Graph Cuts

Graph Cuts is an algorithm that employs graph theory to segment images by minimising an energy. In this framework, an image is taken as a grid of interconnected nodes (4 or 8 neighbourhood for example). There are two additional nodes with special properties, the terminal nodes, namely one



(a) (b) (c)

Figure 3.12: Active contour segmentation by Läthén et al. [117]: (a) portion of a retinal fundus image with initial zero-level set in red; (b) phase map; and (c) final segmentation.

source and one sink. These nodes represent the foreground and background objects. Figure 3.13 shows a representation of the graph that is built.

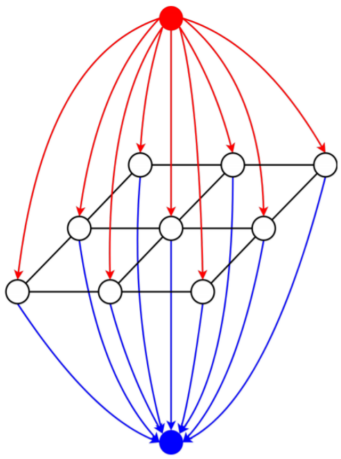


Figure 3.13: Representation of the graph implemented in a Graph Cuts approach for binary segmentation. The red and blue nodes represent the source and sink nodes, respectively.

The underlying idea is that the non-negative weights of edges between pixel nodes denote how difficult it is to separate those pixels (for example based on intensity similarity), while the non-negative weights between terminal nodes and pixel nodes incorporate how likely that pixel belongs to the foreground or background (according to some extracted feature and assuming that the feature densities for the background and foreground are known).

A valid cut on this graph is a partition of the nodes into two disjoint sets such that the terminal

nodes are separated. The cost of a cut is the sum of the edge weights that were severed by it. The minimum cut problem finds the cut having the lowest cost, usually by solving the equivalent maximum flow problem through algorithms such as the Ford-Fulkerson method [121].

Slabaugh and Unal [122] segmented images using Graph Cuts approach, but introduced a shape prior which further constrains the segmentation result, leading to increased noise and weak boundary robustness. The elliptical prior was used being adequate to extract structures such as vessels.

Schaap et al. [123] used Graph Cuts to extract the vessel lumen given its centreline. An estimate of local lumen intensity was made using the intensities of voxels belonging to the centreline. The data term was set according to two differences: one between the local lumen estimate and the voxel intensity, and another between the estimated lumen and surrounding tissue intensities. The boundary term between neighbour voxels was set to the commonly used Gaussian function of the squared intensity difference.

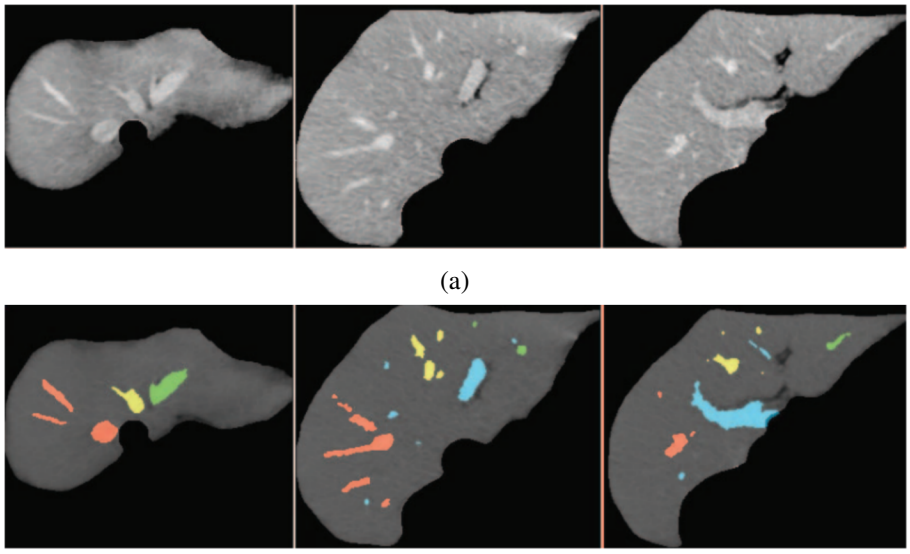
Esneault et al. [124] also considered a model-based constraint into the Graph Cuts framework. Motivated by its integrative nature, and consequent increased noise robustness, the authors used a geometrical moment-based detector of cylinder shapes. This detector was used at multiple scales to retrieve a set of cylinder parameters in the 3D space, which influence the Graph Cuts segmentation procedure, by means of a modified energy minimisation function.

Pamulapati et al. [125] used data coming from two distinct temporal phases of the hepatic vessels image acquisition procedure: the non-contrast phase and the portal venous enhancement phase. Intensity histograms at vessel and background regions were obtained after finding an optimal threshold in the enhanced data. The energy minimisation was modified in order to acquaint for three different region-based terms. A first term computes penalties according to the voxel intensity and the intensity histograms of the foreground and background. A second term associates penalties according to the voxel intensity difference in both data sources, as a vessel voxel should have a significant higher intensity in the enhanced data. Finally, a third term incorporates the multi-scale vesselness measure developed by Sato [74]. Figure 3.14 contains examples of hepatic vessel segmentation using this methodology.

Zhao et al. [126] optimised the energy function proposed in the active contour without edges model of Chan and Vese [111] by means of a Graph Cuts approach. The region term was deviated from a local-phase based enhanced vessel image, after correction of intensity inhomogeneity by a method inspired in the Retinex theory [127].

3.2 Label-driven approaches

Label-driven approaches rely on the availability of data and corresponding annotations to find complex mappings between a set of hand-crafted features and pixel-wise labels (traditional ML) or to automatically find a hierarchy of deep representations of an image that is useful to solve the segmentation problem (DL). The amount of required data varies with the complexity of the model



(a)

(b)

Figure 3.14: Extraction and labelling of hepatic veins on axial slices of liver CT data: (a) portal venous enhancement phase images; and (b) vessel network extraction and labelling. Adapted from [125].

and the degree of supervision. This is true for ML but mostly for DL. The ground-breaking performance levels achieved by DL in many applications were only possible due to the availability of huge amounts of labelled data. Nonetheless, there have been efforts to reduce the data eagerness of these methodologies, through semi- and weak-supervision. Semi-supervised methods also use unlabelled portions of the data during training, usually by including auxiliary tasks like image reconstruction, or classification [128]. Even though unlabelled images do not directly contribute to the segmentation loss, they end contributing to the learned feature spaces. Weakly-supervised methods use less complex and easier to obtain annotations, which still comprise relevant information for the segmentation problem [129].

Transfer learning has also been allowing to employ DL in datasets with little annotated data. Models trained with large amounts of data are likely to learn low-level features which are relevant for a broad array of scenarios. Thus, it may suffice to fine-tune a previously learned model on the target data to obtain good performance [130]. A particular instance of transfer learning known as domain adaptation is an active topic of research by the ML community [131; 132]. Here, there are different datasets of related tasks available, and there is an interest in deriving feature representations from labelled datasets (known as sources) which are useful in the unlabelled ones (known as targets). This topic is of utmost importance in practice, since it would allow to effectively learn models in labelled datasets which are already available and use them in new data acquisitions which probably have different distributions, due to new acquisition procedures and/or protocols.

In this Section, the main traditional ML and DL approaches that have been proposed for the blood vessel segmentation task will be described.

3.2.1 Traditional machine learning

ML has a great impact currently, as the extraction of knowledge from data is nowadays a common procedure in almost every practical field. Computer vision is no exception, and the particular case of vessel segmentation was also focused by this type of techniques. Here, supervised learning is focused, where a training dataset comprising labelled examples is exploited. This allows to use a myriad of learning techniques to create decision boundaries that separate the classes according to the extracted features. In the vessel segmentation literature, ML techniques have been mainly applied to the extraction of the vasculature from retinal fundus images, since labelled databases exist for such scenario [38; 39; 40; 133].

Niemeijer et al. [134] used a k-Nearest Neighbours (kNN) classifier with k = 30 to distinguish between vessel and non-vessel pixels, based on a feature vector of 31 dimensions. Besides the pixel intensity in the green channel, the features were obtained by performing a convolution between the image and the following filters at 5 different scales: the Gaussian kernel and its first and second derivatives (in each axis direction).

Staal et al. [38] extracted potential vessel locations by analysing the ridges of the image at different scales. A ridge is a location where an intensity extremum exists along the direction of the largest surface curvature, i.e. the direction given by the eigenvector corresponding to the eigenvalue of highest magnitude of the Hessian. Theoretically, such extrema should exist near the centrelines of the vessels. The found extrema were used to compose line primitives and assign pixels of the image to the closest primitive. Afterwards, a set of appearance and geometrical features were extracted, including features related to the perpendicular profiles of the primitives. Feature selection was implemented by means of sequential forward selection and a kNN classifier was used to perform soft classification. By varying the threshold applied to the soft output, the authors were able to obtain a Receiver Operating Characteristic (ROC) curve.

Soares et al. [135] considered a smaller number of features, as only the inverted green channel intensity and the 2D Gabor wavelet transform responses at the considered scales were used. Such wavelet was selected due to its selective enhancement of oriented features and the capacity to choose the specific frequencies to find. Hence, for each scale, the maximum modulus of the wavelet transform over all the angles is taken as feature. Using the training data, the authors considered GMMs to find the distribution of the features over the vessel and non-vessel classes. A Bayesian approach was then employed to classify test feature vectors into one of the classes. The authors obtained a ROC curve by varying the threshold on the posterior pixel probabilities. The experiments showed that using k = 20 clusters in the likelihood learning phase lead to the best performance.

Ricci and Perfetti [136] found line-like profiles by using a line operator of length equal to 15 pixels and rotating it over all angles with a step of 15^{o} . Considering vessels brighter than the background, the intuition is that for some angle, the line operator fits only vessel pixels and the sum of intensities over the line operator is larger than at other locations. Thus, one feature is the difference between such sum (the maximum over all angles) and the average intensity at

a 15×15 neighbourhood. Since this also outputs rather high values at regions near vessels, a second line operator of length equal to 3 pixels and orthogonal to the first operator is regarded (see Figure 3.15). A feature from this operator is extracted in the same way as described before. Additionally, the inverted green channel intensity is added to the feature vector, leading to a 3-element predictor for each pixel. A linear Support Vector Machine (SVM) was considered for the classification task, and the bias was varied in order to generate a ROC curve.

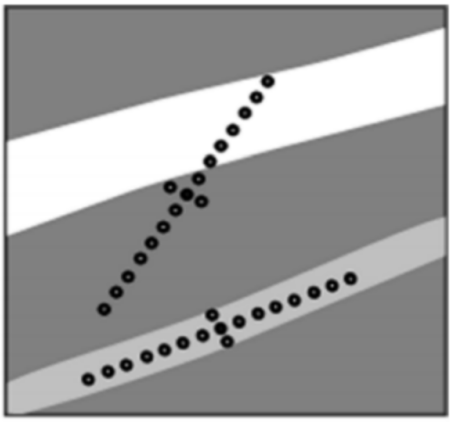
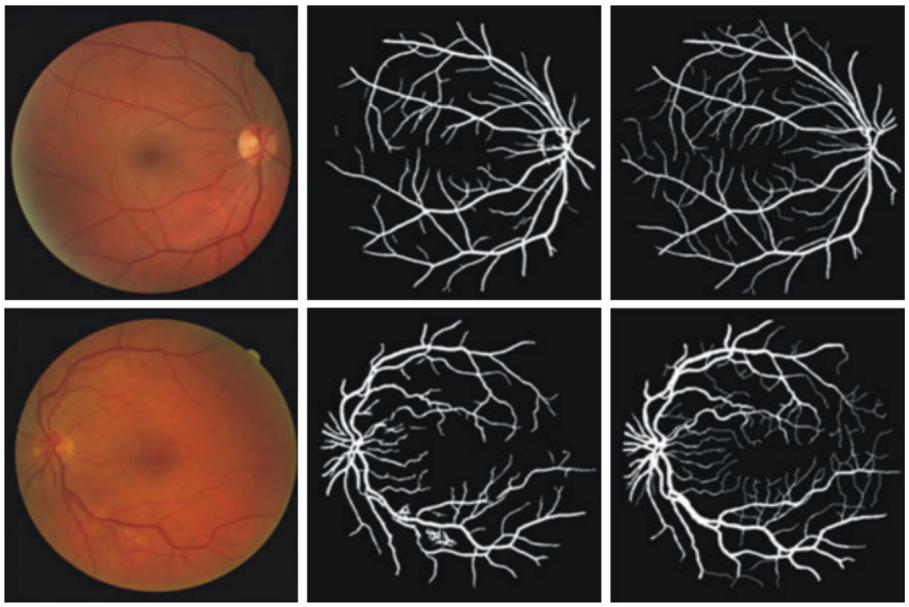


Figure 3.15: Line operators considered by Ricci and Perfetti [136].

Lupascu et al. [137] extracted a significant number of features related to local intensity structure, spatial and geometrical properties. To deal with vessels of different size, some of them were extracted at four different scales, leading to predictors with 41 dimensions. An Adaboost classifier was designed, where a decision is the result of a linear combination of the outputs coming from simple classifiers, usually known as weak learners. In this approach, a first weak learner finds the thresholds that lead to a smaller loss. Then, at each iteration, a new weak learner is introduced, with the detail that the weight of misclassifying a given training sample is larger if it was wrongly classified by the previous learner. The authors considered 100 iterations in their work. Changing the threshold of the final output allows to find the ROC curve of the model.

Marín et al. [138] represented each pixel as feature vector with 7 dimensions, comprising information related to grey-level and moment invariants. The features were extracted after intensity homogenisation and vessel enhancement by means of a top-hat transform. The classifier designed by the authors was a multilayer feed-forward neural network with an input layer (7 neurons), three hidden layers (15 neurons in each) and an output layer (a single neuron). The non-linear logistic sigmoid function was regarded as the activation function of all the neurons, such that the network output ranges between 0 and 1. The authors interpreted such value as a posterior probability and a ROC curve was obtained by varying the decision threshold. The training set was not randomly obtained as occurred in other works, in order to avoid possible reference errors. Instead, the authors carefully selected training samples from the different patterns: possible vessel, background, and noise.

Fraz et al. [139] considered 9 features: the intensity in the inverted green channel; one deviated from the analysis of the orientation of the gradient vector field; another obtained with a morphological transformation; two from line strength measures; and, finally, the remaining four resulting from the Gabor filter response at multiple scales. Similarly to Lupascu et al. [137], the authors considered ensemble-based classifiers. Besides AdaBoostM1 and LogitBoost (which differ in the loss function, but their intuition is the same and was explained before), the authors also considered a bootstrap aggregation strategy, where the weak learners are independently and simultaneously trained, without changing the weights of misclassified samples. The authors used 200 decision trees as weak classifiers and their experiments concluded that the LogitBoost was the strategy leading to better performance (two results are compared to the ground truth in Figure 3.16).



(a) (b) (c)

Figure 3.16: Comparison between the segmentations obtained by Fraz et al. [139] and the ground truth: (a) original retinal fundus images; (b) segmentations; and (c) ground truth.

Roychowdhuri et al. [140] start by extracting the major blood vessels of the retina using two different approaches, one resorting to a high-pass filter and the other to a top-hat reconstruction. The intersection of segmentations is deemed as the preliminary detection of the wider vessels. Neighbourhood-based features as considered in [138] and gradient-based descriptors are extracted from pixels which are not included in the preliminary segmentation. A GMM classifier is then used to distinguish between narrow vessels and the background. The authors argue that this design helps the classifier to focus on the challenging part of the vessel detection task - distinguishing narrow vessels from background, since the major vessels are usually easily extracted.

Strisciuglio et al. [141] considered a supervised approach resorting to the B-COSFIRE descriptor [68]. A bank of 42 filters (21 for vessel segments and 21 for vessel endings) is initially built. Pixel-wise feature vectors containing the local response to each of the 42 filters, and also the

green-channel intensity value, are obtained. Several experiments are performed concerning feature selection, from criteria based on the entropy to genetic algorithms. Finally, a SVM classifier with a linear kernel is used to classify each pixel into vessel or background.

Orlando et al. [142] posed the vessel segmentation problem as minimising the energy of a fully connected Conditional Random Field (CRF). In contrary to traditional CRFs, whose pairwise potentials only encode local information, fully connected CRFs take into account long-range interactions between pixels. This comes at the cost of increased computational complexity during the inference step, however, it is alleviated by restricting the pairwise potentials to linear combinations of Gaussian kernels over an Euclidean feature space [143]. The authors jointly learn the weights for the unary, bias and pairwise terms, by applying the Structured Output SVM [144]. The considered features are those referred in other works [84; 135; 145].

Zhang et al. [146] consider both full-scale and scale-selective wavelet transforms to obtain orientation score representations of the images. In addition to features taken from these representations, the authors also extract Gaussian scale-space features, up to a total of 29 features per pixel. For the classification step, a Random Forest containing 500 trees was used.

Recently, Wang et al. [147] proposed a complete pipeline for the segmentation of retinal vessels, including pre- and post-processing. Regarding the pre-processing stage, in addition to the typical background normalisation, the authors have used an image-detail preserving filter [148] for removing isolated noise. Afterwards, a large number of features was collected, following the descriptors used in other works (matched filters, Gabor wavelets, local grey-level statistics, Frangi's vesselness features, and difference of Gaussians). To remove redundant features and being aware of the imbalance between vessel and background pixels, the authors perform an asymmetric principal component analysis, reducing 300 features to 100 descriptors. A cascade of classifiers was considered, where each consecutive classifier assigns a label to pixels in uncertain regions. Post-processing removes structures whose geometric properties do not follow the structure of vascular networks.

3.2.2 Deep learning

DL is currently a massive trend in the field of computer vision, having achieved top performance in tasks such as hand digit and face recognition. As the term suggests, deep networks comprise a large number of neuron layers which allow to encode complex and non-trivial non-linear functions of the data. In comparison to traditional ML, where the task of feature engineering is fundamental, DL is able to automatically find adequate representations of the data for a particular problem, given sufficient and relevant data. Much of the attention was due to the performance achieved by Convolutional Neural Networks (CNNs) in visual challenges. A CNN comprises sequences of convolutional, pooling, and activation layers, which are responsible for extracting increasingly high-level descriptors of the data, and a final set of fully connected neurons, which learn complex non-linear functions of the learned features in order to solve a particular visual task. Convolutional layers convolve the image with a set of filters, producing a new feature layer for each filter. Pooling layers reduce the size of the images, in order to decrease the number of parameters to learn, to

increase the robustness to small translations of the features in the original image, and to capture information from a larger neighbourhood. Activation functions allow to learn non-linear representations of the data. These networks were already a reality at the 1990s decade, when LeCun et al. [149] designed LeNet-5 (see Figure 3.17) and applied it to the task of character recognition. However, the hardware limitations at the time hindered their true potential. The trend re-emerged

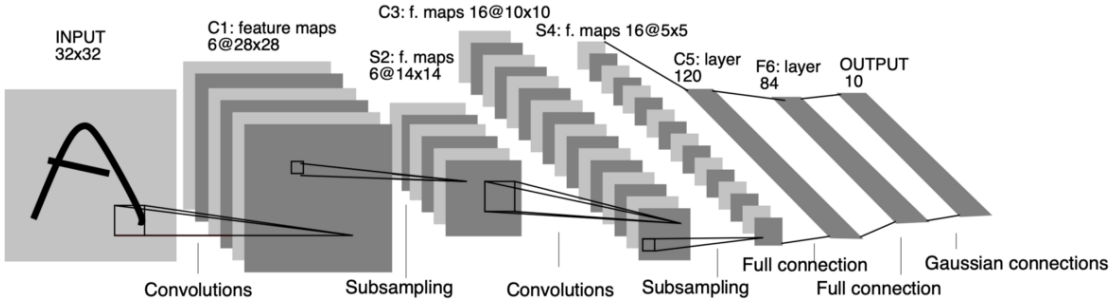


Figure 3.17: The LeNet-5 CNN applied in digit recognition. The final output layer represents a non-normalised distribution over each class. Adapted from [149].

after Krizhevsky et al. [150] significantly outperformed other competitors in the ImageNet classification problem. This was achieved using a CNN designated as AlexNet (see Figure 3.18) and possible due to the improvement hardware had been facing and the introduction of parallel processing. The continuous technological evolution and increasing amount of available data has been allowing the design of networks with larger complexity and increased performance.

Melinščak et al. [152] were among the first to apply DL for the segmentation of blood vessels. The authors resorted to an architecture containing 4 blocks of convolutional, max-pooling and Rectified Linear Unit (ReLU) activation layers, followed by 2 fully connected (FC) ones that output the probability of each class. The use of FC layers turns mandatory the division of the image into small tiles (patches) of equal size, and the problem of segmentation is converted into a multiple patch-classification one, where the network outputs the probability of the centre pixel of the patch belonging to each class. Liskowski and Krawiec [153] considered a similar approach. Among their experiments, they have tested the impact of data augmentation, balancing the training samples, removing pooling layers, and considered structured prediction, i.e. instead of predicting just the label of the centre pixel of the patch, labels are assigned to a neighbouring window centred at that pixel (see Figure 3.19). In their experiments, pooling was often prejudicial. The most promising performances were obtained in the balanced data setting and in the one without pooling layers, both considering structured prediction.

Li et al. [154] pose the problem of vessel segmentation as a cross-modality data transformation one. The authors divide images into 16×16 patches and parametrise the mapping function with a 5-layer artificial neural network comprising, respectively, 256, 400, 400, 400, and 256 neurons. This approach differs from the ones above, as classes are predicted for each pixel of the patch.

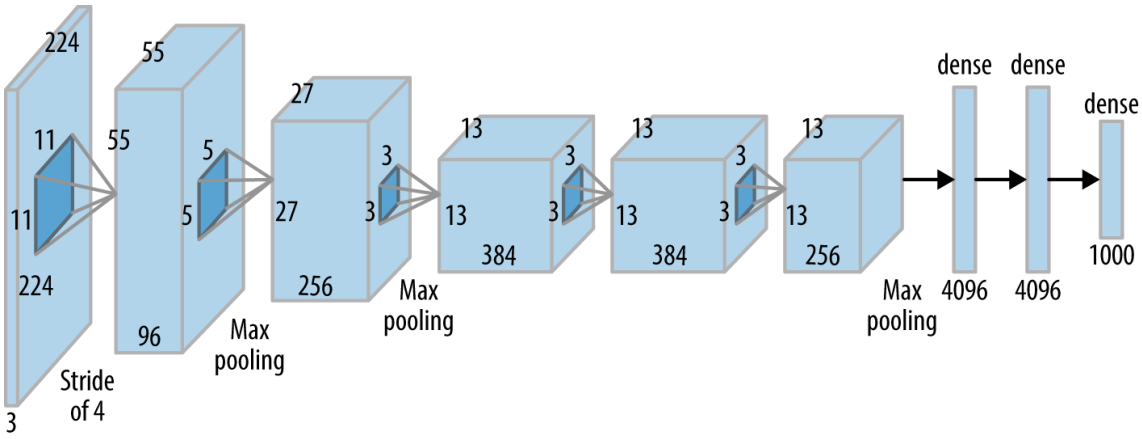
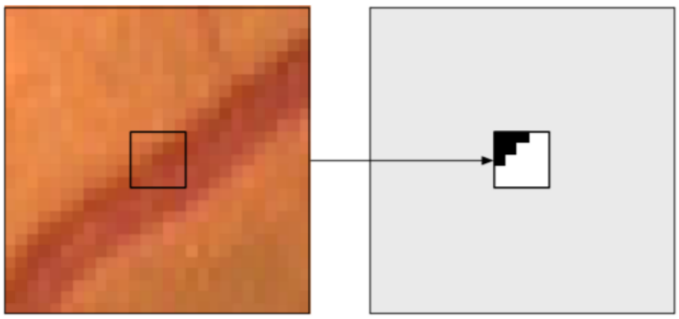


Figure 3.18: The AlexNet CNN applied in the ImageNet competition. A softmax function follows the last fully connected layer in order to obtain a probability for each class. Adapted from [151].



(a) (b)

Figure 3.19: A training sample in the structured prediction approach: (a) 27 × 27 patch with a 5×5 output window; and (b) the corresponding ground-truth. Adapted from [153].

The authors state that it is difficult to train this network from scratch using the backward propagation algorithm, therefore the first layer is pre-trained using an approach based on denoising auto-encoders.

Fully Convolutional Networks (FCNs) [155] were an important milestone in semantic segmentation applications, since they have enabled obtaining dense predictions efficiently. They replace FC layers with convolutional counterparts. Dasgupta and Singh [156] followed the structure prediction approach of Liskowski and Krawiec [153] but resorted to a FCN instead. As discussed by Long et al. [155], training with mini-patches of a given image instead of feeding it into the network, corresponds to sampling the loss function. Many authors resort to patch-based training due to memory limitations of processing units and to increase the randomness of data feeding the models. Feng et al. [157] considered this when training a FCN, however the sampling of patches was not entirely random. Since the vessel class is scarcer, the authors provided a mechanism to guarantee that at least a portion of the patches included regions where both classes were well

represented. For that end, the label entropy of the patches was assessed, given that a higher entropy indicated a larger equilibrium between class representation. The authors considered skip connections too, as introduced in U-Net [23], a state-of-the-art architecture for the segmentation of biomedical images.

The U-Net has been successfully applied to blood vessel segmentation in several works [158; 159] and its original architecture is shown in Figure 3.20. It can be divided in two halves, a first

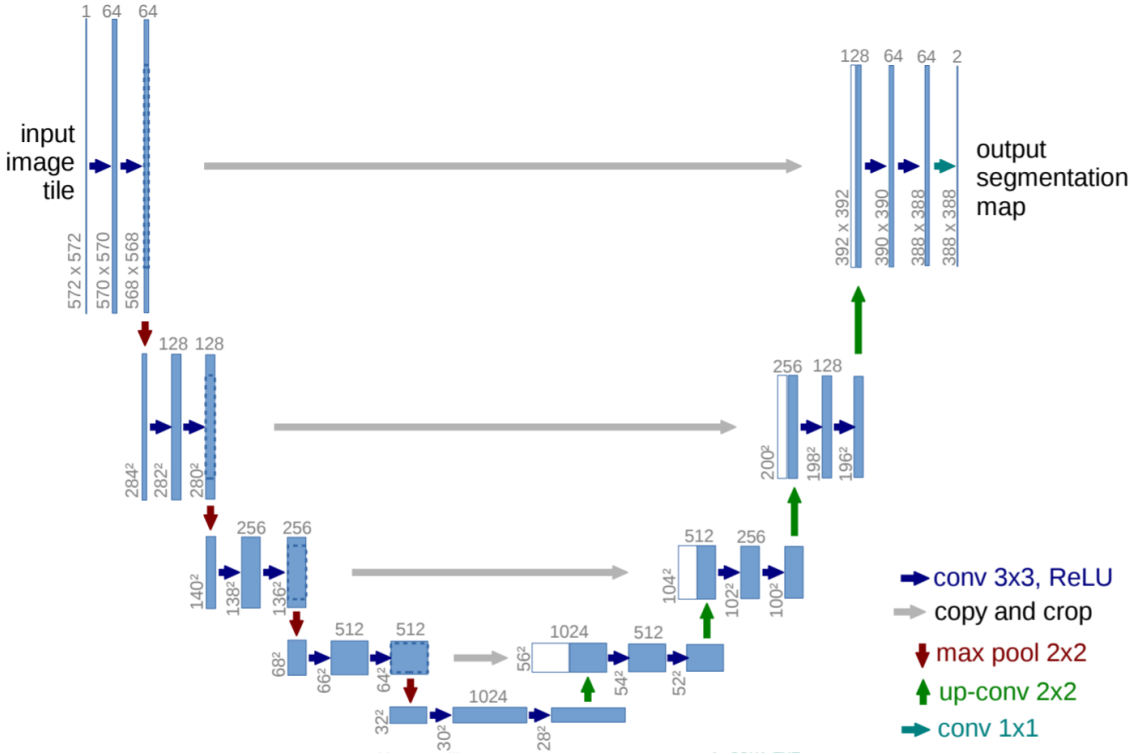
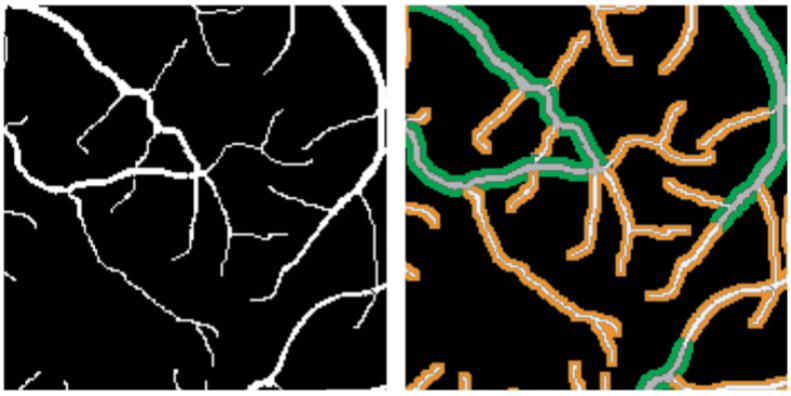


Figure 3.20: The U-Net architecture. Adapted from [23].

one encoding a hierarchy of features that consecutively represents larger portions of the image, and a second one where the high-level representations are used to learn features at up-sampled resolutions until the last layer, where a final probability over the image is obtained. Note that the skip connections allow to reuse the features from the encoding portion of the network, in order to improve the recovery of fine details. Zhang and Chung [158] resorted to this network but they extended the problem of vessel segmentation to a 5-class task, as a mechanism to increase the relevance of pixels in narrow vessels and vessel boundary regions. The considered classes were large vessels, small vessels, neighbourhood of large vessels, neighbourhood of small vessels, and the remaining background tissues, as illustrated in Figure 3.21. Jin et al. [159] also used a U-Net architecture, but they considered the deformable variant of the convolutional kernels, where the neighbourhood offsets are no longer a fixed grid but also learnable parameters in order to be more adequate to local structure (see Figure 3.22).

Other works have not relied on the capabilities of a single deep neural network. Oliveira et



(a) (b)

Figure 3.21: Extension of the binary vessel segmentation task to a 5-class one: (a) example patch from a retinal fundus image; and (b) the transformed ground truth. White, gray, orange, green, and black pixels represent, respectively, narrow vessels, wide vessels, background neighborhood of narrow vessels, background neighborhood of wide vessels, and the remaining background tissues. Adapted from [158].

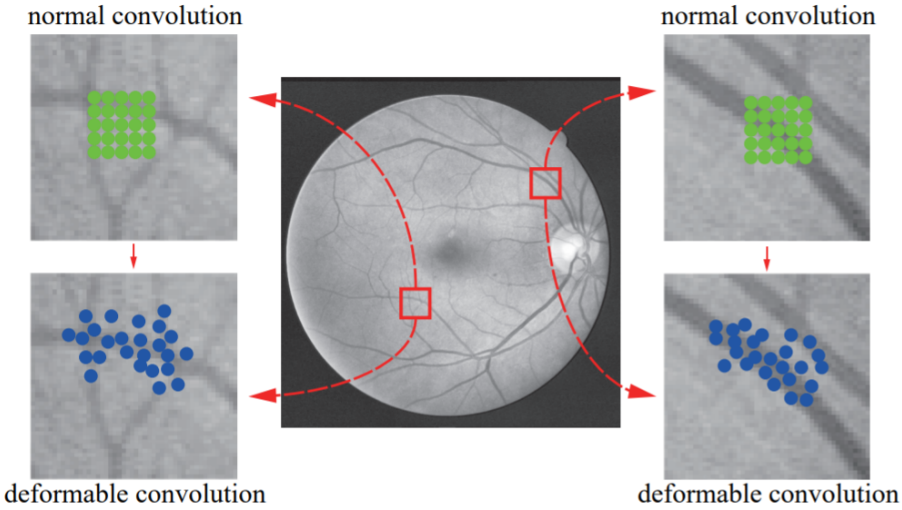


Figure 3.22: Sampling locations in regular (green) and deformable (blue) 5×5 convolution.

al. [160] provide to the network not only the image information but also the outputs of a stationary wavelet transform. The motivation behind this approach is that providing multi-scale information from the start potentially releases some of the network capacity, and it becomes easier to fuse such details with other feature representations in order to obtain the final predictions. Moreover, the authors consider data augmentation through rotations in the prediction phase, claiming that a final prediction obtained by averaging the responses is more robust. This comes from the fact that typical deep networks are not rotation invariant. Fu et al. [161] use a CNN with side-output supervision as a feature extractor, and resort to a CRF encoded as a recurrent neural network [162]

to account for non-local pixel correlations. This encoding allows to train the whole pipeline in an end-to-end fashion.

Transfer learning has been explored in retinal vessel segmentation, in order to attenuate the small amount of available images in public datasets, which have at most a few dozens of images. Jiang et al. [130] considered a pre-trained fully convolutional version of the AlexNet and fine-tuned it with up-sampled versions of retinal fundus patches. Despite using a FCN, the authors conduct patch-based prediction and merge overlapping patches with the OR operation.

3.3 Comparison between methodologies

It is very hard to objectively compare the described approaches, as they are frequently dedicated to different vessel structures and data sources. Additionally, vessel extraction frameworks result from the combination of several steps, such that it becomes non trivial to understand where particular improvements are coming from. However, an objective comparison may be established between works that have been evaluated in the publicly available DRIVE, STARE, and CHASEDB1 databases, containing retinal fundus images. Tables 3.3, 3.4, and 3.5 present objective metrics of, respectively, unsupervised, traditional ML, and DL approaches that have been here reviewed and applied to those databases.

Methodologies using traditional ML are able to slightly improve the performance when compared with unsupervised ones, mainly regarding the Area Under the ROC Curve (AUC) criterion. The DL paradigm further improved the state-of-art performance. Even though this was somewhat expected, as DL has been breaking performance barriers in many problems of Computer Vision, the improvement may seem too large for a problem that is reasonably well understood by humans. Such large improvement is expected in problems where it is very hard for a human to design features that are adequate to separate the classes. DL shines in such scenarios as it automatically finds representative mappings of the data. Regarding vessel segmentation, DL seems to find better mechanisms to distinguish vessels from other structures and noise.

Even then, the data dependency of these approaches makes them prohibitive in most real settings. For instance, the segmentation of 3D vasculature is essentially made by unsupervised techniques, as obtaining a training set in 3D would even require more effort from expert annotators. Additionally, trained models expect to evaluate images that are similar to the ones they have seen in the training step. This does not easily fit into clinical practice, where it is not feasible to produce manual annotations of images every time a new protocol or acquisition device is used. Domain adaptation techniques should increase in a near future the applicability of DL in a wider range of settings, since they improve the generalisation capabilities of these networks.

Some authors proposing label-driven approaches perform cross-training evaluation in order to analyse how well their methods generalise to data distributions that are different from the ones used during the training procedure. When comparing these methods with unsupervised ones, it is more fair to use this cross-training performance. Even then, many unsupervised approaches have hyper-parameters that are fine-tuned to the different datasets, such that most of the times it

Table 3.3: Available performances in the DRIVE, STARE, and CHASEDB1 databases, among the unsupervised works. The represented metrics are sensitivity (*sen*), specificity (*spe*), accuracy (*acc*) and AUC. Best results are highlighted in bold.

		DRIVE				**STARE**				**CHASEDB1**			
Authors	Year	AUC	*acc*	*sen*	*spe*	AUC	*acc*	*sen*	*spe*	AUC	*acc*	*sen*	*spe*
2nd observer		-	94.7	78.0	97.2	-	93.5	84.5	93.8	-	-	-	-
Hoover et al. [39]	2000	-	-	-	-	-	92.6	67.5	95.6	-	-	-	-
Mendonça and Campilho [86]	2006	-	94.5	73.4	97.6	-	94.4	70.0	97.3	-	-	-	-
Martínez-Pérez et al. [103]	2007	84.5	93.4	72.5	96.6	85.3	94.1	75.1	95.7	-	-	-	-
Al-Diri et al. [116]	2009	84.2	-	72.8	95.5	86.0	-	75.2	96.8	-	-	-	-
Palomera-Pérez et al. [104]	2010	81.1	92.2	66.0	96.1	86.0	92.4	77.9	94.0	-	-	-	-
Zhao et al. [108]	2014	-	94.8	73.5	97.9	-	95.1	71.9	97.7	-	-	-	-
Zhao et al. [126]	2015	86.1	95.3	74.4	97.8	88.1	95.1	78.6	97.5	-	-	-	-
Zhao et al. [119]	2015	86.2	95.4	74.2	**98.2**	87.4	95.6	78.0	**97.8**	-	-	-	-
Azzopardi et al. [68]	2015	**96.1**	94.4	76.6	97.0	**95.6**	95.0	77.2	97.0	**94.9**	93.9	**75.8**	95.9
Neto et al. [163]	2017	-	-	**78.1**	96.3	-	-	**83.4**	94.4	-	-	-	-
Fan et al. [164]	2019	85.9	**96.0**	73.6	98.1	88.1	**95.7**	79.1	97.0	81.5	**95.1**	65.7	**97.3**

Table 3.4: Available performances in the DRIVE, STARE, and CHASEDB1 databases, among the traditional machine learning works. The represented metrics are sensitivity (*sen*), specificity (*spe*), accuracy (*acc*) and AUC. Best results are highlighted in bold.

		DRIVE				**STARE**				**CHASEDB1**			
Authors	Year	AUC	*acc*	*sen*	*spe*	AUC	*acc*	*sen*	*spe*	AUC	*acc*	*sen*	*spe*
2nd observer		-	94.7	78.0	97.2	-	93.5	84.5	93.8	-	95.6	74.2	97.9
Niemeijer et al. [134]	2004	92.9	94.2	-	-	-	-	-	-	-	-	-	-
Staal et al. [38]	2004	95.2	94.4	-	-	96.1	95.2	-	-	-	-	-	-
Soares et al. [135]	2006	96.1	94.7	-	-	96.7	94.8	-	-	-	-	-	-
Ricci and Perfetti [136]	2007	96.3	**96.0**	-	-	96.8	**96.5**	-	-	-	-	-	-
Lupascu et al. [137]	2010	95.6	**96.0**	-	-	-	-	-	-	-	-	-	-
Marín et al. [138]	2011	95.9	94.5	-	-	**97.7**	95.3	-	-	-	-	-	-
Fraz et al. [139]	2012	**97.5**	94.8	74.1	98.1	**97.7**	95.3	75.5	97.6	-	-	-	-
Roychowdhury et al. [140]	2015	96.2	95.2	72.5	**98.3**	96.9	95.2	77.2	97.3	95.3	95.3	72.0	**98.2**
Strisciuglio et al. [141]	2016	96.0	94.5	77.8	97.0	96.4	95.3	**80.5**	97.1	-	-	-	-
Orlando et al. [142]	2017	95.1	94.5	**79.0**	96.8	95.7	95.2	76.8	97.4	94.8	94.7	75.6	96.6
Zhang et al. [146]	2017	97.0	94.7	78.6	97.1	97.4	95.5	78.8	97.3	**97.1**	95.0	76.4	97.2
Wang et al. [147]	2019	-	95.4	76.5	98.2	-	96.4	75.2	**98.8**	-	**96.0**	**77.3**	97.9

Table 3.5: Available performances in the DRIVE, STARE, and CHASEDB1 databases, among the deep learning works. The represented metrics are sensitivity (*sen*), specificity (*spe*), accuracy (*acc*) and AUC. Best results are highlighted in bold. The four best architectures of Liskowski and Krawiec are here presented. They are, respectively, BALANCED-SP with output window of side 3, NO-POOL-SP with output window of side 3, BALANCED-SP with output window of side 5 and NO-POOL-SP with output window of side 5.

			DRIVE				**STARE**				**CHASEDB1**			
Authors	Year		AUC	*acc*	*sen*	*spe*	AUC	*acc*	*sen*	*spe*	AUC	*acc*	*sen*	*spe*
2nd observer			-	94.7	78.0	97.2	-	93.5	84.5	93.8	-	95.6	74.2	97.9
Melinščak et al. [152]	2015		97.5	94.7	72.8	97.8	-	-	-	-	-	-	-	-
Fu et al. [161]	2016		-	95.2	76.0	-	-	95.8	74.1	-	-	94.9	71.3	-
Li et al. [154]	2016		97.4	95.3	75.7	98.2	98.8	96.3	77.3	98.4	97.2	95.8	75.1	97.9
Liskowski and Krawiec [153]	2016	(1)	97.9	95.1	84.6	96.7	**99.3**	96.7	**92.9**	97.1	98.2	94.4	**91.6**	94.7
		(2)	97.7	95.2	80.2	97.6	99.2	97.1	85.1	98.5	98.2	96.2	79.1	98.2
		(3)	97.9	95.3	81.5	97.5	**99.3**	97.0	90.8	97.7	98.4	95.8	87.9	96.7
		(4)	97.9	95.4	78.1	98.1	**99.3**	**97.3**	85.5	98.6	98.2	96.3	78.2	98.4
Feng et al. [157]	2017		97.9	95.6	78.1	98.4	-	-	-	-	-	-	-	-
Dasgupta and Singh [156]	2017		97.4	95.3	76.9	98.0	-	-	-	-	-	-	-	-
Oliveira et al. [160]	2018		98.2	95.8	80.4	98.0	99.0	96.9	83.2	98.6	98.6	96.5	77.8	98.6
Zhang and Chung [158]	2018		98.0	95.0	**87.2**	96.2	98.8	97.1	76.7	99.0	**99.0**	**97.7**	76.7	**99.1**
Jiang et al. [130]	2018		98.1	96.2	75.4	98.2	99.0	**97.3**	83.5	98.5	98.1	96.7	86.4	97.4
Jin et al. [159]	2019		**98.6**	**97.0**	78.9	**98.7**	98.7	**97.3**	74.3	**99.2**	98.6	97.2	82.3	98.2

Table 3.6: Available performances of label-driven works when performing cross-training. The dataset in bold is the test set and the one inside parenthesis is the training set. Best results are highlighted in bold.

		DRIVE (STARE)				**STARE** (DRIVE)			
Authors	Year	AUC	*acc*	*sen*	*spe*	AUC	*acc*	*sen*	*spe*
2nd observer		-	94.7	78.0	97.2	-	93.5	84.5	93.8
Soares et al. [135]	2006	-	94.0	-	-	-	93.3	-	-
Marín et al. [138]	2011	-	94.5	-	-	-	95.3	-	-
Fraz et al. [139]	2012	97.0	94.6	72.4	97.9	96.6	95.0	70.1	**97.7**
Roychowdhury et al. [140]	2015	-	94.9	-	-	-	95.1	-	-
Li et al. [154]	2016	96.8	94.9	**72.7**	98.1	96.7	95.4	70.3	98.3
Oliveira et al. [160]	2018	**97.5**	**95.0**	67.1	**99.2**	**98.5**	**96.0**	**84.5**	97.3

is not trivial to compare the robustness of different methods to new data. Table 3.6 presents the performance achieved by different authors when considering cross-training between the DRIVE and STARE datasets. As can be seen, there is a natural decrease of performance, however the top-performing approach still achieves a larger AUC value than unsupervised and traditional ML algorithms, showing that DL can also generalise rather well.

3.4 Summary

This Chapter covered the main approaches for vessel enhancement and segmentation. Matched filtering, Hessian-based filters, and mathematical morphology were considered for the enhancement task. Matched filtering shines when the objects to be found are flawlessly described by a particular shape and local contrast. However, vessels present variability in terms of appearance, such that the enhancement of different cross section profiles may require the design of multiple filters. These methods easily reach high computational cost as a response is computed over the entire data and at multiple orientations. In addition, handling very tortuous segments is not trivial with this kind of approach. Hessian-based filters have been widely applied in the vessel enhancement field. Their ability to evaluate local shape in a scale-space provides a great opportunity to enhance vessels of different sizes. However, as the Hessian relies in the derivative operator, it propagates noise, and it becomes extremely challenging to detect small vessels without increasing false detections. Moving on to mathematical morphology, methods applying this theory are usually fast and noise resistant, even though they do not take into account the available cross sectional information. This may be crucial when distinguishing vessels from other structures that locally resemble a tubular pattern, as is the case of some pathologies.

Concerning vessel segmentation, centreline tracking, region growing, active contour, graph cut, and machine learning approaches were presented. Centreline tracking methods iteratively detect points along the centre of a vessel, focusing a single segment at each time. This framework is able to naturally present more detailed information on a certain vessel segment and is generally

fast, given that only a portion of the data needs to be analysed. Even then, automatic initialisation may be a source of error and bifurcation detection requires special considerations. Missing a bifurcation will result in the loss of portions of the vascular network. Region growing approaches heavily suffer from leakage problems. Even then, they have high efficiency since a sparse search is employed. ACMs are dependent of the initialisation and parameter tuning, such that they have a strong application-dependent character. Parametric models have to deal with strong elongations, and contour splitting and merging, which is not trivial. Implicit representations avoid the parametrisation of the contour, handling better those problems. However, computational cost is increased and special algorithmic care has to be taken to ensure convergence. Graph Cuts are affected by the shrinking bias problem, where there is a bias towards small final contours. This is more evident in the segmentation of elongated structures, such as vessels. Additionally, the memory usage quickly increases with the size of data. Label-driven methods achieve better results, especially when the images are very similar to the ones seen during the training step. Nonetheless, they demand that manual annotations are available. Given that the manual segmentation of data by an expert is a very time-consuming and tedious task, it is not a common scenario in many real-life applications. Even then, advances in their generalisation capabilities are increasing their use cases.

Chapter 4

Application-specific Blood Vessel Segmentation

The content of this Chapter is based on the following works:

- R. J. Araújo and H. P. Oliveira, "Segmentation of the rectus abdominis muscle anterior fascia for the analysis of deep inferior epigastric perforators", In *Iberian Conference on Pattern Recognition and Image Analysis*, 2017.
- R. J. Araújo et al., "Computer aided detection of deep inferior epigastric perforators in computed tomography angiography scans", *Computerized Medical Imaging and Graphics*, 2019.
- R. J. Araújo, J. S. Cardoso, and H. P. Oliveira, "A single-resolution fully convolutional network for retinal vessel segmentation in raw fundus images", In *International Conference on Image Analysis and Processing*, 2019.
- C. Mavioso et al., "Automatic detection of perforators for microsurgical reconstruction", *The Breast*, 2020.

The majority of works comprising the blood vessel segmentation literature are designed to solve a particular use-case and struggle to generalise properly for different scenarios, such as other vessel trees, acquisition devices, and used protocols. This is especially true when supervised machine learning methodologies are used, where the models are likely to learn particularities of the training data which will probably do more harm than good when those models are applied to data from other scenarios. Nonetheless, frameworks focusing a single application may further rely on prior knowledge and establish assumptions which are known to be useful for a given scenario. Hence, application-specific algorithms have been showing great success and are possibly the best choice for controlled environments where the upcoming data to be analysed meet the assumptions considered during design.

In this Chapter we describe two methodologies that we designed for particular use-cases of blood vessel segmentation: i) a semi-automatic algorithm for the segmentation and characterisation of the DIEPs, the blood vessels that vascularise the anterior portion of the abdominal wall and are crucial for a successful breast reconstruction through a DIEP flap (Section 4.1); ii) a supervised approach for the fast segmentation of retinal vessels, which is relevant in screening programmes where large volumes of data are collected (Section 4.2).

4.1 Deep Inferior Epigastric Perforators

About 2.1 million newly diagnosed female breast cancer cases were expected worldwide in 2018, accounting for almost 1 in 4 cancer cases among women. The disease is the most frequently diagnosed cancer in the vast majority of the countries and it is also the leading cause of cancer death in over 100 countries [165]. Women who were diagnosed with breast cancer have higher chance of suffering from anxiety and depression resulting from the fear of recurrence, body image disruption, sexual dysfunction and mortality concerns [166]. Although breast conservative methods have recently shown a survival rate superior to mastectomy, especially in early breast cancer cases [167], the latter is still a highly recurrent procedure and has even been increasing in some institutions [168; 169]. Mastectomy is conducted in cases where the relation between the size of the resected breast and the global volume of the gland is too large for a conservative procedure, in cases where radiotherapy is contra-indicated, and also when the patient does not desire breast conservation [170].

Reconstruction methods allow to recreate the breast, improving the way women feel about themselves and their image after their breast(s) was(were) removed. There are different techniques for breast reconstruction but basically two major groups can be defined: reconstruction with implants and reconstruction with autologous tissues. Each has its own merits but the latter commonly lasts longer, provides a more natural result due to the similarity of tissues, and avoids foreign body reactions. Nonetheless, as it also involves a donor-site, it leads to longer surgery and recovery time, and may lead to donor-site complications. The DIEP flap has become the state-of-art technique for autologous-based breast reconstruction [22]. It makes use of the skin and fat (it does not include the muscle, its big advantage over the Transverse Rectus Myocutaneous flap) of the lower abdomen to build a new breast either in the same surgery when the mastectomy is performed (immediate reconstruction) or in a second one after the initial procedure (delayed reconstruction). The transposition of the lower abdominal skin and fat is free of any attachment to the end anatomic structures of the donor site - the abdomen. Micro-surgical connections are done at the recipient site between the vessels of the transposed skin and fat, and the vessels of the thorax, where the new breast will replace the void left by the mastectomy (see Figure 4.1). A scheme with the abdominal anatomy of interest for conducting a DIEP flap is shown in Figure 4.2.

Before a DIEP flap, preoperative imaging studies are performed to evaluate the branches of the DIEAs, which are known as perforators (DIEPs) and are the vessels responsible for the vascularisation of the tissue that will be used in the reconstruction of the breast. The viability of the new

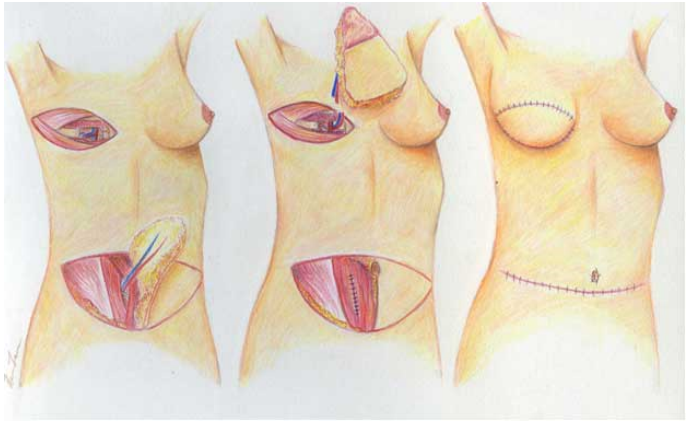


Figure 4.1: Representation of a DIEP flap procedure.

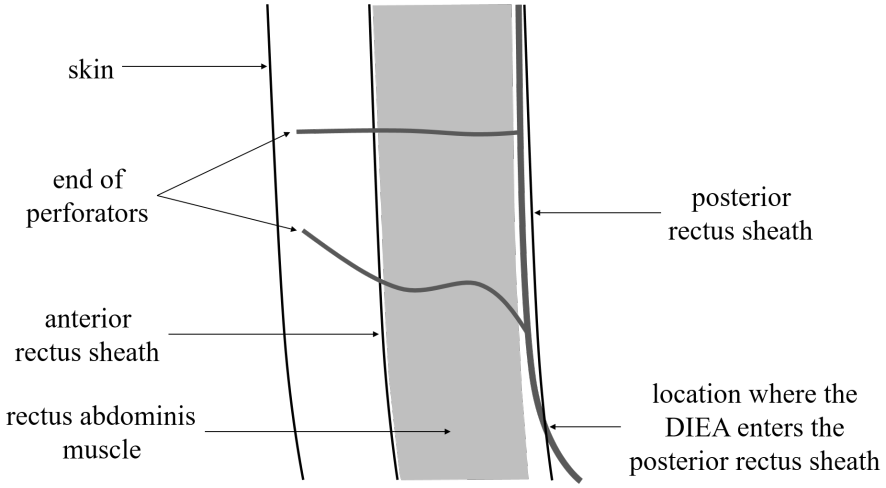


Figure 4.2: Sagittal representation of the anatomy of the anterior portion of the abdominal wall, between the pelvic and umbilicus regions.

breast is related to several features of the included perforators [26]. The calibre of a perforator is an indicator of its capacity to ensure a good vascularisation of the new breast, and perforators with a larger number of ramifications and anastomoses with others are usually preferred. The surgical team also requires the location where the perforators pierce the anterior fascia of the muscle, in order to know how to extract the perforators to be included in the flap. Usually, perforators having linear intramuscular courses are preferred, since the dissection becomes more challenging as the tortuosity increases. Through MRI or CTA, the perforators are manually identified and characterised by the radiological team, and a report is delivered to the surgeons. The aforementioned task of identifying and characterising the 3D course of the perforators is subjective and time consuming. This is exacerbated by the small size of these blood vessels (1-3 mm of calibre, which translates to 1-5 pixels in a state-of-the-art CTA) and their low Signal-to-Noise Ratio (SNR), especially in their intramuscular region. As a result, incoherences between the preoperative studies

and the surgical findings often exist, and can lead to the need of modifying the strategy intra-operatively. For that reason, computer vision algorithms may play an important role in supporting the activity of radiologists who are responsible for the preoperative study, reducing the subjectivity and time involved in the process. Moreover, more precise and complex representations of the findings can be shown through 3D models, whereas the manually written reports strongly limit the information exchange. Other authors have developed plugins for medical software trying to render the manual analysis of the perforators faster [171] and used virtual tools to facilitate the communication of the manual findings with the surgical team [172]. In this Section, we start by providing a brief description of the framework we developed for this scenario during preliminary work [24], which was the first work aiming to automate the extraction of the DIEP vascular tree, requiring only minimal user input. Afterwards, we present the contributions that were made during this thesis and which allowed the methodology to become closer to the clinical needs, and describe the clinical validation which was conducted to show the potential and impact of the proposed algorithm. All the work has been performed with the collaboration of the Breast Unit of Champalimaud Foundation [1].

4.1.1 Previous work

In the past, motivated by the aforementioned reasons, we have proposed a methodology for the semi-automatic extraction and characterisation of the perforators in CTA scans, by means of a vessel centreline extraction technique particularly designed to address the challenges involved in the detection of these small vessels [24]. The complete pipeline is represented in Figure 4.3.

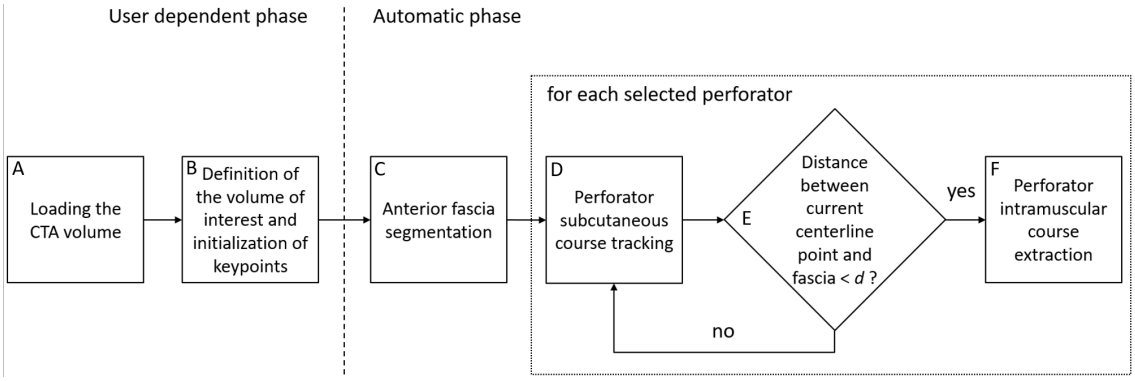


Figure 4.3: The previously proposed pipeline for the semi-automatic detection and characterisation of DIEPs [24]. Threshold d dictates when the tracking procedure of the subcutaneous portion of a perforator finishes.

After loading the data (A), the user is prompted to select the landmarks required by the algorithm to define the volume of interest and posteriorly extract and characterise the DIEP vessel tree in an automated fashion. The user should indicate two points at the end of the subcutaneous course of each perforator and also the locations where each DIEA perforates the posterior rectus

[1] centroclinico.fchampalimaud.org/en/oncology/breast-unit/

sheath (B). This effort is minimal when compared to the current manual analysis that technicians and radiologists face, which involves performing several reconstructions of the data through the Maximum Intensity Projection method [22; 173], in order to follow the course of each perforator and characterise it. The volume of interest should contain the portion of the data that needs to be analysed in this application, which is only a sub-volume of the abdominal CTA scan (anterior region of the abdomen, lower bounded by the hips and upper bounded slightly above the umbilical region). An example region of interest and the structures which exist there were already shown in Figure 2.11.

There is a higher contrast between the vessels and the background in their subcutaneous portion, since the rectus abdominis muscle also responds significantly to the CTA image acquisition. This motivated us to pursue different strategies when addressing the extraction of the subcutaneous and intramuscular courses of the perforators. Anatomically, those regions are separated by the anterior fascia (or anterior rectus sheath) of the rectus abdominis muscle. Unfortunately, that tissue layer is not distinguishable in CTA scans. Nonetheless, it can be approximated as the edge that exists between the subcutaneous region and the muscle. We started by segmenting this layer, thus enabling the employment of a divide-and-conquer strategy to extract the complete course of the perforators.

To segment the anterior fascia of the rectus abdominis muscle (C), we proposed a 2-stage algorithm. First, we obtain a preliminary detection of this layer for each axial slice of the volume of interest by employing an algorithm resorting to intensity and connected component analysis, relying on prior anatomical knowledge. A summary of the steps and illustrative results are provided, respectively, in Figures 4.4 and 4.5. Finally, to obtain a complete and smooth fascia segmentation, refined detections were set as the output of local regressions using the preliminary ones. Figure 4.6 provides illustrative examples of this final step. A detailed description of the methodology followed to segment the fascia can be found in [24; 174].

Having two landmarks on the subcutaneous portion of each perforator to be tracked and a

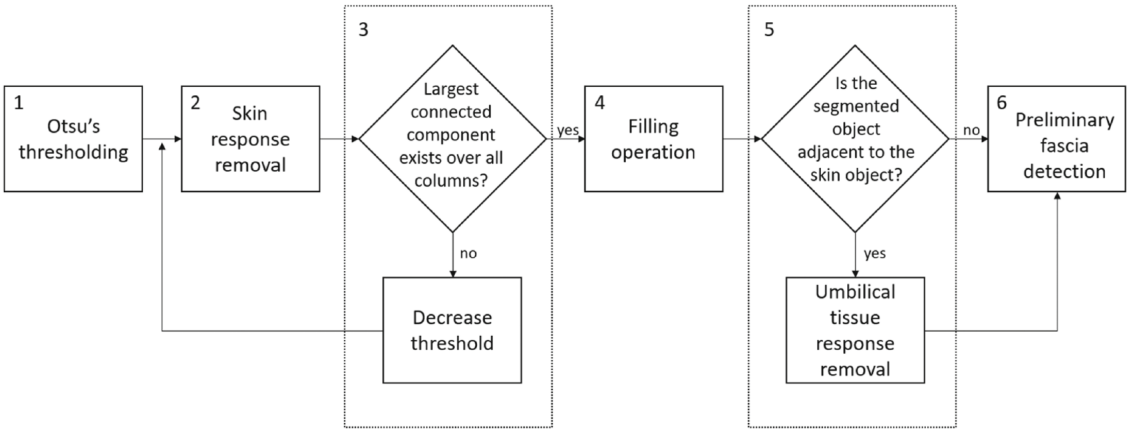
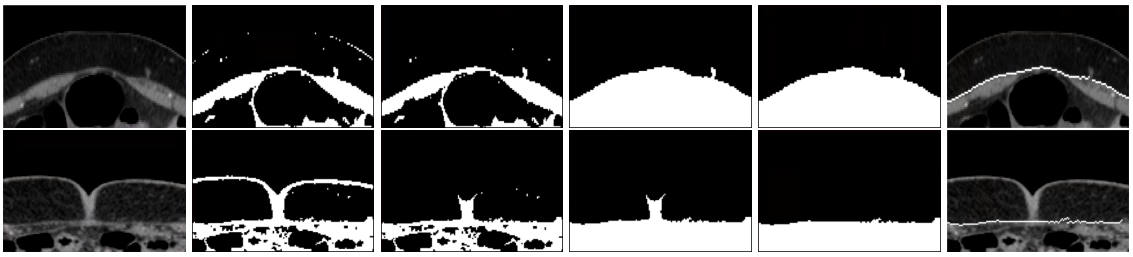
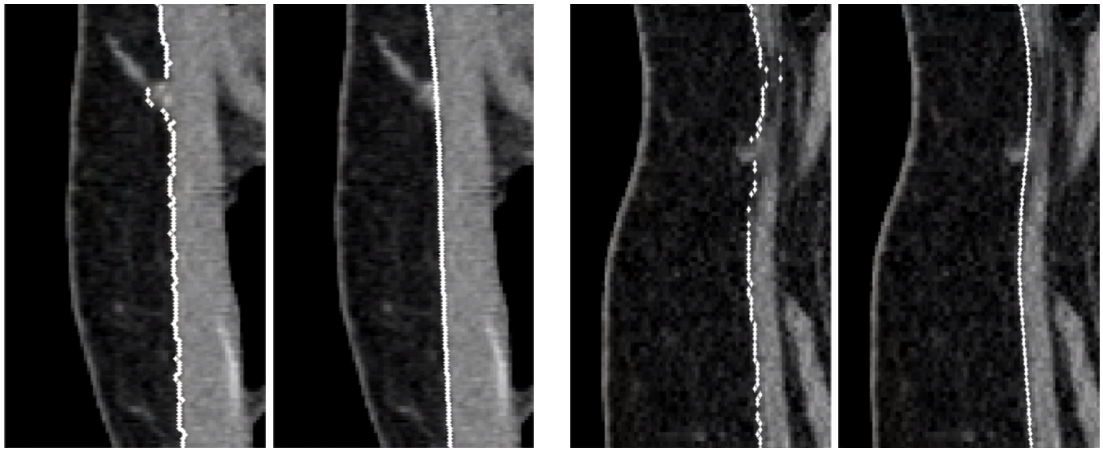


Figure 4.4: Pipeline used to obtain a preliminary fascia detection for each axial slice of the volume of interest.



(a) (b) (c) (d) (e) (f)

Figure 4.5: Example results along the considered pipeline for obtaining a preliminary fascia segmentation: (a) original axial slices of the region of interest; (b-f) segmentations after steps 1, 2, 4, 5, and 6. In the last column, the preliminary fascia layer is shown in white.



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

Figure 4.6: Example results of the interpolation framework implemented to obtain the final fascia segmentation: sagittal slices of the region of interest with (a,c) the preliminary fascia detection; and (b,d) the smooth final fascia segmentation.

segmentation of the fascia, it becomes possible to use a centreline tracking procedure (D) to iteratively estimate new points along the vessel until the distance to the fascia layer becomes lower than a specified threshold. To employ this framework, we estimated the local vessel direction by analysing the neighbourhood gradient vectors, as Agam et al. [175] considered for steering enhancement filters. According to the authors, the local vessel direction $\mathbf{v}_i$ is the one minimising the squared projection of the local gradient vectors into itself:

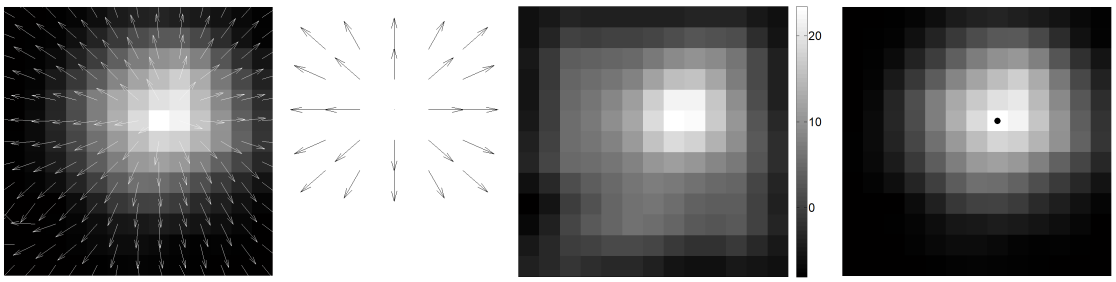
$$E(\mathbf{v}_i) = \frac{1}{n}\sum_{k=1}^{n}(\mathbf{g}_k^T\mathbf{v}_i)^2 = \mathbf{v}_i^T\left(\frac{1}{n}\sum_{k=1}^{n}\mathbf{g}_k\mathbf{g}_k^T\right)\mathbf{v}_i \tag{4.1}$$

where n is the number of gradient vectors inside a local window specifying the neighbourhood, and $\mathbf{g}_k$ is the kth gradient vector. A tracker relying only in this information is prone to start deviating from the centre of the perforator as error accumulates through the iterative procedure. This is further exacerbated by the low SNR characterising this scenario. Thus, we also considered a

correction framework that would estimate the centre of the cross section of the vessel every n iterations, allowing to correct the accumulated bias. It relies on the assumption that voxels on the centre of the vessel have higher intensity, and that it decreases as the distance to the centre increases. In a 2D image of the cross section of a vessel (plane perpendicular to the local estimated vessel direction), it is then expected that the centre location can be found by analysing the divergence of the gradient vector field, through cross-correlation with a template [176]:

$$(\mathbf{f} * \mathbf{g}[\eta]) = \sum_{m} \mathbf{f}^{*}[m]\mathbf{g}[\eta + m] \tag{4.2}$$

where $\mathbf{f}$ and $\mathbf{g}$ represent the gradient orientation vector field and template vector field, respectively, $\mathbf{f}^{*}$ is the complex conjugate of $\mathbf{f}$, and η is the displacement. The centre location estimation corresponds to the maximum response location. Figure 4.7 illustrates this procedure.



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

Figure 4.7: Ridge-based correction framework: (a) example cross sectional image with gradient vector field imposed; (b) template for finding the centre; (c) cross-correlation result; and (d) the detected ridge. Images are interpolated for better visualisation.

As soon as the distance to the fascia criterion is met, the process is terminated and the last point is considered the location where the perforator pierces the fascia, one of the features that the surgeons require in order to know the dissection place of each perforator, in case they choose to include it in the flap. The calibre of the perforator, which is also relevant for the surgical team, is estimated at each centreline point and the average is returned. For a given centreline point, the local calibre is measured by fitting a Gaussian to the intensity profile of a line on the axial plane centred at that point and whose direction is perpendicular to the projection of the local vessel direction into the axial plane.

The intramuscular path of the perforator (F) was found by extracting the minimum cost path between the location where it leaves the fascia (where the subcutaneous tracker ended) and the DIEA of the respective hemiabdomen (provided by the user), using the A* algorithm [177]. This path-finding method includes an heuristic function estimating the distance to the target, in order to decrease the computational cost of the search. We considered the Euclidean distance as the heuristic function. To obtain the cost of traversing each voxel, we started by enhancing the tubular

structures through Frangi's vesselness measure v_F [75]:

$$v_F(n) = \begin{cases} 0, & \text{if } \lambda_2 > 0 \text{ or } \lambda_3 > 0 \\ (1 - e^{-{R_A}^2/2\alpha^2}) \cdot e^{-{R_B}^2/2\beta^2} \cdot (1 - e^{-S^2/2c^2}), & \text{otherwise} \end{cases} \tag{4.3}$$

where n is a given voxel in our data domain, and $\lambda_1, \lambda_2, \lambda_3$ are the eigenvalues of increasing absolute value of the Hessian matrix. The constants α, β, and c control the sensitivity of the vesselness function to the terms R_A, R_B, and S, which are eigenvalue-based ratios accounting for, respectively, the distinction between line-like and plate-like structures, the deviation from a blob, and the local structureness:

$$R_A = \frac{|\lambda_2|}{|\lambda_3|} \tag{4.4}$$

$$R_B = \frac{|\lambda_1|}{\sqrt{|\lambda_2 \lambda_3|}} \tag{4.5}$$

$$S = \sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2} \tag{4.6}$$

Lower costs are given to voxels which more likely belong to vessels, as determined by the following function:

$$C(n) = \begin{cases} 2 - v_F(n), & \text{if } v_F(n) > 0 \\ 10, & \text{if } v_F(n) = 0 \end{cases} \tag{4.7}$$

where the constants have been empirically determined for this application. Many perforators have a significantly tortuous intramuscular course, such that the Euclidean distance heuristic commonly underestimates the true vessel length. This, combined with the difficult enhancement due to the low SNR, leads to a tendency to expand many nodes and consequently a very slow extraction in many cases. To overcome this, an upper bound m was set on the number of nodes allowed in the open list. If the upper bound is reached, only the k nodes closer to the target (Euclidean distance) are kept, whereas the remaining ones are moved to the closed list. It is important to notice that, with this consideration, the final path is not guaranteed to be the optimal one. Nonetheless, it was empirically found that significant speed-ups were possibly without impacting the accuracy of the extracted paths.

4.1.2 Local shape analysis for increased tracking robustness

The aforementioned methodology allowed to obtain promising results [24]. Nonetheless, the tracking of perforators having a significant part of the course adjacent to the fascia often lead to an early stop of the tracking procedure, due to the strong gradient vectors coming from the muscle boundary and their impact in the local vessel direction estimation.

To increase the robustness of the tracker, we aimed to reduce the impact of non-tubular structures in the gradient field estimation. Therefore, we proposed to retrieve the local gradient vectors from a vessel-enhanced data instead, where the signal of tubular-shaped structures is increased, while other local shapes have their signal decreased. Considering the particular case of the muscle boundary, when considering the eigenvalue decomposition described in Section 3.1.2, an eigenvalue with large magnitude (eigenvector perpendicular to the muscle boundary) and two small ones (eigenvectors defining the plane tangent to the muscle boundary) are expected. Thus, by conducting vessel enhancement via the Hessian analysis, we decreased the influence of the muscle boundary and hence increased the robustness of the tracker in the regions nearby. To obtain the vessel enhanced data, we considered once again Frangi's vesselness function [75].

4.1.3 Cost functions for the efficient extraction of intramuscular paths

The intramuscular paths found with the methodology presented in subsection 4.1.1 are not necessarily the optimum ones, due to the limit that was imposed in the open nodes list. This constraint was introduced to avoid visiting a very large amount of voxels, which was quite common, given that some courses had significant length and tortuosity, and sometimes the volume of costs was not well defined due to the extremely low SNR of the intramuscular course of the perforators. It is often possible to find some segments of this course without any signal.

We investigated whether another family of cost functions could induce more interesting compromises between performance and time efficiency. We explored a function which proved to be useful in the centreline extraction of coronary arteries [178]:

$$C(s) = \frac{1}{\nu_F(s)T(s) + \varepsilon} \tag{4.8}$$

where $\nu_F(s)$ is Frangi's Vesselness (4.3) at voxel s normalised to the range $[0,1]$, ε is a small constant to avoid division by zero, and $T(s)$ is a sigmoid function of the intensity:

$$T(s) = \frac{1}{1 + e^{-a_s(I(s)-b_s)}} \tag{4.9}$$

with $I(s)$ being the intensity at voxel s, and constants a_s and b_s controlling the shape of the sigmoid function.

The expression (4.8) produces low costs at voxels which have high probability of belonging to a vessel, according to (4.3), and which have relatively high radio-density, according to the parametrisation of (4.9). Note that (4.8) gives costs in the range $[1,\infty[$, guaranteeing that the heuristic included in the A* formulation is admissible.

4.1.4 Experiments

4.1.4.1 Perforator extraction

Experiments were conducted at two different time lapses. The first assessment was performed with a preliminary version of the database described in Section 2.3, aiming to evaluate the impact of the proposed modifications to the methodology. For every CTA volume of this database, a radiologist provided manual annotations of the existing perforators, by defining some landmarks belonging to the centrelines of those vessels. Across the 21 volumes available at the time, a total of 98 subcutaneous and 50 intramuscular perforator pathways were identified. Since the blood vessel annotations available in this dataset are sparse when compared to the paths extracted by our methodology, we consider the Euclidean and Hausdorff distances from the ground truth annotations to the extracted paths, as metrics indicating the precision of the vessel detection methodologies. The Euclidean distance measures the average distance from the manual annotations to the path, whereas the Hausdorff distance accounts for the maximum distance. Additionally, as we are interested in assessing the time efficiency of the minimum cost path approach designed for intramuscular path extraction, we also consider the time expended in that subtask.

4.1.4.2 Perforator characterisation

A second set of experiments was performed more recently [179] using the extended database (the one detailed in Section 2.3). The goal was to compare the findings (calibre, and the location where the perforator pierces the fascia) retrieved by the proposed framework with those obtained manually by the radiological team. The following statistical tests were conducted:

1. A paired sample t-test with a significance level $\alpha = 0.05$ evaluated the difference in calibres between both methods (null hypothesis H_0 was that the average of the differences is 0, and the alternative hypothesis H_1 is that this does not hold). The same analysis was also performed only in the perforators that ended up being included in the flaps by the surgical team. Effect size was measured through the standardised mean difference and interpreted accordingly (estimates of 0.20, 0.50, and 0.80 denote, respectively, a small, medium, and large effect size) [180];

2. A Wilcoxon rank-signed test with a significance level $\alpha = 0.05$ was used to assess the difference in calibre estimation between both methods when compared to calibres reported by the surgical team intra-operatively (ground truth). This analysis was performed in perforators having a calibre larger than 1.5 mm (frequently included in the flap) and also in the ones having 1.5 mm or less (rarely selected). Effect size was measured through correlation and interpreted accordingly (estimates of 0.10, 0.30, and 0.50 denote, respectively, a small, medium, and large effect size) [180];

3. The methodology was also evaluated on how well it estimated the location where the perforators pierce the fascia of the rectus abdominis muscle. The errors of the software were

divided into vertical and horizontal errors and a one sample t-test was performed for each of them, where in H_0 the average error is 0 and in H_1 it is not. Effect size was measured through the standardised mean difference. Here, the true locations of the perforators were considered to be the ones reported by the radiological team, as these measures cannot be accurately obtained during the surgery. Additionally, the subjectivity is practically negligible for this parameter, in contrast with calibre measurement.

4.1.5 Results and discussion

4.1.5.1 Perforator extraction

Subcutaneous course Our proposed methodology for extracting the subcutaneous portion of perforators is now assessed concerning the metrics described in subsection 4.1.4.1, and compared to both the previous version of the algorithm and the tracking approach of Friman et al. [69], which was tailored for the robust detection of small vessels.

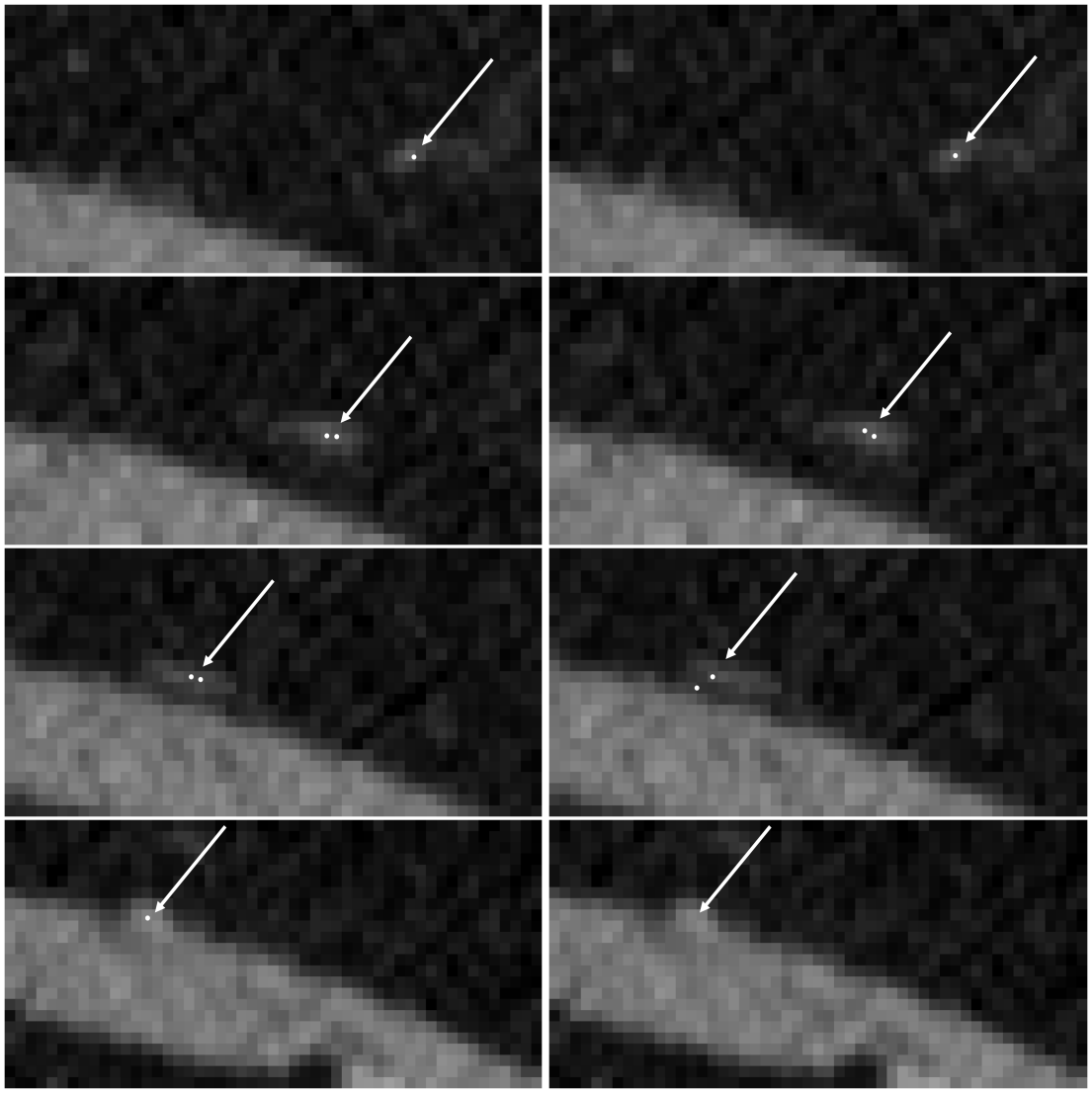
Regarding the implementation of our approach, we empirically set the step δ to 1 mm, the side length of the local window to 4 mm, took a correction measure every 3 iterations, restricted the direction variation to 60°, and obtained the vessel enhanced data by setting $\alpha = 0.5$, $\beta = 10$, and $c = 500$. In order to obtain the results from the approach of Friman et al. [69], we used the implementation available in MeVisLab (MeVis Medical Solutions AG, 2017), and tuned the parameters for this particular application, by setting the minimum and maximum radius to 0.5 and 1.5 mm, respectively, the step length to 1 mm, and the maximum step angle to 60°. The initialisation of all perforator tracking procedures was made at the ground truth landmark which was closer to the end of the perforator. Table 4.1 summarises our findings.

Table 4.1: Results obtained for the extraction of the subcutaneous course of the perforators, having as reference the manual ground truth annotations.

Method	Path error (mm)	
	Euclidean distance	Hausdorff distance
Friman et al. [69]	1.01 ± 0.60	2.38 ± 2.17
Previous work (subsection 4.1.1)	1.35 ± 0.46	2.98 ± 1.46
Proposed	$\mathbf{0.64 \pm 0.25}$	$\mathbf{1.17 \pm 0.88}$

Our refined methodology proved to be more adequate for the subcutaneous tracking of perforators, as the average Euclidean and Hausdorff distances were significantly lower when using it. In fact, the proposed methodology reached sub-voxel accuracy for most of the volumes in our database. The increased performance was mainly due to the fact that our approach is able to neglect more the presence of the muscle when tracking the vessel, as we do it in the vessel enhanced data. In contrast, the methodology of Friman et al. [69] suffered more from this. The average Hausdorff distance gives information about how well the methods are able to correctly track the perforator until it reaches the fascia, as it is near this region that the tracking procedure faces

more difficulties, especially when the perforator has a substantial overlap with the muscle signal. Again, the proposed method behaved better in such circumstance, thus being more adequate for determining the location where the perforators pierce the fascia. Figure 4.8 shows a comparison between the proposed method and the approach of Friman et al. [69] when extracting an example subcutaneous path having a segment close to the fascia.



(a) (b)

Figure 4.8: Tracking the subcutaneous course of a perforator using (a) our approach and (b) the method proposed by Friman et al. [69]. The latter was more prone to terminating sooner when the vessel evolved near the fascia.

Intramuscular course extraction Concerning the extraction of intramuscular paths, we seek to find how appropriate is the volume of costs given by (4.8) when used as terrain costs inside a minimum path framework such as A*. Note that by appropriate, we refer to path accuracy but

also time expended, as minimum path approaches may become prohibitively slow when there is a need to visit many neighbours, especially in a 3D environment. To obtain $v_F(n)$, we empirically set $\alpha = 0.5$, $\beta = 0.5$, and $c = 100$. Regarding $T(n)$, we considered 42 different parametrisations, as given by the possible combinations of taking a_s from 7.5 to 45 with a step of 7.5, and b_s from 0.5 to 0.8 with a step of 0.05 (radio-density was mapped into the range $[0,1]$ as described in subsection 4.1.3). Besides the metrics considered in the case of subcutaneous extraction evaluation, Euclidean and Hausdorff distances between the ground truth annotations and the retrieved paths, we also measured the time expended. Table 4.2 presents the results of our experiments. For the sake of readability, we show only the average value of the metrics of interest.

Table 4.2: A* performance when retrieving the intramuscular courses of the perforators, using Eq. (4.8) to obtain the volume of costs. Each cell contains the performance of a parametrisation (a_s, b_s) regarding the average Euclidean and Hausdorff distances between the ground truth annotations and the retrieved paths (mm), and the average time expended (s), respectively.

		b_s						
		0.50	**0.55**	**0.60**	**0.65**	**0.70**	**0.75**	**0.80**
a_s	**7.5**	0.62 1.40 36.1	0.62 1.38 39.5	0.62 1.38 40.8	0.62 1.37 41.8	0.61 1.36 42.2	0.61 1.36 41.1	0.61 1.36 40.1
	15	0.61 1.32 60.8	0.60 1.30 64.2	0.60 1.29 56.1	0.59 1.27 26.7	0.56 1.13 12.7	0.54 1.06 7.48	0.52 1.01 7.90
	22.5	0.60 1.32 56.7	0.59 1.28 51.2	0.57 1.22 19.6	0.52 0.99 6.70	0.52 0.99 4.80	0.50 0.96 17.1	0.51 0.96 109
	30	0.61 1.38 53.4	0.58 1.22 41.3	0.64 1.38 7.10	0.52 0.98 3.70	0.50 0.95 28.7	0.69 1.31 1340	1.53 3.11 843
	37.5	0.59 1.28 49.5	0.65 1.41 25.9	0.59 1.18 3.51	**0.50** **0.96** **14.6**	0.50 0.94 237	1.48 3.00 756	2.10 4.54 3031
	45	0.67 1.48 46.3	0.65 1.42 10.2	**0.51** **1.00** **3.50**	0.50 0.97 89.1	1.39 2.86 514	1.81 3.81 1154	2.67 5.70 2022

The performance manifold, concerning each of the metrics, is visually represented in Figure 4.9. These manifolds allow us to conclude that the influence of parameters a_s and b_s on the overall performance is not linear. Instead, it is a particular combination of both that may lead to a reasonable volume of costs. This was somewhat expected, as a_s dictates the steepness of the sigmoid function, hence the intensity compression, and b_s sets the threshold of the sigmoid, controlling the range of intensities that produce lower costs. The parametrisations highlighted in bold

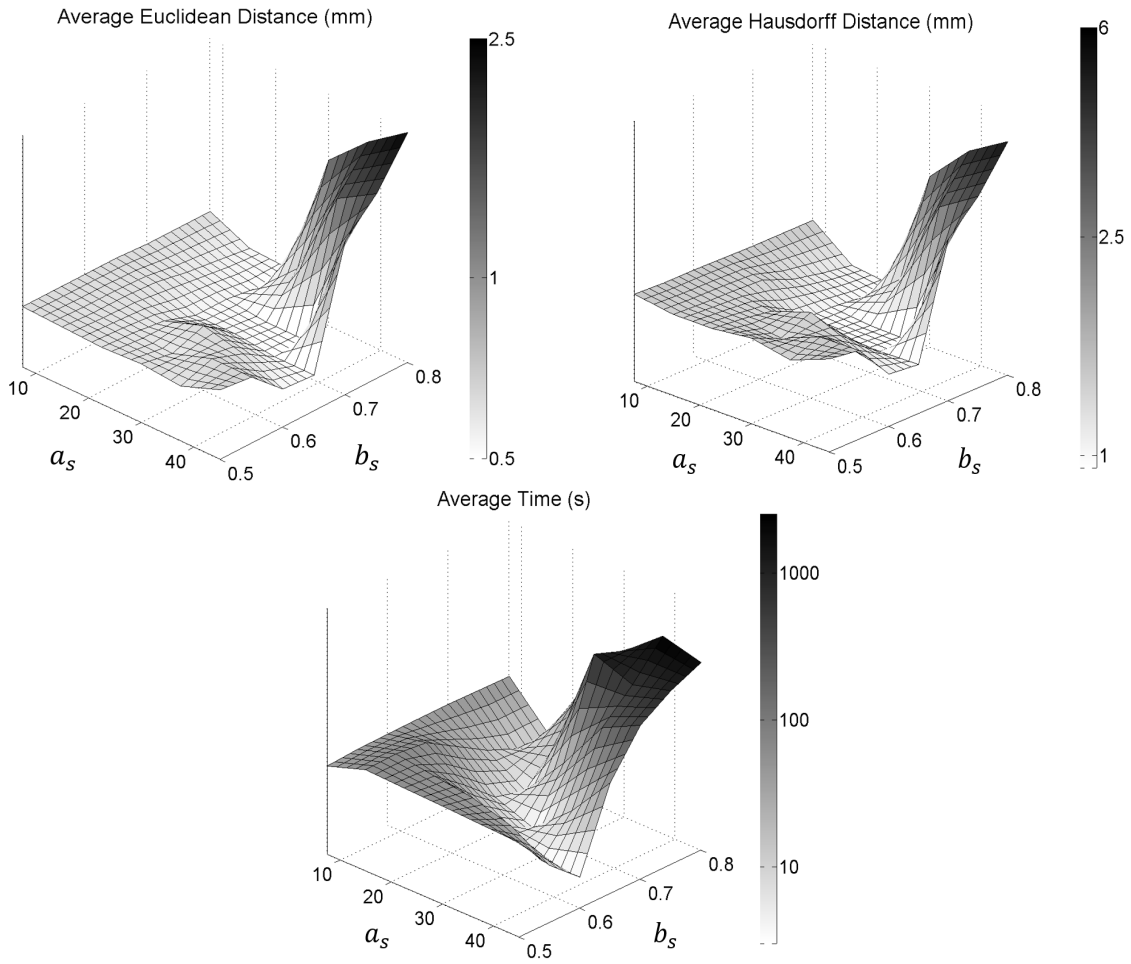


Figure 4.9: Manifolds showing the average A* performance when using Eq. 4.8, and how it varies according to the parametrisation (a_s, b_s). Logarithmic scales were used for the sake of clarity.

in Table 4.2 were the ones reaching better trade-off concerning path detection accuracy and time expended. In a clinical setting like the one described in this Section, where the manual analysis of the data may easily reach a couple of hours, having a semi-automatic algorithm that takes a dozen of seconds to detect an intramuscular path is not problematic. Even then, our methodology was able to reach very interesting compromises. For example, the parametrisation (a_s = 45, $b_s = 0.60$) was able to attain one of the best path accuracies (Euclidean and Hausdorff distances of 0.51 ± 0.14 mm and 1.00 ± 0.39 mm) and also be very fast (3.50 ± 6.5 s). For reference, an Intel Core i7-4500U CPU @ 1.80 GHz 2.40 GHz with 8 GB of RAM was used. An example of an extracted intramuscular path using this configuration is present in Figure 4.10.

The previous version of the algorithm had average Euclidean and Hausdorff distances of, respectively, 1.06 ± 0.32 and 2.44 ± 0.92, taking an average of 15.00 ± 14.76 s. Therefore, for appropriate parametrisations, the proposed minimum cost path approach is able to extract the optimal intramuscular pathways at sub-voxel accuracy, and taking little time to do so. This makes us more confident that these algorithms are suitable to be incorporated into a software aiming to

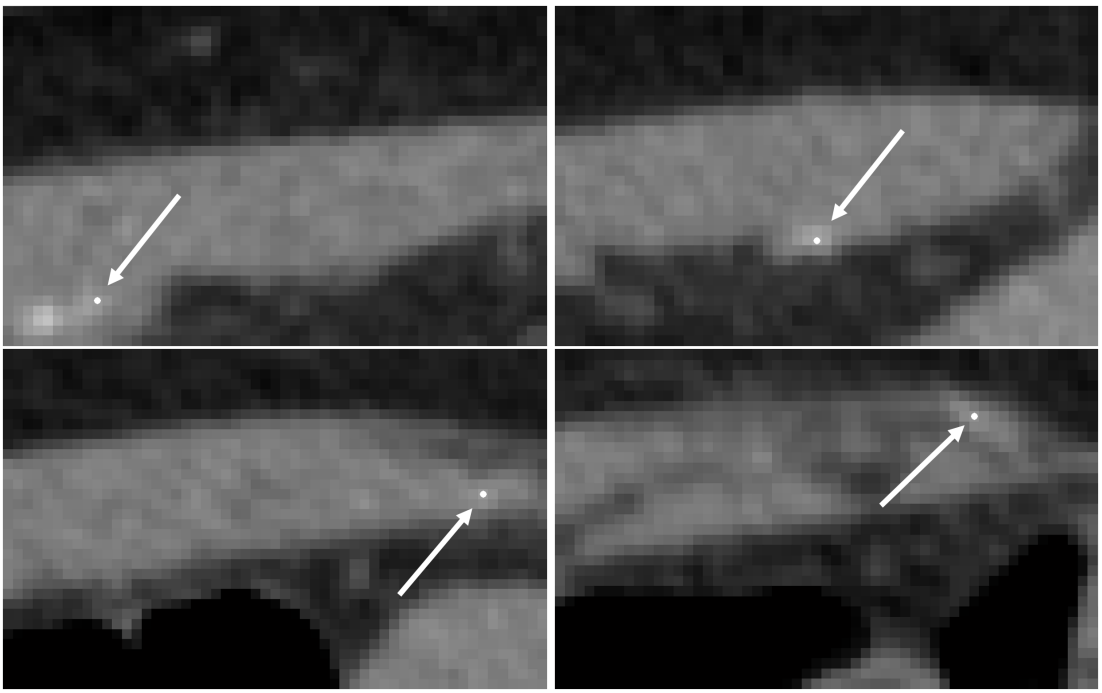


Figure 4.10: Example intramuscular course extracted by the proposed minimum cost path method, for $a_s = 45$ and $b_s = 0.60$.

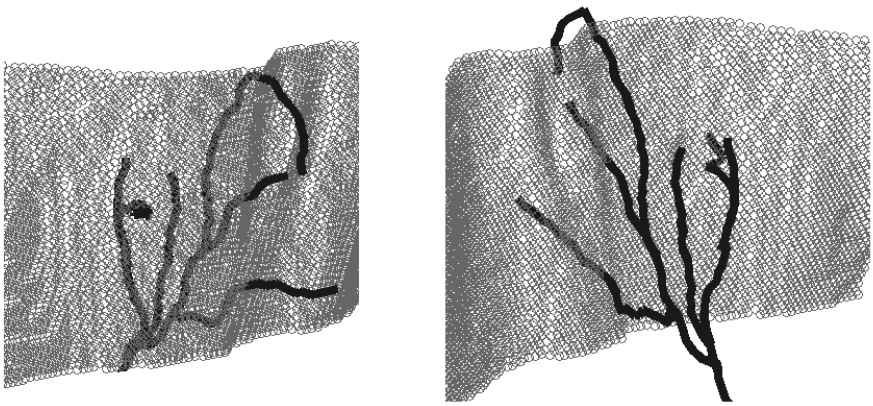


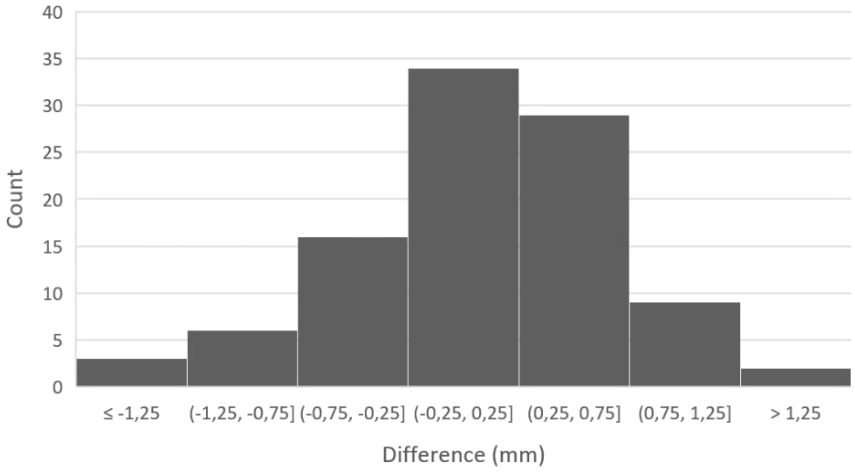
Figure 4.11: 3D representation of the anterior fascia of the abdominal muscle and the extracted DIEP tree from one hemiabdomen.

support the DIEP flap preoperative planning task. The detection of the perforators is also a step towards the efficient creation of 3D models which may be of great relevance to the surgical team. In Figure 4.11, we show a representation of one of the DIEP trees extracted by the methodology presented here.

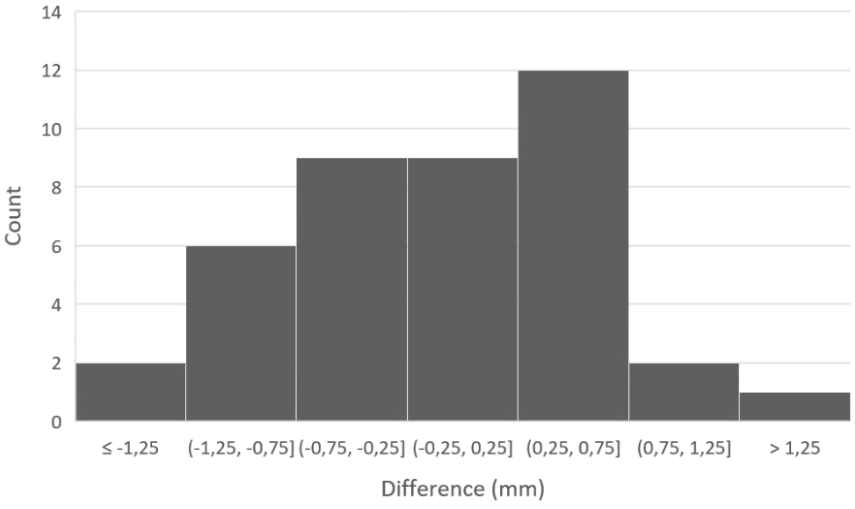
4.1.5.2 Perforator characterisation

From the 40 CTAs available in the considered dataset, 180 perforators have been manually identified pre-operatively by the experts, and 183 by the user running the proposed algorithm. A total of 234 vessels were confirmed intra-operatively (ground truth). From those, a total of 129 vessels were identified simultaneously by both the manual and semi-automated methods.

A statistically significant difference was detected when estimating calibres using the manual and software methods, $p \approx 1e{-}5$ (128 degrees of freedom), and the effect size was medium, $r = 0.39$. However, the paired sample t-test revealed no statistically significant difference when estimating the calibres of the perforators selected for the flap using the manual and software methods, $p \approx 0.44$ (40 degrees of freedom), and the effect size was small, $r = 0.12$. Figure 4.12 shows



(a)

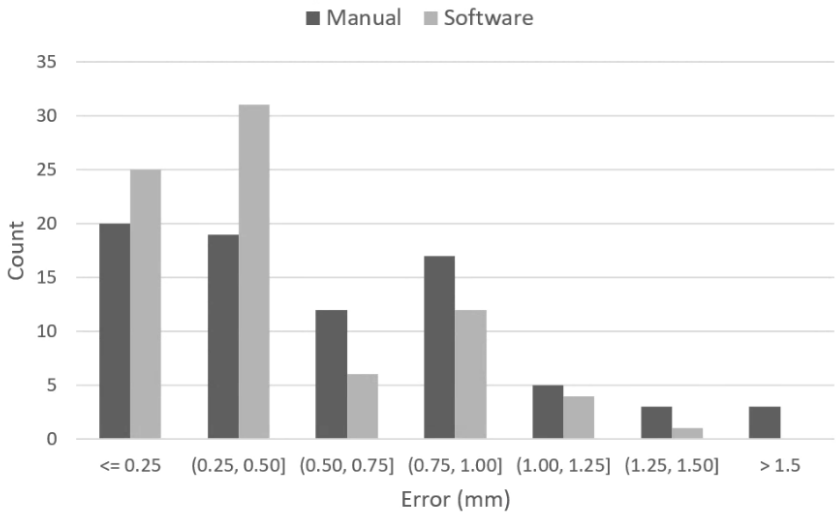


(b)

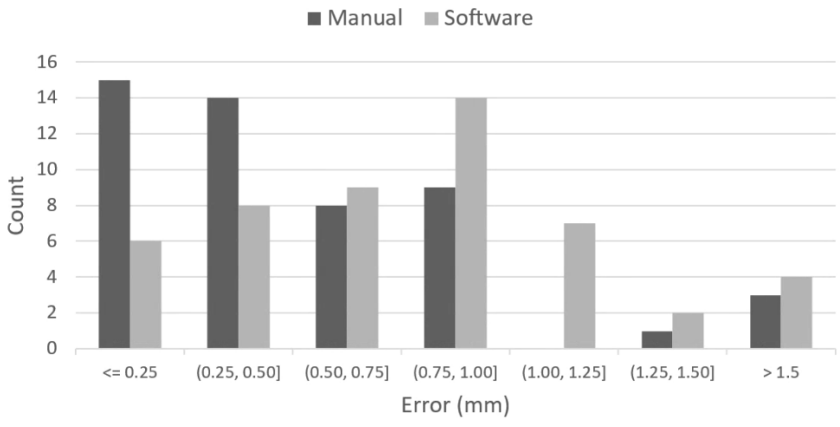
Figure 4.12: Calibre estimation differences between the software and manual reporting when considering: (a) all perforators simultaneously detected by both methods; (b) only the perforators included in the flap.

the distribution of the differences between the calibre estimates of both methods when considering all the perforators or only the ones that ended being included in the flaps.

A statistically significant reduction of the median absolute error of estimated calibres was obtained by using the software, regarding perforators having calibres larger than 1.5 mm, $p \approx 2e{-}3$, with a medium effect size, $r = 0.26$. However, using the software led to a statistically significant increase of the median absolute error when estimating the calibre of perforators having a calibre less than or equal to 1.5 mm, $p \approx 6e{-}4$. The effect size was medium, $r = 0.34$. Figure 4.13 shows the distribution of the absolute errors of both methods when comparing with the intra-operative measures of, respectively, perforators having calibre larger than 1.5 mm, and the ones having calibre less than or equal to 1.5 mm.



(a)



(b)

Figure 4.13: The distribution of the absolute error of calibres estimated by the automated and manual methods when compared with the reference surgical findings: the distribution for perforators with (a) calibre larger than 1.5 mm, and for the ones having (b) calibre less than or equal to 1.5 mm.

Regarding the software estimates of where the perforators pierce the fascia, horizontal error was not statistically significant, $p \approx 0.09$ (170 degrees of freedom), and the effect size was small, $r = 0.13$. However, vertical error was statistically significant, $p \approx 0.02$ (170 degrees of freedom), even though the effect size was also small, $r = 0.18$. Figure 4.14 shows a representation of the differences between the location estimated by our methodology and the manual findings. Regarding error in height, the average absolute error was 3.2 ± 2.4 mm, whereas the horizontal average absolute error was 2.5 ± 2.0 mm.

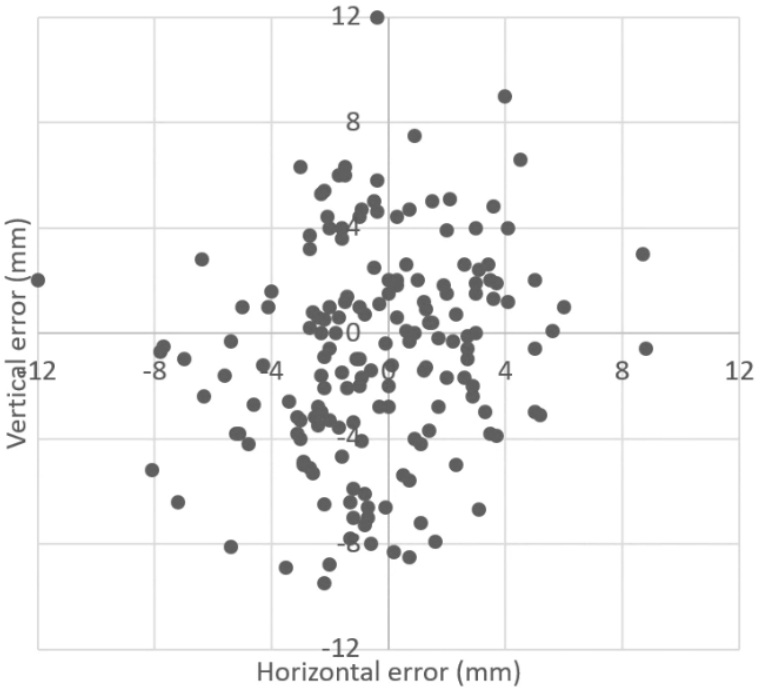


Figure 4.14: Comparison of the automated method with manual analysis when estimating the location where the perforator pierces the anterior fascia. Each point represents for a given perforator the vertical and horizontal error when considering the manually retrieved location as the ground truth.

The results allow to perceive some merits but also limitations of the semi-automatic approach when compared with the manual analysis. It was more adequate at measuring the calibre of larger perforators (≥ 1.5 mm) but also shown a tendency to overestimate the calibre of the smallest ones. Regarding the location where the perforators pierce the fascia, even though there was a small error regarding vertical position (2-3 mm), it was not relevant in clinical practice, since perforators are approached carefully during dissection in order to prevent damage.

The majority (88%) of the perforators included in the flaps were co-identified by both methods, showing that the results of the semi-automatic methodology can be an option for this particular scenario. Despite having very challenging SNR, the DIEP vessel tree has a relatively simple network such that centreline-based approaches are the framework of choice amongst traditional

algorithms for blood vessel segmentation. Additionally, such methodology also facilitates the local characterisation of blood vessels, which is essential in this scenario. Although the proposed pipeline is not completely automatic, the effort required by the user is significantly reduced. The manual analysis takes between 2-3 h to be conducted, whereas the semi-automatic approach usually takes around 30 min. Moreover, these algorithms make it easier for the user to interact with them in case errors are spotted. As they output a sequence of vessel centreline points, it is easy to manually manipulate the output to correct eventual mistakes.

These results are promising in the sense that the semi-automatic methodology was capable of delineating and characterising the perforators with an adequate performance, reducing significantly the operator-dependent analysis and many of the time-consuming steps (gaining at least 2 h per patient).

4.1.6 Prototype software

The promising results obtained in the aforementioned experiments motivated us to design a preliminary prototype that shall be included in the clinical pipeline and undergo further validation. The aim will be to gather input from clinicians about the usability of such tool and other nice-to-have functionalities.

The interface of the prototype is shown in Figure 4.15, where the data of a patient has already been loaded. As can be noticed, this version of the prototype already includes the required functionality to conduct the main steps of the algorithm, such as defining the volume of interest, automatically finding the fascia in that volume, and also conducting the perforator extraction step. In addition, controls to select which detections are visible and also to clear previous findings are available.

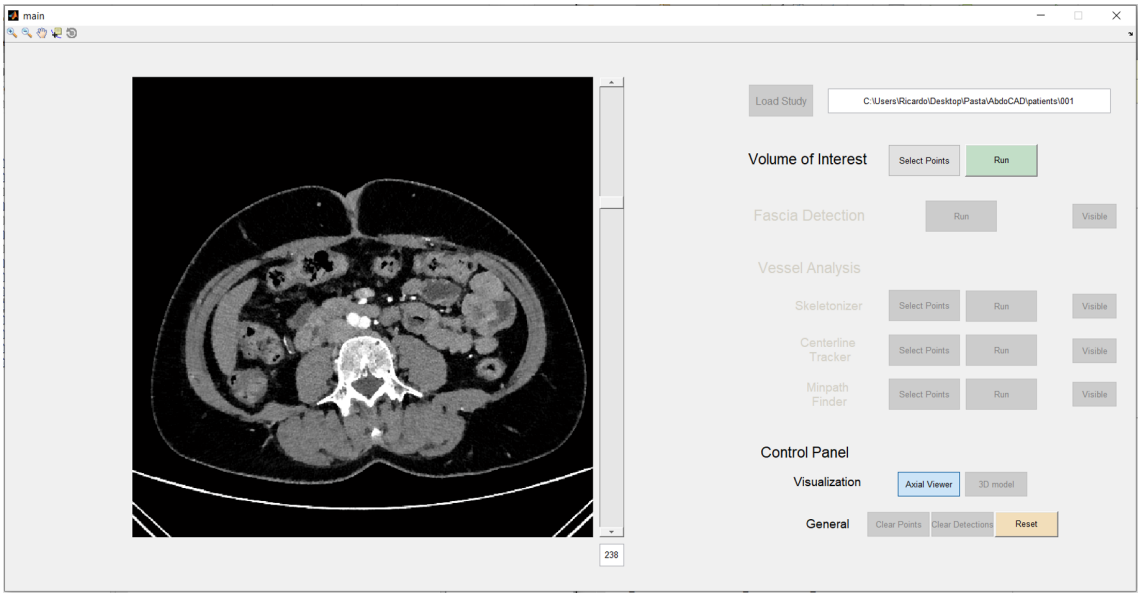


Figure 4.15: Main prototype interface showing the data of a patient.

After the volume of interest has been defined, the visualisation panel shows only the intensity data at the region of interest, as shown in Figure 4.16. This allows the clinicians to more easily focus on the structures of interest while scrolling through the data.

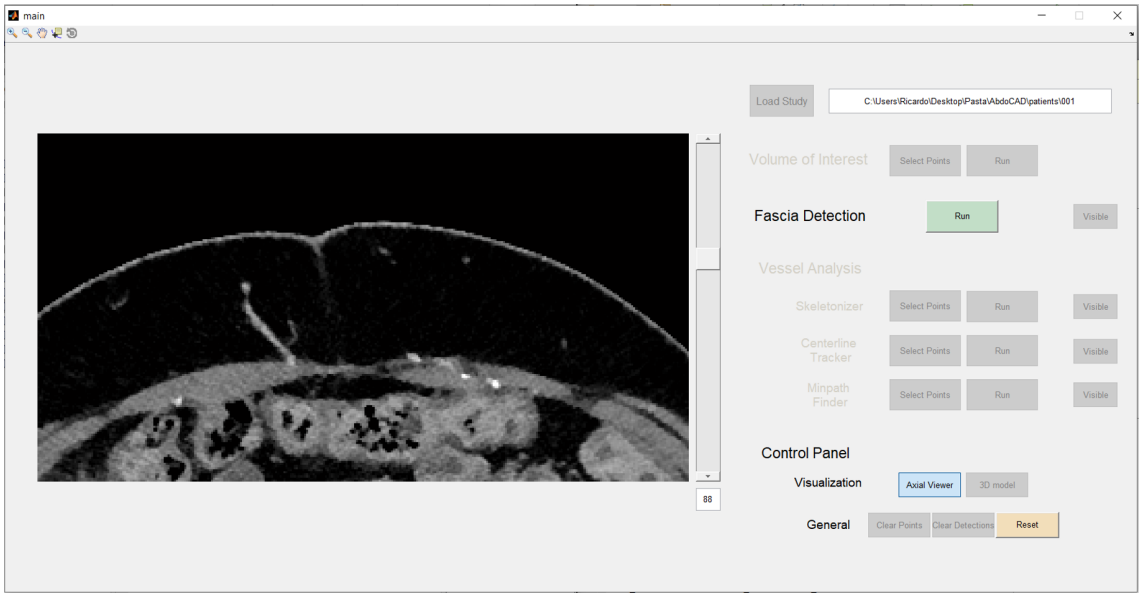


Figure 4.16: Main prototype interface after selecting the volume of interest.

The detection of the fascia is initiated with a simple button click. As soon as the fascia is detected, it is possible to conduct the perforator extraction step, after the required user landmarks have been defined. Figure 4.17 illustrates what is seen in the visualisation panel after conducting fascia and perforators detection.

We provide a 3D visualisation option, due to the relevance of observing how the different structures are arranged in a 3D environment. This is one of the main limitations of the current manual pipeline, as only delivering a report (an example is provided in Figure 4.18) and observing 2D slices limits 3D comprehension. The prototype allows to render the extracted fascia layer and perforators in a 3D environment, being possible to control the zoom and rotate the data. Figure 4.19 demonstrates different perspectives of a particular case.

In the future, we would like to include other important functionalities in the prototype, such as describing the characteristics of the perforators (calibre, tortuosity and general course), providing an automated reporting of the findings, and improving the 3D visualisation options. It would be interesting to detect and show other neighbouring structures, in order to better depict the local anatomy and provide more relevant landmarks. Additionally, we would like to explore more automated ways of characterising the subcutaneous part of the perforators, which seems the most plausible to do so, as the SNR at that region is not as low as the more challenging intramuscular one. For that, we plan to use the volume of vessel enhanced data and consider a segmentation based on hysteresis thresholding. Some preliminary experiments lead to interesting results and we would like to further explore this idea. As shown in Figure 4.20, this algorithm is capable

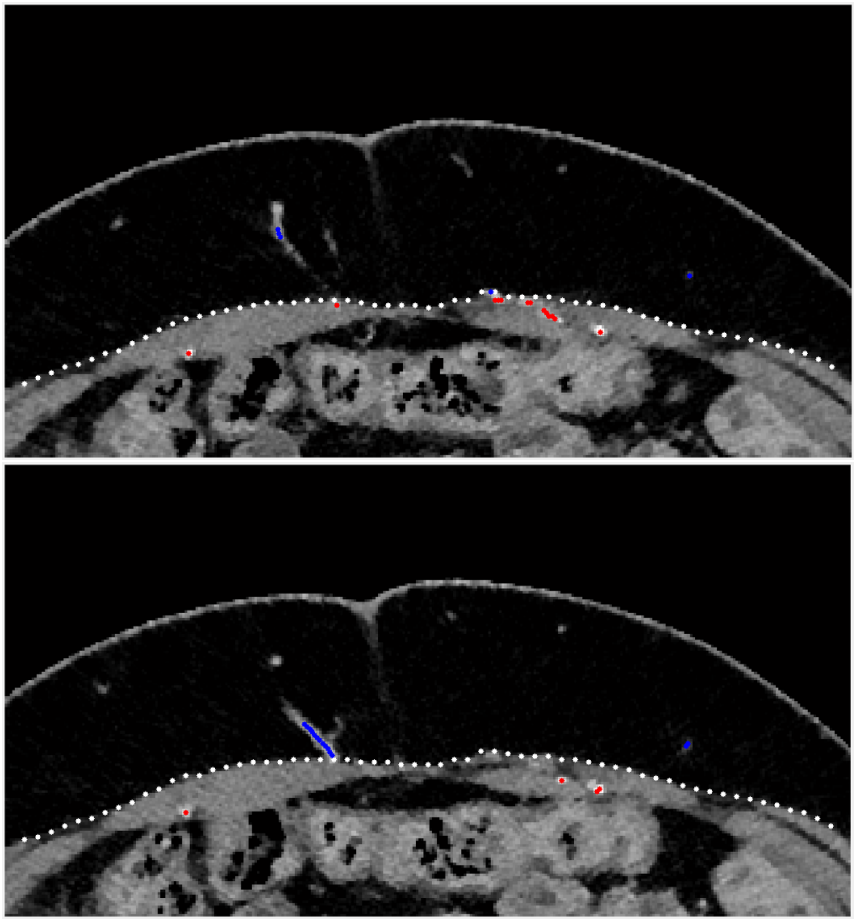


Figure 4.17: Detection of the fascia (white points) and the centrelines from the subcutaneous (blue points) and intramuscular (red points) courses of perforators.

Perforator number	Subcutaneous orientation	Calibre (mm)	Localization (mm)	Intramuscular course
1	Evolves slightly downwards	2.8	24 to the right 4 below	15 mm length; slightly tortuous
2	Evolves downwards	2.4	31 to the left 7 below	Not opacified
3	Evolves slightly upwards and slightly laterally	2.0	34 to the left 14 below	Not opacified
⋮				
n	Evolves laterally	1.6	60 to the right 42 below	10 mm length; linear

Figure 4.18: An example report of the manual analysis of the DIEPs.

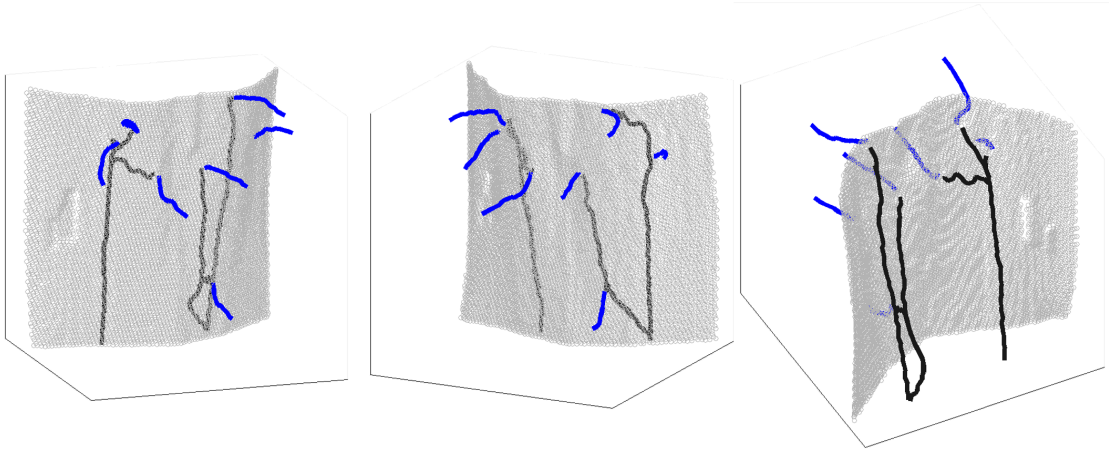


Figure 4.19: 3D visualisation of the extracted fascia layer and perforators. Blue and black lines represent, respectively, the subcutaneous and intramuscular courses of the perforators.

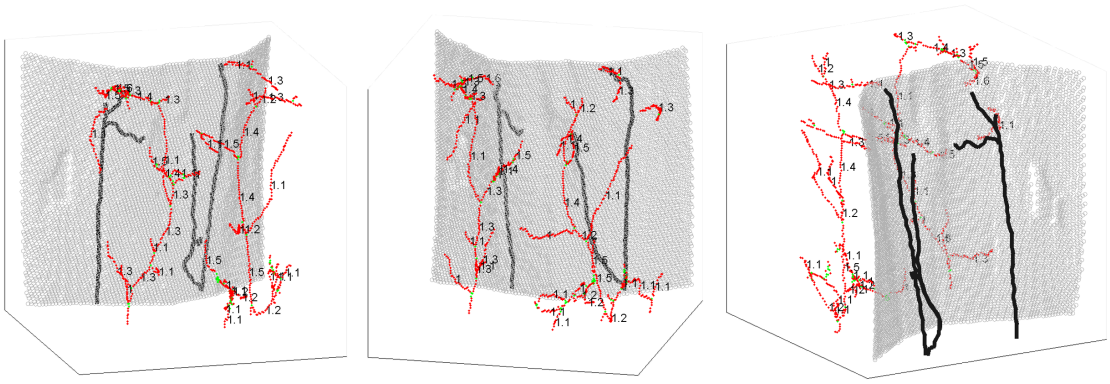


Figure 4.20: 3D visualisation of the results obtained when using the automated approach for subcutaneous course extraction. Red and black lines represent, respectively, the subcutaneous and intramuscular courses of the perforators.

of extracting the more complex subcutaneous structure of the perforators, including ramifications and anastomoses between different perforators. This allows the surgical team to better understand the volume each perforator vascularises and lead to better preoperative planning when picking the optimal perforators to include in the flap.

4.1.7 Summary

In this Section we proposed an improved version of our pioneer semi-automatic algorithm for the extraction and characterisation of the DIEP vessel tree, which is essential during the preoperative planning of breast reconstruction through the DIEP flap procedure.

A first set of experiments allowed to conclude that the proposed modifications improved the extraction of these blood vessels, both in terms of accuracy and time expended. A second experiment in a larger version of the dataset looking to compare the findings of the proposed method with

those retrieved manually by the radiological team was conducted. Regarding calibre estimations, there was a statistically significant difference between both methodologies, however, when considering only perforators which ended being included in the flap, no statistical significant difference was found. Having the calibres measured during surgery as reference, the automated method showed smaller median error in larger perforators (the most commonly included in the flap) yet a larger median error in the smaller ones, when comparing with the manual analysis. With regard to the estimation of the location where the perforators pierce the anterior fascia, having the manual findings as reference, the automated method had no statistical significant horizontal error, however a statistical difference in vertical error was found. Nonetheless, the effect size was small and it does not have much impact in practice.

These experiments are promising and support our conviction that it is feasible to design a CAD algorithm to support the clinicians who are responsible for the preoperative analysis of the DIEPs, leading to more objective and fast results. A prototype software providing a simple interface to conduct our proposed methodology was created. In the future, we want to handle this software to the Breast Unit of Champalimaud Foundation in order to acquire input on its usability and suitability, and also further validating the algorithm in a more clinical context. Augmented reality systems would be interesting to consider in this scenario, given that the simultaneous visualisation of the data and the extracted DIEP tree would give a better perception to the surgeons of the 3D arrangement of the structures of interest.

4.2 Retinal fundus vessels

The retina is a tissue layer in the eye of vertebrates that participates in the production of nerve impulses that go to the visual cortex of the brain. Its vascularisation is easily assessed in a non-intrusive manner by photography-based mechanisms, such that fundus imaging is often used as a diagnostic means of medical conditions affecting the morphology of vessels, such as hypertension, diabetes, arteriosclerosis, and cardiovascular disease [181]. A more detailed discussion on the relevance of this topic can be found in Chapter 1.

The manual analysis of blood vessels by experts is a very time consuming and tedious process. This becomes even more evident in scenarios where large volumes of data are collected in a short time lapse, as is the case of diabetic retinopathy screenings [182]. Hence, CAD algorithms for the segmentation, analysis, and identification of relevant findings are highly welcome, in order to support the activity of these clinicians.

Despite the success of early DL approaches applied in the scenario of retinal vessel segmentation [153], they are quite slow ($\approx$ 90 s per image) during prediction as they must divide an image into a set of patches and classify each of them to obtain a segmentation map. Even though that time may not be prohibitive in many applications, it can quickly induce a bottleneck during a screening programme, where many people are assessed and, ideally, results should be quickly obtained.

In this Section, we describe the supervised FCN we have proposed to automatically segment the retinal vessels at a single step, even when trained in a patch-wise fashion (see Figure 4.21). Our

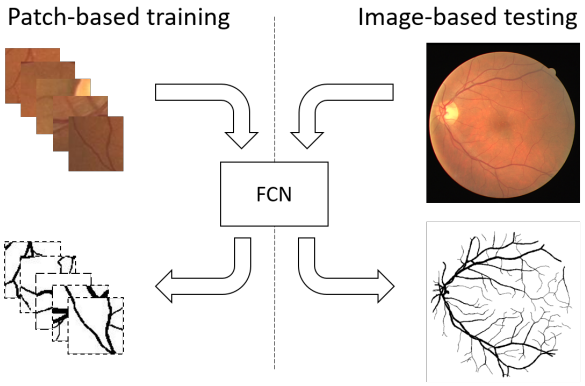


Figure 4.21: FCNs take images of arbitrary size, allowing to combine patch-based training and image-based prediction.

design strongly takes into account the time efficiency at the prediction phase, given that we seek an automated model that is adequate for screening programs. In addition, we also avoided employing pre- and post-processing steps, which are commonly taken into account in the literature. Note that by pre-processing we refer to methods such as Contrast Limited Adaptive Histogram Equalisation and the use of Wavelet transforms, not to data normalisation. Our motivations are three-fold:

- in theory, a large number of non-linear functions is capable of modelling the more irregular structure of raw data, even though, in practice, an adequate pre-processing tends to facilitate the learning process, especially in scenarios where the availability of data is more limited. Nonetheless, it is not completely clear how this translates to each application and network architecture, and whether it can be mitigated by employing adequate data augmentation.

- most pre- and post-processing steps introduce a further significant burden in terms of computation time, adding up to the total time expended in the prediction step;

- it becomes difficult to understand in the end how each module is contributing to the final performance obtained if a proper experimental setup is not conducted. This is especially true when using a particular Deep Neural Network (DNN) design and also post-processing, and only reporting the performance of the complete pipeline, which does not allow to grasp the capabilities of the network itself.

Following the discussion above, we use raw color fundus images in our experiments, in order to understand if our network design is able to reach state-of-the-art performance and, simultaneously, keep the prediction process as simple as possible. A FCN was proposed in the past [156], however its performance is significantly inferior to the best performing methods, indicating that other specific network design options may not have been ideal for retinal vessel segmentation.

4.2.1 Single-resolution Fully Convolutional Network for Vessel Segmentation

Here, we discuss the motivations and preliminary empiric findings that led us into designing a FCN adapted to the specific task of vessel segmentation in raw color fundus images.

4.2.1.1 Fully Convolutional Networks

CNNs have revolutionised the field of computer vision, given their combination of deep hierarchical feature extraction (sequence of convolutional layers) and classification (FC layers) blocks. This was the type of deep neural network used in [153], where very small patches of the retina were fed into the model and it outputted the probability of the centre pixel being a vessel. This highlights one of the problems of using typical CNNs for segmenting vessels, which is the need to divide a given image into a very large number of small patches and classify each of them, yielding a tremendous computational cost. A second problem is that fully connected layers force all the input images to have the same size.

A FCN design is a more adequate choice for segmentation problems, since it does not use fully connected layers. Thus, it is not mandatory to divide an image in order to obtain a complete segmentation map, which is crucial whenever we require fast predictions, as is the case of retinal screening programs, where a high volume of data is quickly generated. The inputs may also have varying size, making this design much more adaptable to different imaging conditions. It allows us to train on smaller patches of the images and later still be able to obtain single-pass predictions of the entire images, as is represented in Figure 4.21. Note that performing patch-wise training is an engineering option which facilitates avoiding wasting computational effort with portions of the images that do not contain information of the retina fundus, and helps managing the available memory.

4.2.1.2 Specific design considerations

After motivating the use of a FCN design for the segmentation of retinal vessels, now we delve into more specific aspects of the proposed network architecture, discussing some options we took based in previous works and empirical findings.

Spatial resolution Pooling or strided convolutions are commonly used to induce higher-level features to encode more neighborhood information. Recent results [153] suggest that pooling operations do not improve the performance of networks that are trained in small images. In preliminary experiments, we found that indeed a single-resolution deep network was more capable than a U-Net shaped model when extracting small capillaries. Even though the latter is able to combine low- and high-scale features, it seems that a deeper network at a fine scale is able to obtain better representations of small structures of interest, as is the case of very small vessels. Thus, in this work, the image resolution was kept across the entire network, contrarily to the previously proposed FCN of Dasgupta and Singh [156].

Activation units All intermediate non-linearities were given by a Leaky ReLU:

$$f(x) = \begin{cases} x & \text{if} \quad x > 0, \\ ax & \text{otherwise} \end{cases} \tag{4.10}$$

where x represents the outcome of the previous convolution and a was set to 0.2. It was used over a ReLU just to allow the network to learn even for negative inputs. In the last layer, we used a Sigmoid activation unit, since we are dealing with a pixel-wise binary problem.

Batch normalisation Whenever the statistics at test time differ from the ones found during training, batch normalisation becomes problematic. In fact, this is the case when a model is trained in small retinal patches and, during prediction, is applied to entire retinal images, whose statistics will be inevitably different. In preliminary experiments, we found that using batch normalisation was indeed hurting the performance of the models, thus it was not considered in the final design.

Dropout Turning off some computational connections along the network was useful to create more redundancies and thus obtain more robust models. We found it was also useful to apply dropout at the initial levels of the model, in order to add some noise to the initial representations.

Loss function Neural networks targeting binary segmentation problems usually minimise the BCE loss, a pixel-wise criterion that exponentially increases as the network becomes more confident when committing a mistake. Notice however, that this loss is agnostic to class imbalance, thus it naturally biases models to be more confident identifying the most common class, which in our case, is the background. We are interested in alleviating this effect, in order to obtain models with good sensitivity and that do not simply ignore narrow vessels. Weighting differently each class was an option we considered for reaching fairer models. Furthermore, we used the recently proposed Focal Loss (FL) [183], an extension to the BCE loss that puts more focus in the misclassified examples:

$$L_{focal}(p, y; \alpha, \gamma) = -\Big(y \cdot \alpha(1-p)^{\gamma} \cdot \log(p) + (1-y) \cdot (1-\alpha) \cdot p^{\gamma} \cdot \log(1-p)\Big) \tag{4.11}$$

where $p \in [0,1]$ is the probability of class 1 (vessel) outputted by the network, $y \in \{0,1\}$ is the binary target variable, $\gamma \geq 0$ is a focusing parameter, and $\alpha \in [0,1]$ allows to give more weight to samples of a certain class. γ was set to 2 in this work. Even though the FL by itself is also agnostic to class imbalance, by performing hard training, it helps inducing the model to not ignore the potential hardest cases, such as small capillaries.

After all these considerations, architecture and hyper-parameter tuning was conducted (see Section 4.2.2.2). The final design we considered for segmenting vessels from raw color fundus images is represented in Figure 4.22.

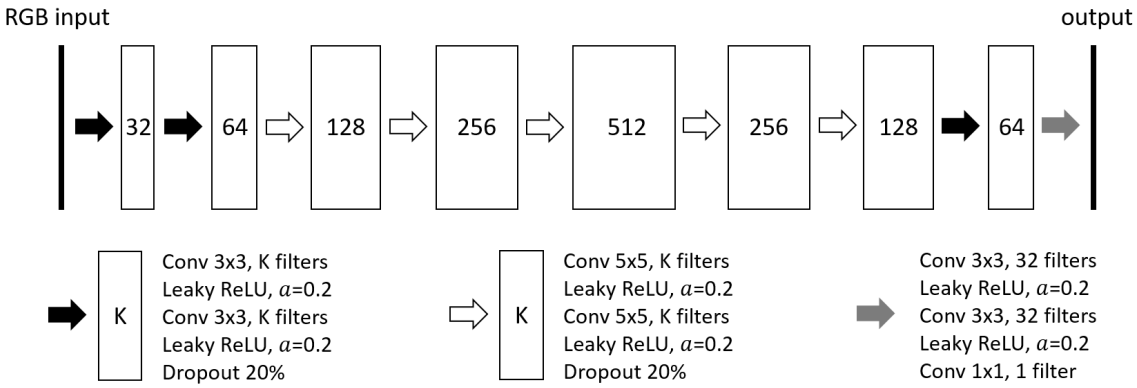


Figure 4.22: Single-resolution fully convolutional network proposed for segmenting vessels in raw fundus images.

4.2.2 Experiments

The model is evaluated on the retinal fundus datasets described in Section 2.1. The metrics used to assess its performance will be briefly described here, and then we provide details regarding how hyper-parameter tuning was conducted to obtain the final neural network design.

4.2.2.1 Model Evaluation

To evaluate how well a map of vessel probabilities fits the ground truth, we calculated the metrics that are commonly used in this task, which are accuracy, sensitivity, and specificity:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{4.12}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{4.13}$$

$$Specificity = \frac{TN}{TN+FP} \tag{4.14}$$

where TP, FP, FN, and TN are the true positive, false positive, false negative, and true negative detections. A limitation of these metrics is that they are evaluated at a threshold of 0.5. Thus, we also considered the commonly used AUC, which we believe is more ideal for this task, as it better depicts how well a method separates both classes in different operating points.

4.2.2.2 Implementation details

The architecture and hyper-parameters were tuned by randomly picking three images from the DRIVE training set for validation purposes and using the remaining ones to train varying model configurations, according to the considerations detailed in subsection 4.2.1.2. Colour images were solely normalised to the range $[0,1]$. At each training epoch, 500 batches of N patches of size

$M \times M$ were fed to the network. Patches were randomly extracted from images at valid positions, where valid means the centre pixel belongs to the retinal fundus. Data augmentation was conducted via random transformations including vertical or horizontal flipping, and rotations in the range $[-\pi/2, \pi/2]$. We used the Adam optimiser with the parameters as provided in the original work [184], with the exception of the learning rate, which was initialised to $1e{-}4$ and decreased to half every time the validation loss did not decrease for 10 epochs. A loss decrease was only considered if it surpassed the threshold of $1e{-}4$. Early stopping occurred if there were 30 epochs without improvement. Our preliminary experiments achieved best performance in the validation set using the network design present in Figure 4.22, and for $N = 16$ and $M = 64$, even though these hyper-parameters did not have a significant impact in the performance of the model.

We trained our final FCN design for 30 epochs. Starting from epoch 10, we performed linear learning rate decay by multiplying it by a constant of 0.75, and after epoch 20 the constant was changed to 0.5. Concerning DRIVE, we trained the network in the 20 images of the training set and evaluated it in the 20 images comprising the test set. Regarding STARE and CHASEDB1, datasets with few images and where a prior division does not exist, we followed the same approach of other researchers [153], which resorted to the leave-one-out validation.

4.2.3 Results and discussion

The results obtained by conducting the described methodology in the referred databases are present in Table 4.3, along with the performance of the most relevant state-of-the-art approaches. A brief description of each of the works being compared is present in Chapter 3. It is important to note that the method of Azzopardi et al. [68] belongs to the unsupervised category, and that the work of Fraz et al. [139] uses traditional machine learning. The rest of the methods included use deep learning techniques.

The analysis of the results shows that our FCN design is able to combine efficiency and strong predictive capabilities, even when using raw fundus images and not considering pre- and post-processing. By comparing the AUC of the methodologies, it is possible to conclude that the proposed methodology achieved competitive performance in all the databases, even though it is not the top performing model at any of them. It is interesting to see that it is not possible to pick a single best method, since the best-performing ones are different among the considered datasets. In fact, some of these have lower performance than the proposed design in the datasets where they do not reach state-of-the-art performance. Among the databases, our method deviated more from the best-performing one in the STARE dataset. We believe this is due to the high variability of the raw color information among the images of this database. This may indicate that preprocessing techniques leading to more uniform images are relevant here. Regarding DRIVE, we also tested $\alpha = 0.6$ (give more weight to the vessel class) to better show the compromises we can get between sensitivity and specificity. Note that by varying α, we could easily achieve models with very high sensitivity or specificity, thus we emphasise that it is the compromise that is relevant. Besides, this shows that the AUC metric is the most adequate to inspect the true model's capacity to distinguish both classes. We did not conduct this experiment in the other databases, since the number of

Table 4.3: Performance of the proposed methodology and state-of-the-art approaches in the DRIVE, STARE, and CHASEDB1 databases. Accuracy, sensitivity and specificity are abbreviated as acc, sen, and spe, respectively. Best results are highlighted in bold.

Method	**DRIVE**				**STARE**				**CHASEDB1**			
	AUC	acc	sen	spe	AUC	acc	sen	spe	AUC	acc	sen	spe
Azzopardi et al. [68]	96.1	94.4	76.6	98.1	95.6	95.0	77.2	97.0	94.9	93.9	75.8	95.9
Fraz et al. [139]	97.5	94.8	74.1	98.1	97.7	95.3	75.5	97.6	97.1	94.7	72.2	97.1
Liskowski and Krawiec [153]												
balanc.-SP, $s = 3$	97.9	95.1	84.6	96.7	**99.3**	96.7	**92.9**	97.1	98.2	94.4	**91.6**	94.7
balanc.-SP, $s = 5$	97.9	95.3	81.5	97.5	**99.3**	97.0	90.8	97.7	98.4	95.8	87.9	96.7
no-pool-SP, $s = 5$	97.9	95.4	78.1	98.1	**99.3**	**97.3**	85.5	98.6	98.2	96.3	78.2	98.4
Oliveira et al. [160]	98.2	95.8	80.4	98.0	99.0	96.9	83.2	98.6	98.6	96.5	77.8	98.6
Zhang and Chung [158]	98.0	95.0	**87.2**	96.2	98.8	97.1	76.7	99.0	**99.0**	**97.7**	76.7	**99.1**
Jiang et al. [130]	98.1	96.2	75.4	98.2	99.0	**97.3**	83.5	98.5	98.1	96.7	86.4	97.4
Jin et al. [159]	**98.6**	**97.0**	78.9	**98.7**	98.7	**97.3**	74.3	**99.2**	98.6	97.2	82.3	98.2
Proposed												
$\alpha = 0.5$	98.2	95.6	80.3	97.9	98.7	96.5	82.9	98.0	98.6	96.5	82.1	98.1
$\alpha = 0.6$	98.2	95.4	85.0	96.9	-	-	-	-	-	-	-	-

models that are trained in a *leave-one-out* validation setting is very high. The use of FL over BCE lead to an improvement of 0.2 percentage points regarding the AUC metric, when evaluating the system in the DRIVE database for $\alpha = 0.5$. The other metrics did not significantly change with this loss, meaning that it mostly induced the system to become slightly more confident on its predictions. Then, this seems to support that the single-resolution deep architecture was the main reason for our system to significantly outperform the FCN proposed in [156]. Figure 4.23 shows the best and worst predictions outputted by the proposed methodology for the considered databases, regarding AUC. It is possible to visualise that the model is able to cope with challenging imaging conditions, and even with the presence of severe pathology (4th row of Figure 4.23).

Using a Nvidia GeForce GTX 1080 Ti GPU, it took us 2.1, 2.7, and 6.5 s to make a prediction for an image in DRIVE, STARE, and CHASEDB1 databases, respectively. The method of Liskowski and Krawiec [153] takes on average 92 s using the Nvidia GTX Titan GPU. Even though the GPUs are not identical, this strongly suggests that our method is significantly faster, thus being more adequate for real-time applications. The other deep learning methods in Table 4.3 also report that it only takes a couple of seconds to conduct prediction, however it is not clear whether this time includes the pre- and post-processing steps.

4.2.4 Summary

In this Section, we described the motivation behind using a FCN for the particular case of retinal blood vessel segmentation, and discussed the design options we took to do it using raw retinal fundus images. The proposed network is more convenient and efficient than some of the state-of-the-art approaches, as it allows to make predictions for unseen images of different sizes at a single step and avoids time-consuming pre- and post-processing steps, traits that become relevant in screening scenarios.

Even though the proposed method did not surpass the best-performing state-of-the-art methods in any of the considered datasets (considering AUC), it reached competitive performance. It is important to refer that the best performing methods in each dataset are different, and our algorithm surpasses some of them in the other datasets. This shows that the considerations we made do not necessarily compromise the performance in this task, as the proposed design competes with the state-of-the-art approaches. Among all the datasets, it was in STARE that our methodology deviated more from the best performing method. In this dataset, the raw images are significantly different from each other, such that an adequate pre-processing may indeed be helpful to achieve better results. Thus, for future work, it would be interesting to study cost efficient pre-processing techniques, in order to avoid introducing a relevant bottleneck during prediction time.

4.3 Main Contributions and Final considerations

This Chapter described the work conducted in two very distinct applications. First, the segmentation of the DIEP vessels in CTA scans, a problem involving the analysis of 3D data and where there is no publicly available datasets with labels. This motivated the design of a more traditional

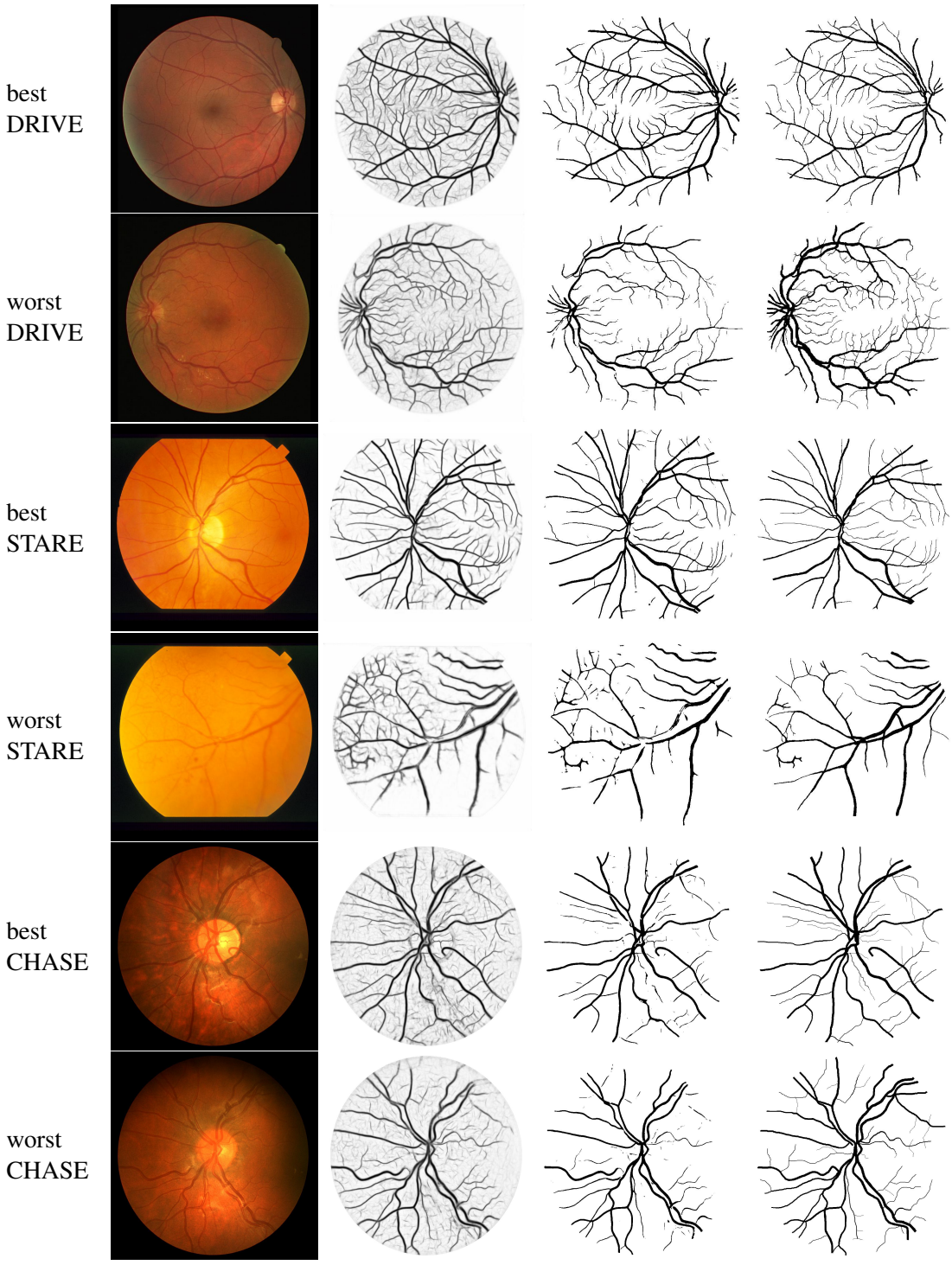


Figure 4.23: Best and worst results for each database, concerning the AUC metric. From left to right, raw color fundus image, probability map outputted by the proposed methodology, segmentation obtained by thresholding probabilities at 0.5, and ground truth. Notice that the blood vessel masks are inverted for visualisation purposes.

computer vision pipeline to address the problem. Our pioneer work on this topic started during the MSc thesis [24] and was improved during the PhD. In addition, a first set of clinical validation was conducted in order to better understand the potential of the framework.

Regarding the second scenario, the segmentation of retinal blood vessels in fundus photos, 2D data is targeted instead and there are several available databases including gold standard annotations which can be used for supervised machine learning research. We explored the efficient FCN architecture and showed that our particular design choices were suitable for this task, even without conducting pre- and post-processing steps.

These works were developed having specific applications in mind, such that the design of the methodologies was inevitably influenced by the particularities of each of them. In the case of DIEPs, the prior knowledge concerning the local anatomy and the significantly different SNR of their subcutaneous and intramuscular parts motivated us to employ a divide-and-conquer methodology to simultaneously consider accuracy and efficiency. The little amount of data has favoured the employment of a more traditional-based computer vision algorithm. Even though the complete pipeline was specifically designed for the DIEP extraction, it is still expected that its single components are useful in other similar scenarios, such as the extraction of perforating vessels from other regions of the body.

Concerning the retinal vessels, the wide availability of data opens the possibility of successfully performing experiments based on deep learning. A model that was learned with supervision has a natural tendency to capture particularities of the training data which will not be useful when making predictions over sufficiently different images. Therefore, models learned on a given dataset commonly do not generalise properly to different ones, even if they depict the same blood vessel tree, given that illumination, imaging artefacts, noise, and other particularities of the data make them very different from a computer point of view.

These limitations are not relevant when we know beforehand that the upcoming data is not significantly different from the data we used to obtain prior knowledge or train a DNN. In this case, application-specific algorithms tend to excel as they strongly focus the given task. This is the reason why this kind of methods has been more successful in real-life scenarios, as most of the settings are heavily controlled and the data conforms to an expected pattern. Even then, there are situations where models generalising well are indeed crucial. In the particular case of blood vessel segmentation, it is infeasible to re-calibrate the hyper-parameters of a traditional computer vision method or obtaining gold standard annotations of data for retraining a DNN, whenever the acquisition procedure is changed and/or a slightly different vessel tree shall be targeted. Therefore, there is a natural growing interest in methodologies that are powerful and simultaneously generalise well to similar tasks.

Chapter 5

Deep Vesselness Measure for Increased Generalisation

The content of this Chapter is based on the following work:

- R. J. Araújo, J. S. Cardoso, and H. P. Oliveira, "Deep Vesselness Measure from Scale-Space Analysis of Hessian Matrix Eigenvalues", In *Iberian Conference on Pattern Recognition and Image Analysis*, 2019.

Blood vessel imaging, as every procedure conducted in clinical practice, naturally evolves throughout time. The introduction of new medical equipment and/or the proposal of novel imaging protocols impact the appearance of blood vessels and the surrounding tissues in these medical images. To better perceive this, it suffices to assess the differences between retinal images coming from the different datasets presented in Section 2.1. Moreover, the advancement of technology and medical expertise creates the possibility for the emergence of novel treatments and techniques, which may require the imaging of structures which were not studied in the past. One example of this is the case of the DIEPs, whose analysis gained relevance after the introduction of the DIEP flap breast reconstruction, as covered in Section 4.1.

The aforementioned factors introduce an important challenge to the computer vision community, which is the development of algorithms which generalise properly to similar domains. Having methodologies which are capable of performing well in different but related domains enables easy and fast deployment of computer vision solutions in new imaging pipelines. However, the generalisation to unseen data is not easy to address. Approaches relying on deep learning and which are able to get supervision from datasets of relevant size have been the ones achieving best performance in the domains they have seen, nonetheless they typically generalise poorly to different ones. This was shown in Figure 1.4, where it is possible to see that a U-Net [23], a state-of-the-art model for segmentation of biomedical images, has its performance significantly decreased on unseen domains. This is not surprising, since supervised models will typically use their entire capacity to fit the training data distribution. Yet, this problem also affects many traditional computer vision algorithms, especially the most competitive ones. Even though these

rely mostly on strong prior knowledge about a given application, they commonly employ several heuristics and rule-based methods to tune the methodology to a particular scenario and see their performance indicators increase. Even then, most unsupervised approaches do generalise better to related domains, since properly encoding prior knowledge tends to generalise better than data-hungry methodologies (when domain shift is sufficiently large).

Vessel enhancement based on the Eigen decomposition of the Hessian matrix is one of the most widely used enhancement processes, due to natural formulation for both 2D and 3D scenarios, good generalisation to different data distributions, and also high noise suppression capabilities. Several hand-designed metrics combining the eigenvalue information were proposed in the past, as already described in more detail at Section 3.1.2. Even though these metrics rely on strong prior knowledge of the problem, they usually end discarding a lot of details due to how they combine multi-scale information, thus being suboptimal.

In this Chapter, we describe a methodology aiming to find a more complex and optimal vesselness measure mapping the eigenvalue information at different scales into the final vessel enhanced image, by means of a DNN. By using supervision, our goal is to obtain a deep vesselness measure that combines the advantages of both deep learning methodologies (finding deep complex functions) and using prior knowledge (increased robustness to data coming from different distributions). Recent research considered the implementation of Frangi's algorithm as a neural network, by careful initialisation of its weights [185]. The authors then used supervision to update weights responsible for the computation of the Hessian, and coefficients controlling the relevance of the different eigenvalue ratios used in Frangi's vesselness. Note, however, that the first option strongly relaxes the use of prior knowledge, as the network is able to learn features completely different from the Hessian, thus regularisation may be lost. Additionally, the authors do not consider exploring other functions mapping the Eigen maps to the final vesselness, being restricted to the use of the maximum operator across the responses obtained at different scales, which is suboptimal.

5.1 Traditional vesselness measures

Given a D-dimensional image $\mathbf{I}$, the type of structure present at a given location $\mathbf{x} = (x_1, x_2, \ldots, x_D)$ may be inferred through the analysis of the Hessian matrix at $\mathbf{x}$, a $D \times D$ matrix encoding the second order derivatives of the intensities:

$$H_{ij}(\mathbf{x}, \sigma) = \sigma^{\gamma} \mathbf{I}(\mathbf{x}) * \frac{\partial^2 G(\mathbf{x}, \sigma)}{\partial x_i \partial x_j}, \quad i, j = 1, \ldots, D \tag{5.1}$$

where G is a D-variate Gaussian, σ denotes its standard deviation, dictating the scale at which the image is being analysed, γ is a constant that normalises responses obtained at different scales, allowing a fair comparison [80], and $*$ represents the convolution operation.

The Eigen analysis of $H(\mathbf{x}, \sigma)$ produces D eigenvectors representing the principal directions that decompose the second order structure of the image at $\mathbf{x}$. Each of them has an eigenvalue associated, a scalar whose magnitude and signal allow to characterise the intensity curvature along

the corresponding eigenvector. From now on, let us consider that the Eigen decomposition of the Hessian at a location $\mathbf{x}$,

$$L(\mathbf{x},\sigma) = \text{eig}(H(\mathbf{x},\sigma)) \tag{5.2}$$

produces a set of eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_D$, such that, $|\lambda_1| \leq |\lambda_2| \leq \ldots \leq |\lambda_D|$. These provide a concise description of the local geometry at $\mathbf{x}$, allowing the design of functions that respond to particular geometries. In this context, a vesselness measure is any function f of the eigenvalues that is suited for the enhancement of blood vessels:

$$v(\mathbf{x},\sigma) = f(L(\mathbf{x},\sigma)) \tag{5.3}$$

A common assumption on vessel geometry is it being piecewise linear, that is, locally, it resembles a cylinder. As in this work we deal only with 2D images, we restrict the following discussion to this scenario. Nonetheless, extension to 3D is straightforward and addressed in the aforementioned works. The most commonly used vesselness measure is Frangi's [75], which for the 2D case is given by:

$$v_F = \begin{cases} 0 & \text{if} \quad \lambda_2 \leq 0, \\ \exp\left(-\frac{R_B^2}{2\beta^2}\right) \cdot \left(1 - \exp\left(-\frac{S^2}{2c^2}\right)\right) & \text{otherwise} \end{cases} \tag{5.4}$$

where $R_B = |\lambda_1|/|\lambda_2|$ is a ratio measuring local similarity to a blob through eccentricity of the second order ellipsis, $S = \sqrt{\lambda_1^2 + \lambda_2^2}$ is the amount of local structure, and β and c, control the relevance of those quantities, respectively. This formulation highlights vessels which are darker than the background, but inverting the conditions of (5.4) is enough to detect brighter vessels instead. In the case of Jerman's vesselness [79], assumptions are slightly relaxed in order to better model aneurysms and bifurcations:

$$v_J = \begin{cases} 0 & \text{if} \quad \lambda_2 \leq 0 \vee \lambda_p \leq 0, \\ 1 & \text{if} \quad \lambda_2 \geq \lambda_p/2 > 0, \\ \lambda_2^2 \cdot (\lambda_p - \lambda_2) \cdot \left[\frac{3}{\lambda_2 + \lambda_p}\right]^3 & \text{otherwise} \end{cases} \tag{5.5}$$

where λ_p is a regularised eigenvalue for ensuring that robustness to noise is achieved in regions with uniform intensity.

Regardless of the considered vesselness measure, the final enhanced image, $\mathbf{V}$, is obtained by combining the responses obtained at different scales σ, through a pixelwise maximum operation:

$$\mathbf{V}(\mathbf{x}) = \max_{\sigma_1,\ldots,\sigma_n} v(\mathbf{x},\sigma) \tag{5.6}$$

The traditional pipeline here described is represented in Fig. 5.1.

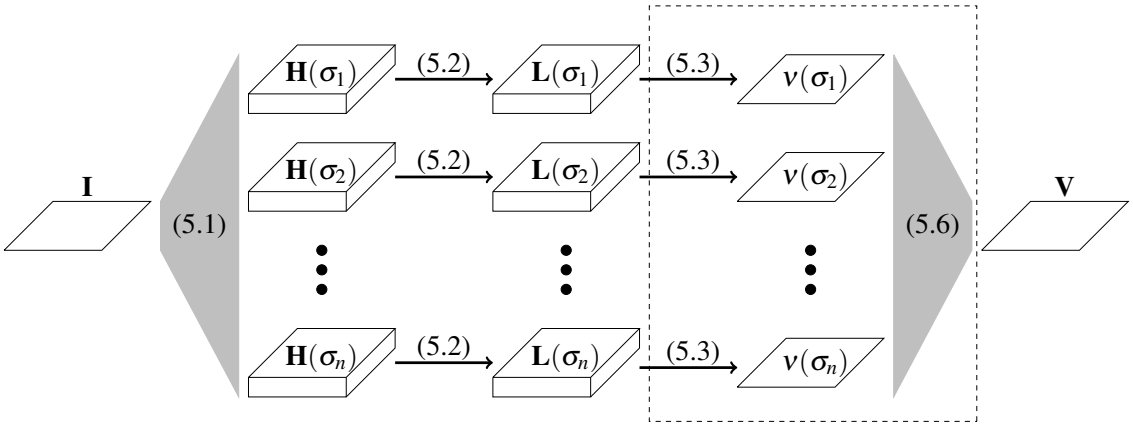


Figure 5.1: Multiscale pipeline of traditional vessel enhancement methodologies, where n denotes the number of scales. Pixelwise Hessian matrix and corresponding eigenvalues are represented as feature vectors for convenience. Our proposed design replaces the functions inside the dashed rectangle by a DNN.

5.2 Proposed deep vesselness measure

Our proposed methodology replaces hand designed vesselness measures (see region delimited by dashed lines in Fig. 5.1) by a DNN. Our motivation is twofold. First, mapping eigenvalue information into a vessel probability (5.3) through hand-designed functions, despite being based on prior intuition, is most likely suboptimal. Second, combining the responses at different scales by a pixel-wise maximum operation (5.6) discards much of the information encoded at all scales and fails to capture high-level local information that may be useful in challenging regions. Thus, we replace those functions by a neural network having as input the concatenation of the eigenvalue description, and as output a vessel probability map, as represented in Fig. 5.2. We use label supervision in order to update its weights, aiming to obtain a more optimal deep vesselness measure, which is still regularised as we only provide the scale-space eigenvalue description.

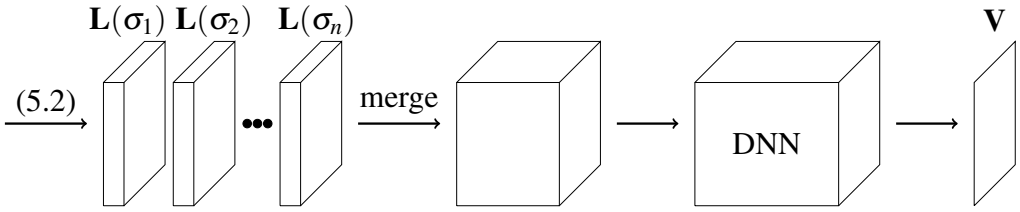


Figure 5.2: Proposed pipeline for vessel enhancement. A DNN learns a more complex vesselness measure from the eigenvalue information.

5.2.1 Neural network considerations

We modeled our DNN as an FCN [155], such that images seen at train and test phases may have different sizes. This also allowed us to train our model in small patches of blood vessel images, and still later obtain predictions for entire images at a single pass. This may be relevant due to memory issues and to avoid training with unnecessary data, such as the black regions in retinal fundus images. We considered patch-based training, which is not expected to affect the performance of an FCN.

Batch normalisation [186] is especially useful in very deep networks, which was not the case here. Additionally, care must be taken when the statistics of the data are not the same in the train and test sets. Such difference may be a result of performing patch-based training, where, for example, entire images of the retina have different statistics than small patches that were just taken from the retinal fundus area. It may also naturally arise from training and testing in different datasets. This last scenario is very relevant as one of the main advantages of traditional Hessian-based methods is their good generalisation to other distributions of data. Thus, we did not consider batch normalisation.

Recent findings [153] seem to support that reducing space resolution via max-pooling or strided convolution does not always improve the performance of networks trying to capture small details, as is the case of blood vessels. Preliminary experiments that we conducted support this, such that slightly increasing the kernel dimension and keeping spatial resolution equal across the entire network proved to be more effective. Having features already encoding neighbourhood information as the input of our neural network may also contribute to learn interesting deep vesselness measures by only looking at a relatively small neighbourhood. Even then, we found that using dilated convolution [187] in the intermediate layers improved the performance of the system.

An ideal vessel enhancement algorithm would output probability of 1 for every pixel belonging to a vessel and probability of 0 otherwise. However, note that, for an adequate scale σ, and when analysing pixels over the cross section of a vessel, it is expected that the Eigen decomposition of the centre pixel is the one matching better the Eigen description of an ideal vessel ($|\lambda_1| << |\lambda_2|$, in the 2D case). This is the reason why vesselness measures such as Frangi's enhance more the central regions of vessels. Even though a DNN is able to find complex relations between eigenvalues and thus learn effectively when hard labels are provided, in preliminary experiments, we found that using soft labels (obtained by blurring the hard labels with a standard normal distribution) was helpful. Nevertheless, as will be shown in Section 5.3, our design is still capable of enhancing the peripheral regions of vessels extremely well.

The ReLU function was used as an activation function throughout the network and the last non-linearity was a Sigmoid function. Regarding the loss function, we considered the BCE, which is also adequate when soft labels are given. The Adam optimiser [184] was used to update the weights. Other state-of-the-art considerations such as Dropout were also tested. The final FCN design is represented in Fig. 5.3. More information on the tuning procedure is given in subsection 5.3.2.

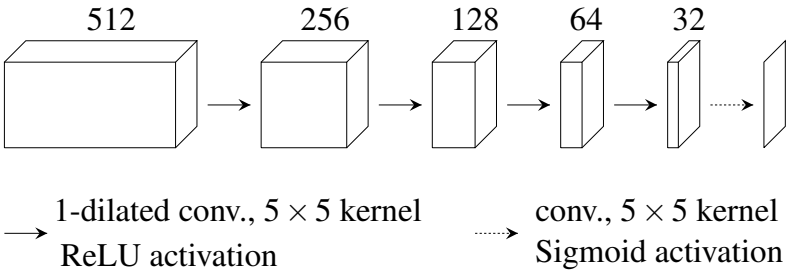


Figure 5.3: FC design used in the experiments after tuning the network architecture. The first set of features is obtained by doing convolution over the eigenvalue pile of features with 5×5 kernels and no dilation.

5.3 Experiments and Discussion

In this Section, we start by briefly describing the datasets and metrics we used during the experiments. Afterwards, we present the procedure that was followed to tune the network architecture and its hyper-parameters, in order to reach the design presented in Fig. 5.3. Finally, we detail the different experiments taken into account to show the properties of the proposed methodology and present the obtained results.

5.3.1 Datasets and Metrics

At the time of this research, to the best of our knowledge, blood vessel 2D imaging datasets containing the ground truth vessel masks only existed for the retinal case. This made us consider supervision only from this type of vascular network. To conduct the experiments, we resorted to the retinal datasets covered in Section 2.1, DRIVE [38], STARE [39], and CHASEDB1 [40]. To compare the different algorithms, we have analysed their ROC curves.

5.3.2 Implementation details

Having in mind the considerations discussed in subsection 5.2.1, we tuned the architecture and hyper-parameters using the DRIVE dataset. We randomly set aside three images from DRIVE's training set for validation purposes and used the remaining ones for training different model configurations. In this work we did not conduct any preprocessing step, we simply selected the green channel information and normalised it to the range $[0,1]$. We considered $\sigma \in [1,11]$ with steps of 2. At each training epoch, we gave the models 300 batches of 8 patches of size 64×64. A total of 100 epochs were conducted. These values were empirically found to be appropriate in preliminary experiments but their variation did not yield significant performance alterations. Patches were randomly extracted from the field of view region of images. We considered data augmentation via random vertical or horizontal flipping, and rotations in the range $[-\pi/2, \pi/2]$. The parameters of

the Adam optimiser were initialised as described in [184]. The best performance in the validation set was obtained when using the design represented in Fig. 5.3.

With the exception of DRIVE, the datasets do not have a prior split into train and test sets. Thus, we considered the first 10 images of each for training purposes and the remaining were set aside for testing. According to the experiment we conduct, different sets are used for training and testing but details will be provided as necessary. The training procedure is conducted as described before for network and hyper-parameter tuning, but this time all the available training data is used. Frangi's and Jerman's enhancement responses were obtained using their Matlab implementations and default parameters.

5.3.3 Results and Discussion

We start by considering the scenario where we train and test our model in the same dataset/domain (yet in different sets of data). With this experiment, we aim to show that for a specific distribution of data, it is possible to use deep learning to obtain more complex and optimal vesselness measures than traditional ones. The ROC curves of the proposed and baseline methods, for the different datasets, are shown in Fig. 5.4.

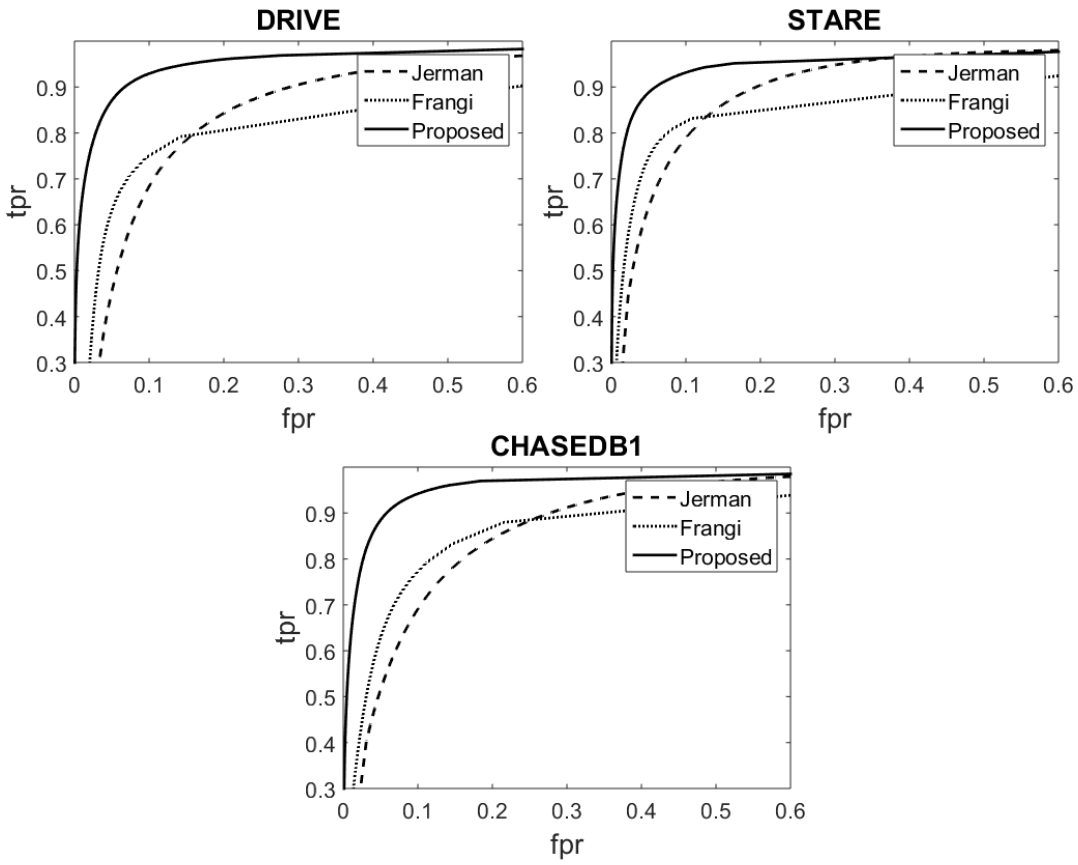


Figure 5.4: ROC curve of the proposed methodology when trained and tested in the same dataset. The ROC curves of the baseline methods are presented for comparison.

This shows that, when specific data distributions are targeted, we are able to obtain more optimal vesselness functions than the traditional ones. This was expected, but note that traditional methods do not target specific dataset distributions, but instead a representation that generalises well. Obviously, the more interesting scenario is to analyse what occurs when the proposed methodology is used to enhance blood vessels in images coming from data distributions other than the one(s) used during training. Thus, we now consider the scenario where we set the test dataset aside and train using the remaining ones. The ROC curves of the system in such conditions are again compared against the baselines in Fig. 5.5. It is possible to conclude that our system is

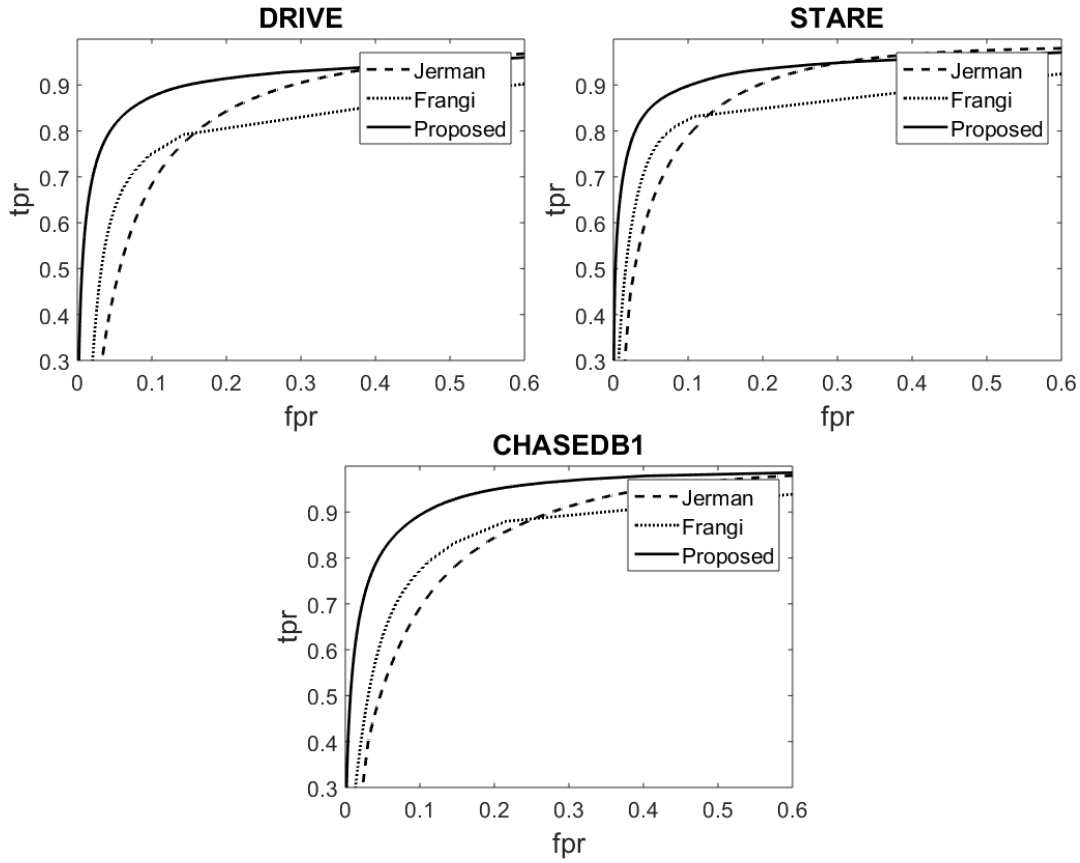


Figure 5.5: ROC curve of the proposed methodology when trained and tested in different datasets. The ROC curves of the baseline methods are presented for comparison.

indeed capable of generalising well to data coming from distributions that were not available during training. The proposed deep vesselness measure is then learning a complex mapping of the eigenvalue information that is useful to extract general tubular-like structures, not only to capture the particularities of a given data distribution. For very high false positive ratios, Jerman's vesselness occasionally achieves higher true positive ratio, however note that such region is not ideal for enhancement functions as it already comprises a large amount of noise. This is clearly seen in Fig. 5.7, where our method proves to be much more robust to noise than Jerman's one.

Finally, we compare the generalisation capability of the proposed method against the well-established U-Net [23] for biomedical image segmentation. Such network has much more capacity and has increased flexibility as it is not restricted to use a given set of features, as we do in the proposed methodology. Instead, it creates representations from the image itself. Fig. 5.6 shows the ROC curves of both methods. They achieved similar performance in STARE but the proposed

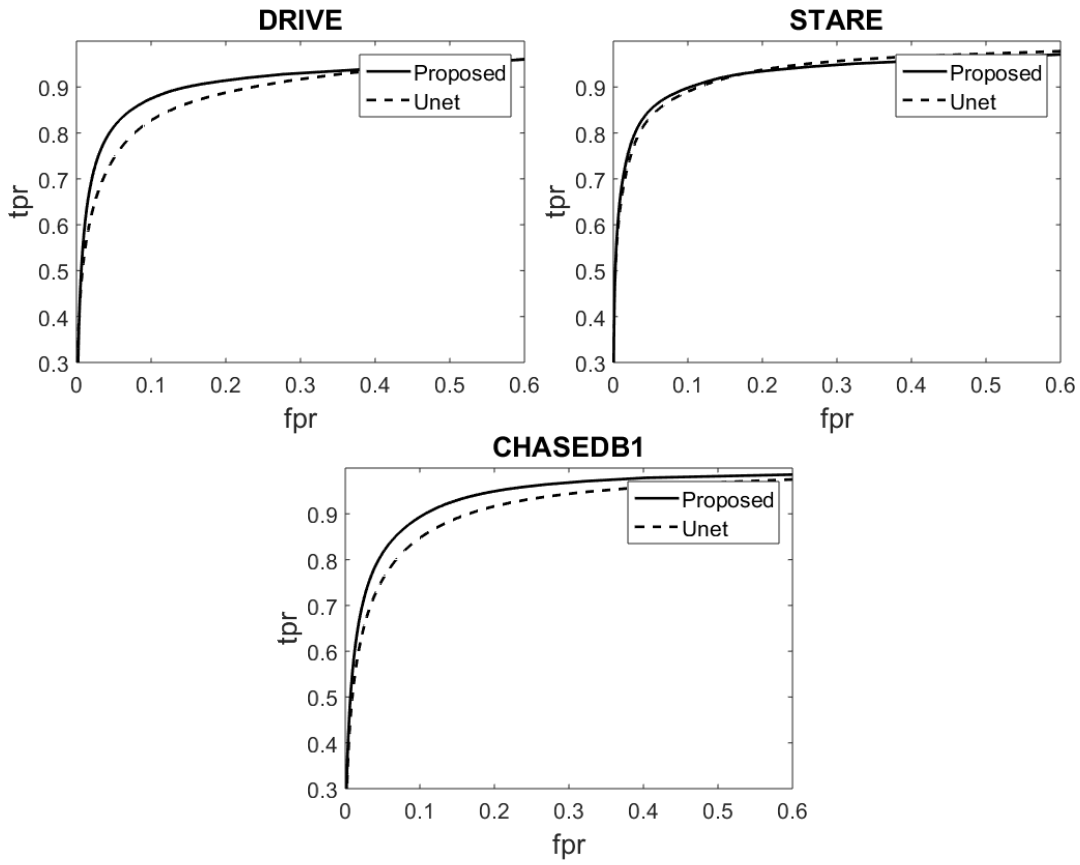


Figure 5.6: ROC curves of the proposed methodology and a regular U-Net, when trained and tested in different datasets.

approach generalised better for DRIVE and CHASEDB1. This shows that a careful regularisation of deep neural networks, as we described in this Chapter, is a relevant mechanism to achieve neural designs that generalise better to similar domains, even using significantly less parameters. Powerful networks having access to raw images and supervision are able to learn a much larger family of functions, and, while this generally allows to achieve strong performance in the same domain of the training data, it inevitably promotes learning features which are more specific and probably will not be as relevant for different distributions. By restricting the network to have as input a feature space which is known to properly encode tubular-like structures, we constrain the type of functions that can be learned. Therefore, the capacity of the model is likely to be used to learn mappings that are more adequate to general tubular structures. We can also view this regularisation as transforming an initial input space where samples coming from different

domains are far apart (due to illumination, noise, artefacts, among other processes) into a proper feature representation where they become more similar. Visual results of our method and the baselines are provided in Fig. 5.7.

5.4 Main contributions and final considerations

This Chapter focused a topic that has been trending in the machine learning area, which is the ability of generalising to new but related data distributions. Our approach to this challenge, in the particular scenario of blood vessel segmentation, combined the strengths of two different types of methodologies: i) the generalisation typical of algorithms relying on prior knowledge; and ii) the capability of modelling complex functions by learning DNN models using supervision. To do so, we extended traditional Hessian-based methodologies for the enhancement of blood vessels in medical images. By replacing hand-designed functions mapping eigenvalue descriptions to the final output with a shallow DNN, we were able to learn more optimal functions than traditional vesselness measures. At the same time, when comparing with a U-Net which was fed images instead of an eigenvalue description, our methodology generalised better to data coming from distributions other than the ones used at training. This showed that our approach was able to embed significant prior knowledge, thus helping to achieve good abstraction of what a blood vessel is.

Lately, there has been a lot of research regarding domain adaptation techniques, which apply mechanisms that allow to better use the knowledge obtained for a given set of datasets (commonly regarded as source datasets) in the analysis of new data which is related but has a different distribution (the target data). These methods are divided into unsupervised and semi-supervised domain adaptation, according to whether no or some supervision from the target dataset is used during model training. These frameworks tend to employ auxiliary tasks [131] and/or adversarial losses [132] to promote the learning of features which are relevant for the different data distributions.

While our approach relied on a powerful yet simple description that is basis for all tubular structures, domain adaptation techniques have been letting complex DNNs reach very promising performances in target domains using a very small amount of labels or even no supervision at all from target data. Therefore, despite the positive findings achieved with our proposed methodology, we believe that domain adaptation techniques are very promising and we would like to consider them in future experiments. Additionally, we would like to extend our research to 3D data, possibly exploring synthetic datasets during the training procedure, since tools such as VascuSynth [188] allow to readily generate large amounts of data and the respective ground truth annotations.

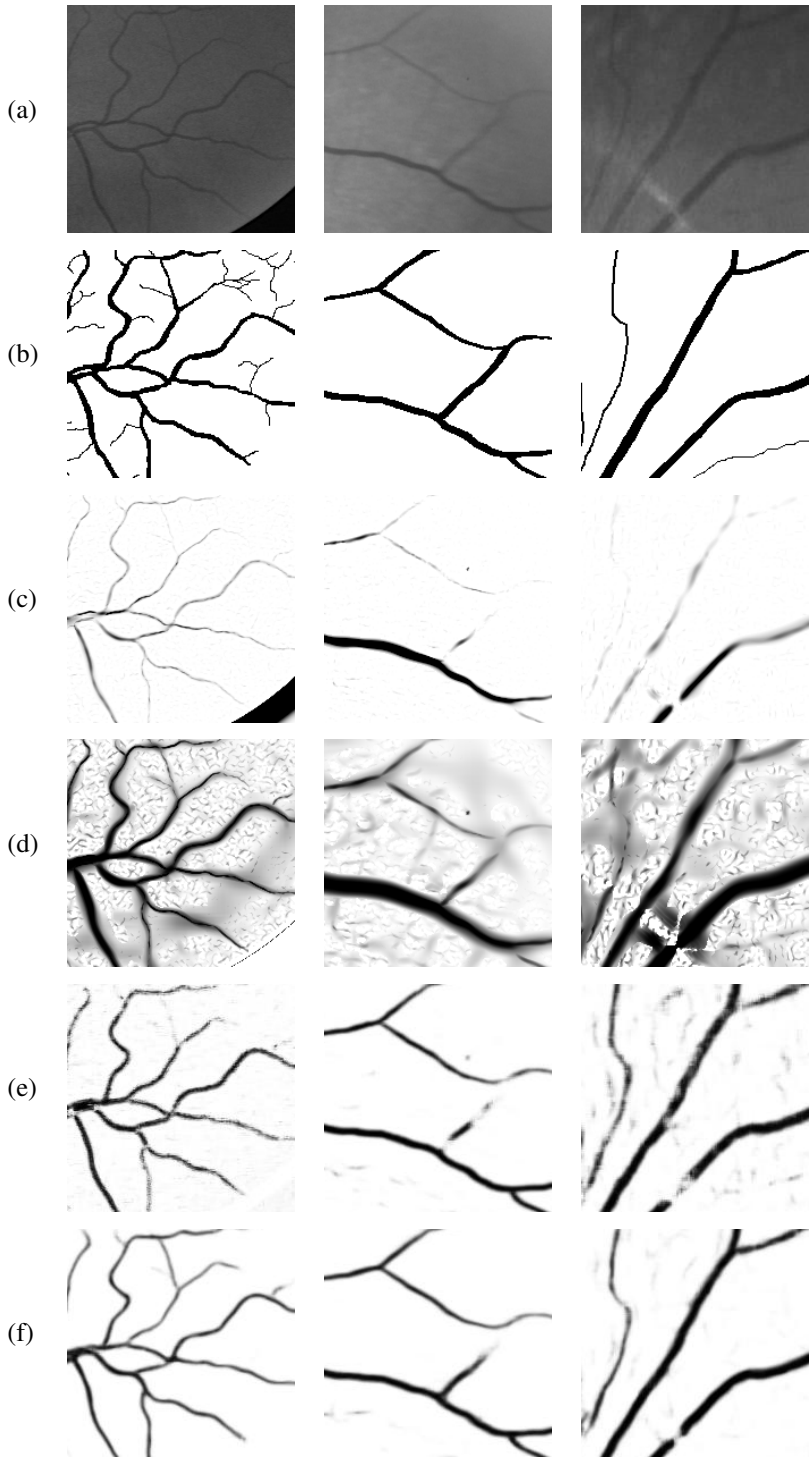


Figure 5.7: Blood vessel enhancements achieved by the baseline methods and the proposed approach: (a) example retinal fundus images; (b) ground truth vessel masks; and corresponding enhancements by, respectively, (c) Frangi's, (d) Jerman's, (e) U-Net, and (f) proposed vesselness. From left to right, image from DRIVE, STARE, and CHASEDB1. Concerning the U-Net and the proposed method, training was conducted in all datasets, except the testing set. Frangi's vesselness was rescaled for visualisation purposes, since the signal at narrow vessels is usually small. The masks are inverted for visualisation purposes.

Chapter 6

Topology Coherence

The content of this Chapter is based on the following work:

- R. J. Araújo, J. S. Cardoso, and H. P. Oliveira, "A deep learning design for improving topology coherence in blood vessel segmentation", In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2019.
- R. J. Araújo, J. S. Cardoso, and H. P. Oliveira, "Topological blood vessel segmentation", arXiv, 2021

Despite all the progress that is being achieved in blood vessel segmentation lately, mostly through the exploration of deep learning, the algorithms are still far from mimicking expert behaviour. When considering the typical pixel-wise metrics, supervised learning has been able to produce models which reach expert performance, even though their generalisation to different datasets should not be expected to match that of a human. Moreover, the analysis of pixel-wise metrics is not sufficient to understand how close the predictions of a model really are from those of an expert. In fact, when observing the predictions of models, we commonly see errors that allow us to understand that they were not capable of grasping the complete concept of blood vessel networks, since they typically miss vessel segments in challenging regions, due to structure overlap, central reflex, or low SNR. Examples of these errors are provided in Figure 6.1. A missing segment may correspond to discarding the termination of a vessel or to inducing a disconnected tree, both of which (especially the latter) causing a modification of the underlying graph representing the vascular network. Thus, we designate these errors as topological errors. In our opinion, the current metrics do not allow to properly understand how well the predictions of models maintain the topology of the ground truth masks. This may be essential in some scenarios, as these errors may put at risk applications that require vessel pathway extraction and/or characterisation [189], and are likely to deeply affect any routine that should follow an automated segmentation. The state-of-the-art methodologies are prone to commit topological errors since they rely on the minimisation of pixel-wise loss functions which do not account for the structure of the errors, such

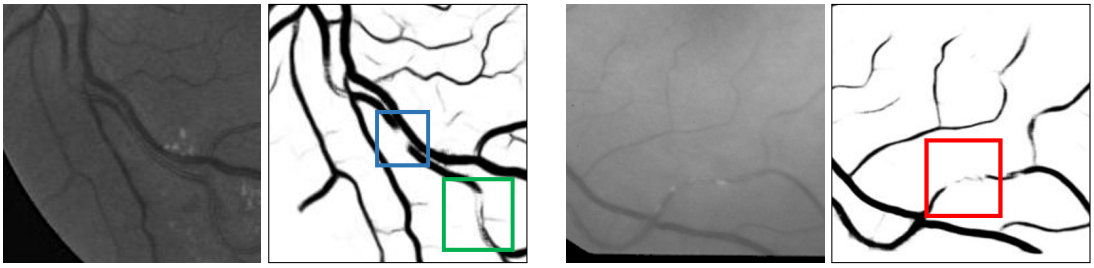


Figure 6.1: Example images and corresponding segmentations obtained with a U-Net model [23] learned by minimising the BCE loss. Topological errors commonly emerge in challenging cases: central reflex (green), the latter plus neighbour structure influence (blue), and poor local contrast (red). Masks are inverted for visualisation purposes.

as the BCE and the soft Dice losses, and mostly use model architectures that are not aware of topological incoherences.

Recently, there have been attempts to incorporate topological awareness in deep learning models targeting different applications. A loss encoding hierarchical relations between labels, such as containment and detachment, was designed to improve the multi-class segmentation of histology glands [190]. In [191], a process for consecutive refinement of a segmentation given the grey-scale image and the previous mask was proposed, guided by the differences between high-level features of the current segmentation and the ground truth. However, none of these works was applied to vessel segmentation. Bifurcation detection has been addressed in a parallel fashion [192], aiming to enhance the overall segmentation process of vascular networks, and consequently, the overall network topology. Nonetheless, topological errors do not arise in bifurcations only, appearing frequently in the middle of branches due to different reasons, as demonstrated in Figure 6.1.

This Chapter has the purpose of raising awareness to the relevance of the topological properties of vascular networks during learning and evaluation steps, and describes the work we have conducted in this context. We start by presenting an end-to-end deep neural network architecture comprising a probabilistic refinement step and showing how it can successfully improve topological consistency via model design. Afterwards we discuss the need for topological metrics, contemplating some attempts that have been made in the literature, and proposing our own similarity index. In the end, we present a novel loss function which is shown to increase the robustness to topological errors.

6.1 Deep probabilistic refinement for increased topological awareness

In this section, we present an end-to-end deep neural network design for improving topological consistency in blood vessel segmentation. The methodology comprises a typical segmentation network followed by a refinement model which aims to enforce the learning of meaningful features from noisy data. We discuss how such design can be used as a strategy to reduce topological

mistakes. In what follows, $\mathbf{X}$ and $\mathbf{Y}$ denote, respectively, the grey-scale input and the ground truth vessel mask, and $\mathbf{Y}'$ and $\mathbf{Y}''$ represent the outputs of the segmentation network and the refinement model, respectively.

6.1.1 Auto-encoding for learning local topology

Errors committed by a segmentation network can be interpreted as a hidden noise process affecting the true vessel signal $\mathbf{Y}$. Thus, we seek an encoding of $\mathbf{Y}'$ that does not model this noise, allowing $\mathbf{Y}''$ to better depict the topology of $\mathbf{Y}$.

Usually, auto-encoding designs encode the entire image $\mathbf{X}$ into a vector $\mathbf{z} \in \mathbb{R}^D$, assuming that complex large scale spatial interactions may be learned. This is not the most adequate option for encoding images where recurring patterns exist, as is the case of blood vessels. In this type of data, for inference purposes, it is better to follow the methodology in [193], where we have latent variables $\mathbf{z}$ as a 3D tensor (stack of feature maps) instead, for explicitly capturing spatial information.

Let us start by considering the generation process of ground truth vessel masks $\mathbf{Y}$. It consists of sampling latent variables from a prior distribution $p_{\theta^*}(\mathbf{z})$ and generating masks according to a conditional distribution $p_{\theta^*}(\mathbf{Y}|\mathbf{z})$. We assume these distributions belong to parametric families of distributions $p_\theta(\mathbf{z})$ and $p_\theta(\mathbf{Y}|\mathbf{z})$. Given observations $\mathbf{Y}$, we want to perform inference in this model, $p_\theta(\mathbf{z}|\mathbf{Y}) = (p_\theta(\mathbf{Y}|\mathbf{z})p_\theta(\mathbf{z}))/p_\theta(\mathbf{Y})$, to obtain distributions over the latent space explaining the different observations.

The described approach leads to an intractable problem because evaluating the marginal likelihood of the data, $p_\theta(\mathbf{Y})$, requires integrating over the entire latent space. This limitation can be circumvented using variational inference by approximating the posterior probability with a family of distributions $q_\lambda(\mathbf{z})$. The optimal parameters λ^* are the ones minimising the Kullback-Leibler divergence between the two distributions:

$$\begin{aligned} \mathbb{D}_{\mathrm{KL}}(q_\lambda(\mathbf{z})||p_\theta(\mathbf{z}|\mathbf{Y})) &= \mathbb{E}_{q_\lambda(\mathbf{z})}\left[\log\left(\frac{q_\lambda(\mathbf{z})}{p_\theta(\mathbf{z}|\mathbf{Y})}\right)\right] \\ &= \mathbb{E}_{q_\lambda(\mathbf{z})}\left[\log q_\lambda(\mathbf{z}) - \log p_\theta(\mathbf{z},\mathbf{Y})\right] + \log p_\theta(\mathbf{Y}) \end{aligned} \tag{6.1}$$

However this optimisation problem also requires computing the marginal likelihood, thus being once again intractable. By noting that $\mathbb{D}_{\mathrm{KL}}$ is a non-negative quantity and rearranging (6.1):

$$\begin{aligned} \log p_\theta(\mathbf{Y}) &= \mathbb{D}_{\mathrm{KL}}(q_\lambda(\mathbf{z})||p_\theta(\mathbf{z}|\mathbf{Y})) + \mathbb{E}_{q_\lambda(\mathbf{z})}\left[\log p_\theta(\mathbf{z},\mathbf{Y}) - \log q_\lambda(\mathbf{z})\right] \\ &\geq \mathbb{E}_{q_\lambda(\mathbf{z})}\left[\log p_\theta(\mathbf{Y}|\mathbf{z})\right] - \mathbb{D}_{\mathrm{KL}}\left(q_\lambda(\mathbf{z})||p_\theta(\mathbf{z})\right) \end{aligned} \tag{6.2}$$

we obtain the Evidence Lower BOund (ELBO), which can equivalently be maximised, allowing us to do approximate posterior inference.

The Variational Auto-Encoder (VAE) [194] conditions the approximate posterior on the data. This distribution, $q_\phi(\mathbf{z}|\mathbf{Y})$, and the data likelihood one, $p_\theta(\mathbf{Y}|\mathbf{z})$, are both parametrised by neural networks, which are commonly designated as recognition (or encoder) and generative (or decoder) models, respectively. The weights of both networks are jointly learned using the Stochastic Gradient Variational Bayes estimator introduced in the same work [194]. Since parameters ϕ are shared among all observations, this model performs amortised inference.

Until now, we considered the case of auto-encoding the vessel masks $\mathbf{Y}$. However, the aim of this work is to use the VAE as a segmentation refiner. Therefore, our recognition model is conditioned by the segmentation output $\mathbf{Y}'$, while the generative model produces masks $\mathbf{Y}''$ closer to the ground truth $\mathbf{Y}$. As we shall discuss next, this formulation is a particular case of a Denoising VAE [195].

6.1.2 Refinement model as a Denoising VAE

The Denoising VAE (DVAE) [195] is trained on noisy observations, where the noise is modeled by a distribution conditioned on the data, $p_\gamma(\mathbf{Y}'|\mathbf{Y})$. In our use case, the outcome of the segmentation network, $\mathbf{Y}'$, is interpreted as a noisy version of the true vessel signal, $\mathbf{Y}$. In a DVAE, the recognition model is given by:

$$\tilde{q}_\phi(\mathbf{z}|\mathbf{Y}) = \int q_\phi(\mathbf{z}|\mathbf{Y}')p_\gamma(\mathbf{Y}'|\mathbf{Y})d\mathbf{Y}' \tag{6.3}$$

The modified ELBO of the DVAE comes as:

$$\mathbb{E}_{\tilde{q}_\phi(\mathbf{z}|\mathbf{Y})}\left[\log\left(\frac{p_\theta(\mathbf{z},\mathbf{Y})}{\mathbb{E}_{p_\gamma(\mathbf{Y}'|\mathbf{Y})}\left[q_\phi(\mathbf{z}|\mathbf{Y}')\right]}\right)\right] \tag{6.4}$$

but a more practical lower bound was proven to be eligible for optimisation by the authors [195]:

$$\mathbb{E}_{\tilde{q}_\phi(\mathbf{z}|\mathbf{Y})}\left[\log\left(\frac{p_\theta(\mathbf{z},\mathbf{Y})}{q_\phi(\mathbf{z}|\mathbf{Y}')}\right)\right] \tag{6.5}$$

which is equivalent to training a regular VAE on noisy examples. From (6.5) follows the conclusion that the recognition model in the DVAE is trying to learn meaningful features from the noisy observations, in order to obtain latent representations that allow the generative model to produce an output that is similar to the noiseless data. Our proposed refinement model can be seen as a particular case of a DVAE, where the noise model is not known and is encoded in the observations $\mathbf{Y}'$ instead. In Figure. 6.2, we present the proposed design for obtaining segmentation masks that are topologically more coherent.

6.1.3 Experiments and Discussion

The U-Net model [23] is very popular for segmenting biomedical images, given its capability of accounting for both low and high-level features of the images. In this work, the U-Net was used as the segmentation network. Our proposed method (*prop*) was compared against two baselines:

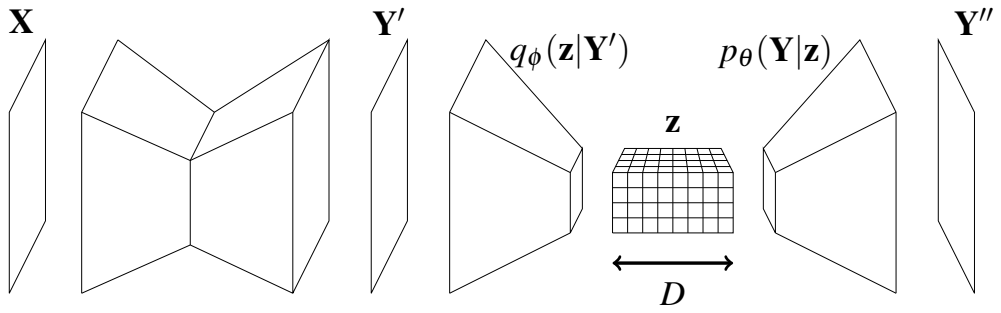


Figure 6.2: Design of the proposed model for blood vessel segmentation. $\mathbf{X}$ and $\mathbf{Y}'$ are, respectively, the input image and the mask obtained using any particular segmentation network. $\mathbf{Y}''$ is the output of the proposed DVAE receiving as input $\mathbf{Y}'$.

i) a single U-Net producing vessel masks (*unet*), and ii) a cascade of two U-Net models which performs segmentation and refinement tasks (*dunet*). The losses of these models are, respectively:

$$L_{prop} = \alpha \cdot L_1(\mathbf{Y}', \mathbf{Y}) + (1-\alpha) \cdot \big(L_2(\mathbf{Y}'', \mathbf{Y}) + \mathbb{D}_{\mathrm{KL}}(q_\phi(\mathbf{z}|\mathbf{Y}')||p_\theta(\mathbf{z}))\big) \tag{6.6}$$

$$L_{unet} = L_1(\mathbf{Y}', \mathbf{Y}) \tag{6.7}$$

$$L_{dunet} = \alpha \cdot L_1(\mathbf{Y}', \mathbf{Y}) + (1-\alpha) \cdot L_2(\mathbf{Y}'', \mathbf{Y}) \tag{6.8}$$

with $p_\theta(\mathbf{z})$ being the standard Gaussian. We tested the impact of using losses L_1 and L_2 other than BCE to train the models: the class-weighted BCE (BCEw), which penalises more false negatives than false positives (weights of 0.7 and 0.3 were found appropriate for vessel and non-vessel classes, respectively); and the FL [183], which is an extension of BCE that increases the weight of pixels according to the magnitude of the error.

6.1.3.1 Datasets and Metrics

We performed experiments in the three benchmarks for retinal vessel segmentation that have been introduced in Section 2.1: DRIVE [38], STARE [39], and CHASEDB1 [40] databases.

To compare the performance of the models, we considered usual pixel-wise metrics, such as: AUC, sensitivity, and specificity. To evaluate the topological coherence of the masks, we followed a similar approach to [196]. A connected path is randomly chosen from the ground truth and the equivalent path in the binarised prediction mask is analysed. The prediction is classified as infeasible if such path does not exist. Otherwise, it is wrong or correct whether its length differs by more than 10%, or not, respectively. We sampled 1000 paths per test image.

6.1.3.2 Implementation details

The train data of DRIVE was randomly split into 15 training and 5 validation images, in order to tune the models. When these included a refinement step (*dunet* and *prop*), α was set to 1 and decreased by 5e$-$3 each epoch until 0.3, and L_1 and L_2 consisted of the same type of loss function. For stability purposes, when using the FL, we set $L_1 =$ BCE and $L_2 =$ FL. The training procedure lasted for 150 epochs where, in each, 300 batches of 16 patches of size 64x64 were used. Patches were taken from the green channel of images and augmented via random transformations including horizontal and vertical flips, rotations in the range $\left[-\frac{\pi}{2}, +\frac{\pi}{2}\right]$, and addition of an intensity bias.

The original U-Net model comprises 4 condensing and expanding levels; however, we concluded that 2 were ideal in this scenario, when considering the AUC metric. Afterwards, we tuned the refiners in the *dunet* and *prop* models, considering the number of correct paths. Both pipelines ended having around 4M parameters. The best performing *dunet* model was a cascade of two U-Nets as the one described above. Our proposed recognition model was constituted by 4 convolutional layers (3$\times$3 kernels and padding of 1) producing, respectively, 64, 64, 256, and 256 feature maps. Each of the first 3 is followed by a max-pooling layer (kernel size of 2). Then, convolutional layers (1$\times$1 kernels, no padding) learning D feature maps, parametrise the diagonal Gaussian over the latent space. D was tuned to 100. Regarding the generative model, it includes 3 transposed convolutional layers (4$\times$4 kernels, padding and stride of 2) producing, respectively, 256, 256, and 64 feature maps, followed by 2 convolutional layers (3$\times$3 kernels and padding of 1), where the first learns 64 kernels and the last outputs the parameters of a Bernoulli distribution. ReLUs were used in the intermediate layers of the proposed VAE, and a Sigmoid activation function in the last one. The modulating constant of $\mathbb{D}_{\mathrm{KL}}$ was tuned to 1e$-$3.

Having tuned the structure and hyper-parameters of the models, they were trained as before, but this time using all the train data. Note that we perform patch-based training, but the design of the models allows single-pass prediction of a complete image. The implementation of the models, training procedure, and described losses, in PyTorch, is available at https://github.com/rjtaraujo/dvae-refiner.

6.1.3.3 Results and discussion

The average performance of the models on 5 different runs is shown in Table 6.1. The FL slightly increased the AUC of the models, meaning that they became better at separating both classes. However, that did not necessarily translate into better topological masks in the end. This is not surprising, as giving more focus to hard cases does not guarantee we are giving more weight to the pixels that generate topological mistakes. Instead, using BCEw, thus giving more weight to the vessel class, allowed to improve the sensitivity and the topology, as was expected. Proceeding to model design comparison, the proposed method was able to significantly decrease the number of infeasible paths, essentially due to finding the correct topology, as demonstrated by the increase of correct paths. This was achieved without hurting pixel-wise metrics, as may be seen by analysing the AUC. In fact, this metric was even improved in some cases. Note that there is a compromise

Table 6.1: Performance of the models, in percentage, averaged over 5 runs. AUC, *sen*, *spe*, *inf*, and *cor* stand for, respectively, area under the roc curve, sensitivity, specificity, infeasible, and correct paths. The larger these metrics, the better the performance of the model, with the exception of *inf*, where lower values are better. The best obtained performance for each indicator and considered database is highlighted in bold.

		BCE			BCEw			FL		
		unet	dunet	prop	unet	dunet	prop	unet	dunet	prop
DRIVE	AUC	97.7	97.8	97.8	97.9	97.9	97.9	**98.0**	**98.0**	97.9
	sen	79.2	79.6	85.1	87.4	87.8	89.7	78.4	79.0	82.3
	spe	98.1	98.0	96.7	96.2	96.1	95.3	98.1	98.1	97.4
	inf	47.0	45.0	34.1	34.8	31.8	**29.1**	48.6	44.8	40.4
	cor	45.5	47.3	56.7	56.7	59.2	**61.2**	43.8	48.3	51.4
STARE	AUC	98.0	98.2	98.3	98.1	98.4	98.6	98.7	**98.8**	**98.8**
	sen	80.5	82.7	87.3	87.7	89.1	90.1	81.1	82.7	85.2
	spe	98.5	98.4	97.3	97.3	97.2	96.8	98.5	98.4	97.9
	inf	53.4	43.2	27.9	38.9	29.2	**23.1**	49.7	38.4	34.9
	cor	40.8	51.8	61.9	55.3	64.3	**69.2**	43.6	54.4	58.1
CHASE	AUC	97.6	97.7	97.9	97.8	98.0	98.0	97.9	**98.2**	**98.2**
	sen	80.7	80.5	82.8	87.8	88.4	89.8	80.6	80.9	84.2
	spe	97.6	97.6	97.4	95.9	95.9	95.6	97.5	97.7	97.2
	inf	74.9	74.0	64.7	60.5	54.6	**48.0**	73.7	71.4	62.2
	cor	20.9	22.8	29.8	32.4	38.1	**45.6**	21.6	24.5	31.8

between the sensitivity and specificity of the models, such that using them for direct comparison of models is often not trivial. By comparing with the results achieved by the *dunet* model, which is also a model with more capacity, we conclude that our proposed design effectively learned better features for ensuring topological coherence. Fig. 6.3 shows some visual results of the three models trained with BCEw.

6.1.4 Summary

We proposed a design where a VAE is cascaded after a segmentation network, with the purpose of improving the topological coherence of the predicted blood vessel masks. The experiments showed that our methodology achieves that objective by predicting more correct paths and less infeasible paths, without negatively affecting pixel-wise metrics. The results of comparing the proposed method with a cascade of two U-Net models sustain that the improvement comes from the model design and not from the increased complexity of the pipeline.

6.2 Assessing topological coherence

Blood vessel trees are graph-like structures in the sense that, besides local calibre, they are well encoded by a graph $G = (V, E)$, where vertices V represent bifurcations and vessel terminations,

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

Figure 6.3: Example masks obtained by using the BCEw loss: (a) original images, (b) ground truth, and predictions from (c) U-Net, (d) Double U-Net, and (e) proposed method. Notice that the masks are inverted for visualisation purposes.

and undirected edges E encode the segments connecting them. We designate by topological coherence the similarity between two vascular tree graphs, such as the ones corresponding to the ground truth and the predicted segmentation of a given image.

As shown in Section 6.1, one option to evaluate the topological properties of the algorithms is to sample paths from the gold standard mask and determining the proportion of infeasible (impossible to reach the end point from the initial one), wrong (a path exists but is not equivalent to the gold standard one) and correct paths in the predicted mask. Even then, it may not always be trivial to objectively compare two different algorithms, as a decrease in the amount of infeasible paths will be likely followed by an increase in both wrong and correct paths, which naturally constitutes a trade-off, similarly to sensitivity and specificity. Moreover, to capture the amount of false positive paths being introduced by the algorithms, the analysis should be made in the opposite direction too, further increasing the complexity of analysing the behaviour and taking conclusions.

A novel similarity metric derived from the Dice score and focusing the blood vessel centrelines, designated clDice, has been proposed [197]:

$$clDice(\mathbf{P},\mathbf{Y}) = 2 \cdot \frac{clSens(\mathbf{P},\mathbf{Y}) \cdot clPrec(\mathbf{P},\mathbf{Y})}{clSens(\mathbf{P},\mathbf{Y}) + clPrec(\mathbf{P},\mathbf{Y})} \tag{6.9}$$

$$clSens(\mathbf{P},\mathbf{Y}) = \frac{\sum_i \mathbf{P} \cdot \mathbf{Y}_s + \varepsilon}{\sum_i \mathbf{Y}_s + \varepsilon} \tag{6.10}$$

$$clPrec(\mathbf{P},\mathbf{Y}) = \frac{\sum_i \mathbf{Y} \cdot \mathbf{P}_s + \varepsilon}{\sum_i \mathbf{P}_s + \varepsilon} \tag{6.11}$$

where $\mathbf{P}$ and $\mathbf{Y}$ denote the predicted and ground truth masks, respectively, $\mathbf{P}_s$ and $\mathbf{Y}_s$ are their centrelines, and ε is a small constant to deal with the cases where the denominator would be 0.

The authors [197] argue that clDice accounts for the topology of the vascular network; however, we believe that this metric is not a truly topological one, since it yields the same penalty for a missing segment in a vessel termination and one of equal extension that leads to disconnected trees. The latter has a much larger impact on the vascular tree graph, as illustrated in Figure 6.4. Therefore, we believe the literature is lacking a unified metric that better captures these topological properties.

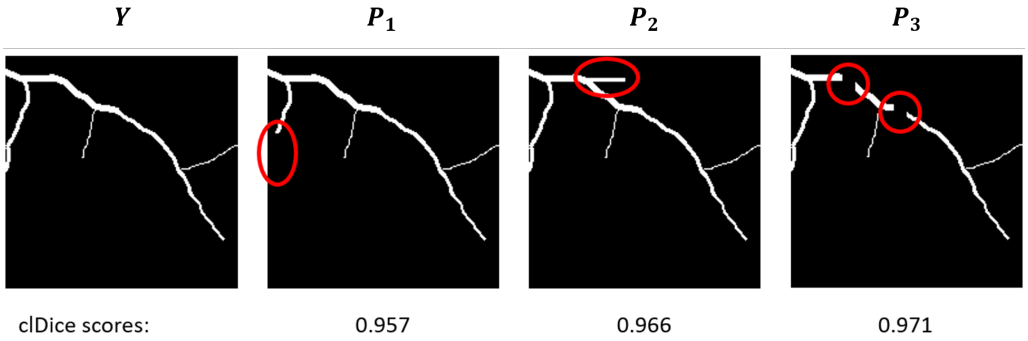


Figure 6.4: The effect of different errors on the clDice metric. $\mathbf{Y}$ is the ground truth mask and $\mathbf{P}_1, \mathbf{P}_2$, and $\mathbf{P}_3$ are segmentations where, respectively, a termination segment is missing, a false positive branch is added, and two missing segments that induce disjoint trees exist (errors highlighted in red). As shown, the clDice metric is sensitive to the extension of centreline error, not to the effect of the errors in the overall graph. A topological metric, in our opinion, should penalise P_3 more than P_1 and P_2.

In our opinion, a proper topological metric or similarity index should highlight the errors affecting the vascular tree graph according to the following properties:

- **Property 1**: broken segments should be further penalised than missing termination segments, since they lead to major changes of the underlying graph;

- **Property 2**: topological errors in the main vascular tree branches should have larger impact, since they may lead to larger sub-trees being lost in automated analysis algorithms.

As mentioned before, the most commonly used metrics for evaluating blood vessel segmentation algorithms - accuracy, sensitivity, specificity, and AUC - do not possess any of these properties. Even though they penalise broken and missing termination segments, these errors tend to be scarcer than calibre-related ones, thus they are strongly dissipated. The clDice metric focuses centrelines and neglects the errors due to calibre assessment. Nevertheless, as shown in Figure 6.4, it equally penalises termination and broken tree inducing missing segments of same extension, and it does also not distinguish errors occurring in major and peripheral branches. Therefore, to the best of our knowledge, there is no metric or similarity index in the blood vessel literature satisfying properties 1 and 2.

6.2.1 Proposed topological similarity index

Let $\mathbf{A}$ be a binary blood vessel mask, which can consist of a single connected component or multiple sub-trees. Let $\mathbf{A}_f$ be the set of vessel (foreground) pixels and $P_{i,j} = \{i, \ldots, j\}$ be the minimum cost path between pixels $i, j \in \mathbf{A}_f, l_i = l_j$, where l_k specifies the sub-tree to which a vessel pixel k belongs.

Removing a termination blood vessel segment $T = \{t_1, \ldots, t_n\}$ would render impossible all paths $P_{i,j}$ where $i \in T \vee j \in T$. Considering a broken segment $B = \{b_1, \ldots, b_n\}$ instead, any path $P_{i,j}$ intersecting B, $P_{i,j} \cap B \neq 0$, would also not be possible anymore. Note how this relates with the first property we have specified, as a broken segment has a larger impact than a missing termination in terms of the amount of paths that become impossible (assuming errors of similar size). In fact, it is also likely to satisfy the second property due to the pattern naturally displayed by blood vessel trees. Given their root, they keep dividing in branches, and calibre decreases with each division. Therefore, there is a natural tendency for the majority of possible paths in a vascular tree to traverse the major branches. This behaviour makes us consider functions taking into account the feasibility of paths when designing our proposed topological similarity indices. Based on this motivation, we now formulate a similarity index $m : \mathbb{R}^D \times \mathbb{R}^D \rightarrow [0,1]$ comparing two D-dimensional binary masks. We are interested in assessing how feasible the possible paths in the ground truth $\mathbf{Y}$ are in a segmentation $\mathbf{P}$ (path sensitivity/recall) and also the amount of false positive paths that exist in $\mathbf{P}$ (path precision). Let N_P and N_Y denote, respectively, the number of possible paths in $\mathbf{P}$ and $\mathbf{Y}$. An expression for a similarity index taking into account what was discussed until now follows:

$$m(\mathbf{P}, \mathbf{Y}) = \sqrt{\frac{\sum\limits_{i,j \in \mathbf{Y}_f,\, l_i = l_j} f(P_{i,j}, \mathbf{P})}{N_Y} \cdot \frac{\sum\limits_{i,j \in \mathbf{P}_f,\, l_i = l_j} f(P_{i,j}, \mathbf{Y})}{N_P}} \tag{6.12}$$

where $f(P_{i,j}, \mathbf{A})$ is any function assessing the coherence between a path $P_{i,j}$ and a mask $\mathbf{A}$, returning 0 and 1 in the extreme cases of, respectively, no and full coherence. In our experiments, we consider two different possibilities for f based on the Hamming distance H, a metric that counts the number of switches that are required to have two n-bit strings match. Notice that our problem can be interpreted as comparing two strings, since that: (i) a path $P_{i,j}$ of length n can be seen as a n-length string of 1s; (ii) and the values that $\mathbf{A}$ takes at $\mathbf{x} \in P_{i,j}$ can also be represented as a n-length string, this time possibly containing both 0s and 1s, according to whether A takes the value of 0 or 1, respectively, at each of the path positions. Given that we want f to produce values in the range $[0,1]$, and 1 to be equivalent to complete coherence, we define the first possibility, f_H, as:

$$f_H(P_{i,j}, \mathbf{A}) = \frac{n - H(P_{i,j}, \mathbf{A})}{n} \tag{6.13}$$

where $n \geq 2$ is the number of pixels in the path $P_{i,j}$. An illustration demonstrating how f_H is calculated is shown in Figure 6.5.

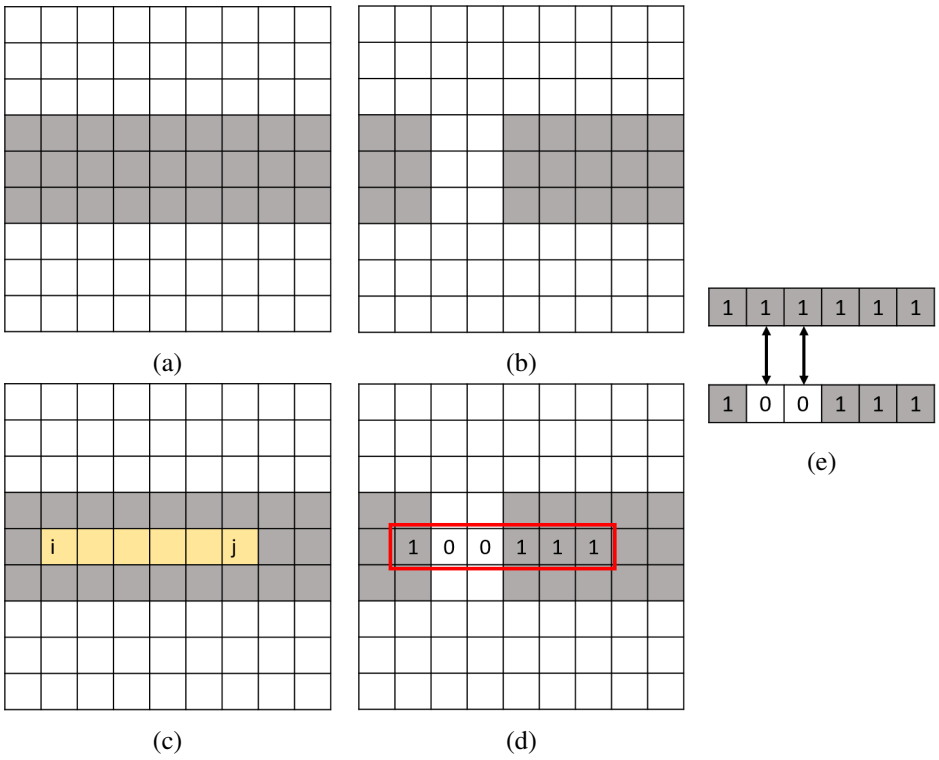


Figure 6.5: Hamming distance between a path $P_{i,j}$ sampled from a mask $\mathbf{A}_1$, and a mask $\mathbf{A}_2$. (a) $\mathbf{A}_1$, (b) $\mathbf{A}_2$, (c) a path $P_{i,j}$ sampled from $\mathbf{A}_1$, (d) the respective string obtained from $\mathbf{A}_2$, and (e) the number of switches needed to have both strings match, in this case, $H(P_{i,j},\mathbf{A}_2) = 2$, therefore $f_H = 4/6$.

The second possibility is a binary function that only outputs 1 if the entire path is possible in the mask or, in other words, if the Hamming distance is 0:

$$f_F(P_{i,j},\mathbf{A}) = \begin{cases} 1, & \text{if } H(P_{i,j},\mathbf{A}) = 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.14}$$

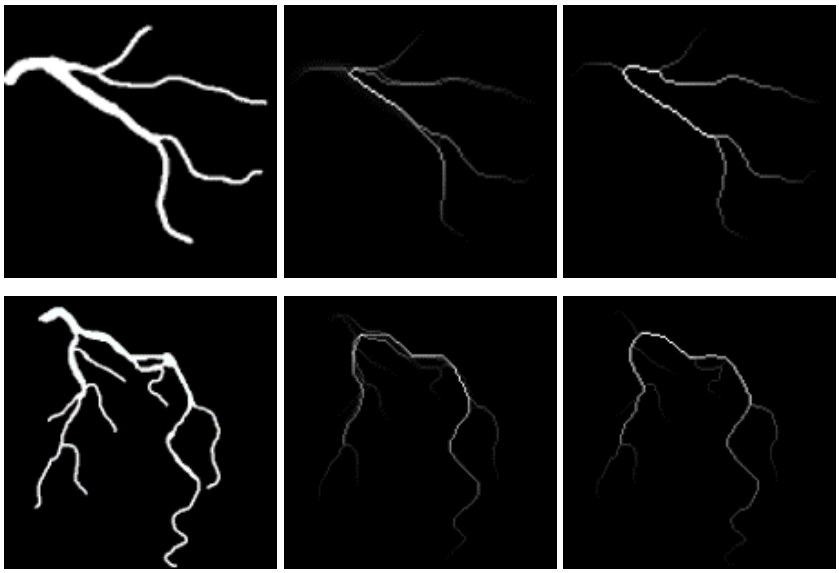
6.2.2 Practical considerations

There are two important considerations to have into account regarding the similarity index defined before: i) to obtain $P_{i,j}$ we resort to a minimum cost path algorithm, which is likely to follow the boundaries of blood vessels when precaution is not taken (see Fig. 6.6b). This is an unwanted behaviour as it would be very sensitive to calibre-based errors and that is not the goal we seek with the proposed similarity index; ii) the number of possible paths grows exponentially with the number of pixels in $\mathbf{A}_f$, with a worst case complexity of $O(n^2/2)$. This turns the use of (6.12) impractical to evaluate real-world images which always have a significant number of blood vessel pixels.

Regarding the tendency of minimum cost path based approaches to follow the vessel boundaries, we consider a simple possibility to address that issue. Instead of considering the complete

set of vessel pixels $\mathbf{A}_f$, we simply take into account the centreline pixels $\mathbf{A}_s$ and all the possible paths there. Concerning the second property we have defined earlier, this approach slightly reduces the relevance of larger vessels in comparison with the narrower ones. Even then, the mentioned natural properties of vascular trees (larger blood vessels ramify into smaller ones) promote centrelines of larger vessels to be visited more times (see Fig. 6.6c). Another option would be using a non-linear distance function between the foreground (vessel) and background pixels, in order to promote the minimum cost path to follow the centrelines of the blood vessels.

Figure 6.6 shows two coronary trees and the number of times each pixel is visited when using the discussed strategy for employing the proposed similarity index. As illustrated, the centreline pixels of the main branches of the vascular trees are visited more frequently, penalising more any error in these segments.



(a) (b) (c)

Figure 6.6: The relative frequency each pixel is visited when considering all the possible paths to be taken in a tree (a higher intensity is equivalent to a higher relative frequency). (a) Example coronary trees; (b) the frequency map when the minimum cost path between every two points of the tree is considered; (c) when the minimum cost path is retrieved from the centrelines for every two centreline points.

Concerning the exponential complexity of the number of paths to be extracted, we can resort to a Monte Carlo approach to approximate (6.12):

$$\tilde{m}(\mathbf{P},\mathbf{Y}) = \left(\frac{1}{n}\sum_{k=1}^{n} f(P_{i,j},\mathbf{P}),\ i,j \overset{i.i.d.}{\sim} U(\mathbf{Y}_s),\ l_i = l_j\ \cdot \frac{1}{n}\sum_{k=1}^{n} f(P_{i,j},\mathbf{Y}),\ i,j \overset{i.i.d.}{\sim} U(\mathbf{P}_s),\ l_i = l_j\right)^{1/2} \tag{6.15}$$

where $U(\mathbf{A}_s)$ denotes the Uniform distribution over a set of vessel pixels $\mathbf{A}_s$.

Regarding the example shown in Fig. 6.4, our proposed similarity index (6.12) evaluates $\mathbf{P}_1, \mathbf{P}_2$ and $\mathbf{P}_3$ with a score of, respectively, $0.984, 0.988$, and 0.933 (using f_H). This is aligned with the properties we seek for a topological benchmark.

6.3 A loss function for increasing topological coherence

BCE and soft Dice are among the most typically used loss functions in the blood vessel segmentation problem. BCE allows to weigh differently false positive and false negative errors, however it treats equally the errors inside each group. The FL [183] further increases the weight given to pixels where the error is larger, such that it can be interpreted as a hard mining technique. Again, the weight given to a pixel is proportional to the magnitude of the error and not its type. The soft Dice loss is interesting for binary problems where imbalance is relevant, which is the case of blood vessel segmentation, since typically vessel pixels are a minority of the total pixel count. Nevertheless, the effect of false positives and false negative errors is still the same independently of the location where they occur. The clDice loss [197], similarly to the behaviour of the corresponding metric, brings increased attention to the blood vessel centrelines, disregarding errors on the wall boundaries of vessels. In conjunction with the soft Dice loss, it was shown to promote the creation of segmentations which were more similar to the reference ones concerning topology. Yet, despite its usefulness, we argue that this loss is still not truly topological due to two reasons: i) it does not distinguish errors inducing disjoint trees and merging distinct ones from those happening at the terminating portions of blood vessel segments; ii) it does not distinguish errors in major vessels from those in minor ones. In this section, we describe our proposed loss function which, besides focusing on the centrelines of the vascular tree, is also capable of distinguishing the errors mentioned in Property 1.

6.3.1 Detecting errors that produce disjoint trees

We aim to design a loss function which, contrary to state-of-the-art losses, is able to highlight errors inducing disjoint trees, such that it can guide models towards producing blood vessel masks which are more topologically coherent. Errors leading to disjoint trees are nothing more than "holes" in a given blood vessel segment, therefore they may be filled by applying mathematical morphology operators, which have already been introduced in subsection 3.1.3.

Let us start by considering the simple case of a single blood vessel segment being affected by this type of errors, as illustrated in Figure 6.7. The disjoint segments can be connected by employing a closing operator (dilation followed by erosion) using a structuring element (SE) of sufficient size. Let $\mathbf{A}_{D(r)}$ and $\mathbf{A}_{C(r)}$ be, respectively, the output of the morphological dilation and closing of $\mathbf{A}$ using a squared SE with radius r. In the example provided in Figure 6.7, a radius of 2 would be necessary to connect all the disjoint segments.

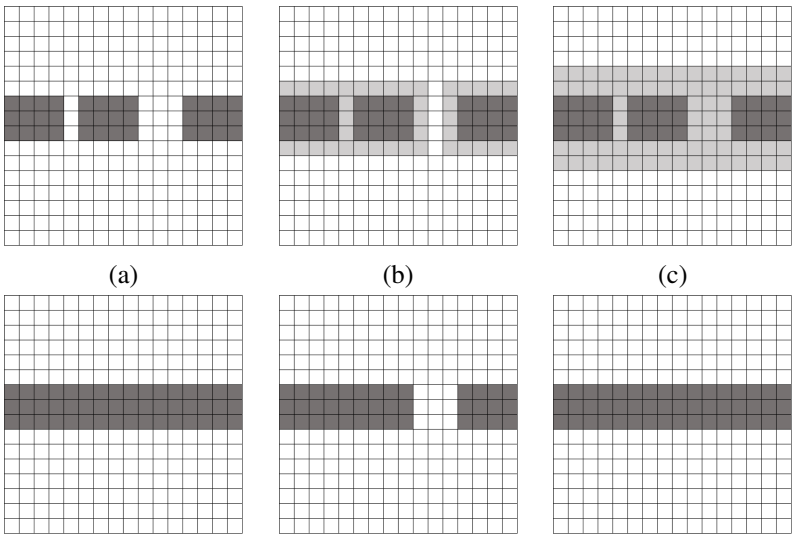


Figure 6.7: Connecting disjoint segments to recover the reference blood vessel segment, by means of a closing operation with a squared structuring element. (a) Predicted, $\mathbf{P}$, and (d) reference, $\mathbf{Y}$, segmentations; morphological dilation of $\mathbf{P}$ with a SE of radius (b) 1, $\mathbf{P}_{D(1)}$, and (c) 2, $\mathbf{P}_{D(2)}$, where light grey represents the appended pixels; outputs of the respective closing operations, (e) $\mathbf{P}_{C(1)}$ and (f) $\mathbf{P}_{C(2)}$.

Most medical images of blood vessels (or even patches, small portions of these images) contain tree-like structures, not single vessel segments as illustrated in the previous simplistic scenario. Hence, the closing operation might merge disjoint segments which should not be connected at all (see Figure 6.8), constituting an unwanted behaviour. This effect is exacerbated when a large SE is required to connect segments far apart. Fortunately, as long as we have a reference segmentation, $\mathbf{Y}$, which is the case of supervised blood vessel segmentation, this limitation is easily circumvented. To do so, we pose the problem as joining separate segments only if the missing segment exists in the reference mask. In our experiments, we consider the centrelines $\mathbf{Y}_s$ instead, in order to put a larger focus on the graph structure of the vascular trees. Hence, we detect missing segments inducing disjoint trees as follows:

$$\mathbf{e}(\mathbf{P},\mathbf{Y};r) = \left(\mathbf{P}_{C(r)} - \mathbf{P}\right)^2 \cdot \mathbf{Y_s} \tag{6.16}$$

Figure 6.9 illustrates this approach on the example provided in Figure 6.8.

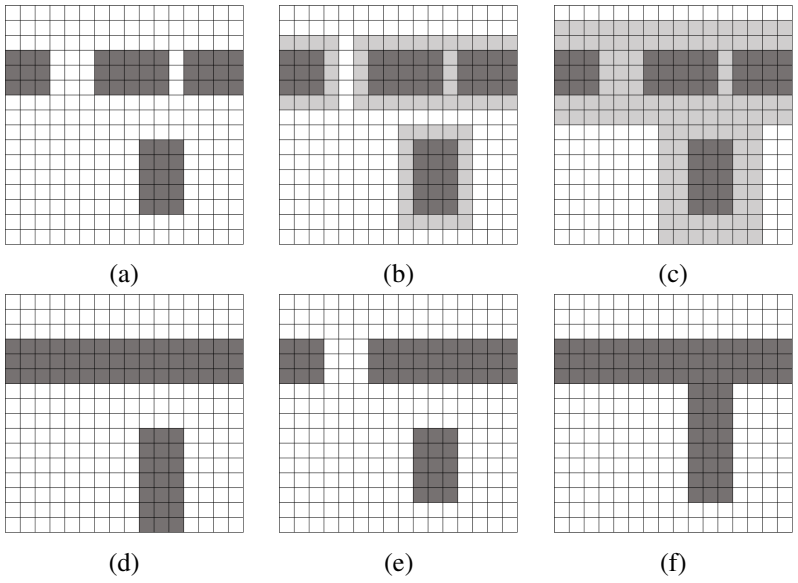


Figure 6.8: Connecting disjoint segments to recover the reference blood vessel tree, showcasing the possibility of joining segments which are not connected in the reference segmentation. (a) Predicted, $\mathbf{P}$, and (d) reference, $\mathbf{Y}$, segmentations; morphological dilation of $\mathbf{P}$ with SE of radius (b) 1, $\mathbf{P}_{D(1)}$, and (c) 2, $\mathbf{P}_{D(2)}$, where light grey represents the appended pixels; outputs of the respective closing operations, (e) $\mathbf{P}_{C(1)}$ and (f) $\mathbf{P}_{C(2)}$.

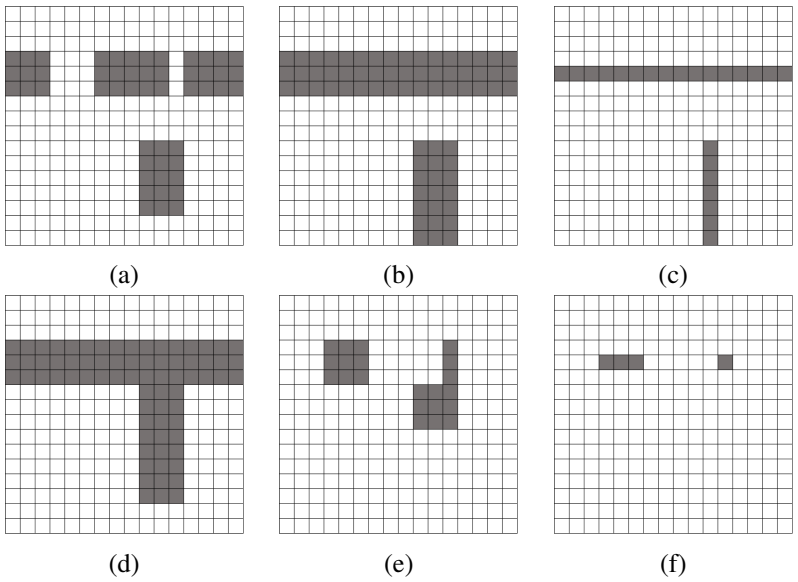


Figure 6.9: Detection of centreline errors inducing disjoint trees. (a) Predicted, $\mathbf{P}$, and (b) reference, $\mathbf{Y}$, segmentations; (c) reference centrelines, $\mathbf{Y}_s$; (d) morphological closing of the prediction, $\mathbf{P}_{C(2)}$; (e) Squared difference between the latter and the original prediction, $(\mathbf{P}_{C(2)} - \mathbf{P})^2$; and (f) the centreline errors that produce disjoint trees, according to the reference mask, $\left(\mathbf{P}_{C(2)} - \mathbf{P}\right)^2 \cdot \mathbf{Y}_s$.

6.3.2 Weighting errors of different size

Despite identifying the errors that originate disjoint trees, according to (6.16), their weight is being proportional to the length of the corresponding missing centreline. Therefore, following the example provided in Figure 6.9, the larger missing segment would end having three times the weight of the smaller one. We argue that this might not be ideal for learning purposes. In our opinion, the model should learn that it is more likely that a small missing segment is in reality a false negative, than segments which are farther apart. Therefore, to reduce this weight bias towards larger missing segments, we consider weight normalisation according to the length of the missing segment.

Let S_r be a missing segment that can be filled through a closing operation considering a SE with radius r or larger. The total error of this segment is given by:

$$e(S_r) = w_r \cdot l(S_r) \tag{6.17}$$

where w_r is a scalar, and $l(S_r)$ is the length of the segment, being either $2r-1$ or $2r$. We seek a normalisation such that the following inequality always holds:

$$w_r \cdot l(S_r) \geq w_{r+1} \cdot l(S_{r+1}) \tag{6.18}$$

which states that a missing segment that can be filled with a SE with radius r must have at least the same total error as a missing segment which can only be filled with a SE with radius $r+1$ or larger. By noticing again that the length of a missing segment S_r can have two different values, the following equality is sufficient to hold inequality (6.18):

$$w_r \cdot (2r-1) = w_{r+1} \cdot (2(r+1)) \tag{6.19}$$

Let us consider that the largest SE radius to be used is r_M and that its associated weight w_{r_M} is 1. It is now possible to consider an iterative approach that highlights all missing segments with length up to $2r_M$ and normalises their weights according to 6.19:

$$\mathbf{e}(\mathbf{P},\mathbf{Y};r_M) = \sum_{r=r_M,r_{M-1},\ldots,1} \varepsilon_r \cdot \left(\mathbf{P}_{C(r)} - \mathbf{P}\right)^2 \cdot \mathbf{Y}_\mathbf{s} \tag{6.20}$$

with ε_r given as:

$$\varepsilon_r = \begin{cases} w_{r_M}, & \text{if } r = r_M \\ w_r - w_{r+1} & \text{otherwise} \end{cases} \tag{6.21}$$

Figure 6.10 exemplifies how this normalising approach would work for the synthetic example shown in Figure 6.9, considering $r_M = 2$.

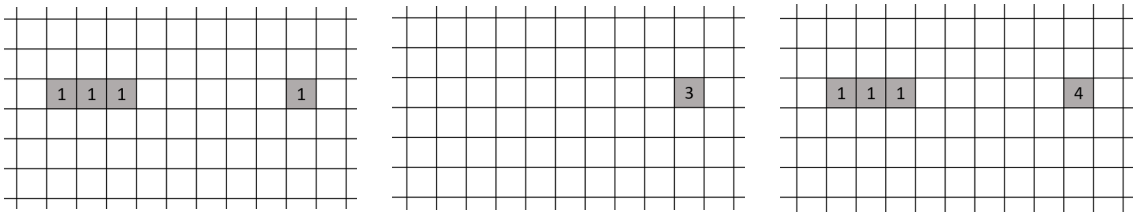


Figure 6.10: Demonstration of the error normalisation approach regarding the synthetic example considered in Figure 6.9. (a) Missing segments detected when using a SE with radius r_M, which was set to 2 in this example, and the corresponding $\varepsilon_2 = w_2 = 1$; (b) missing segments detected when using a SE with radius 1, with corresponding weight $w_1 = 4$, as given by (6.19), hence $\varepsilon_1 = w_1 - w_2 = 3$; (c) total error according to (6.20).

6.3.3 Design of a topological loss

In subsections 6.3.1 and 6.3.2, a method for highlighting errors inducing disjoint trees in a segmentation, according to a reference mask, was presented. Until now, false negative segments in the prediction segmentation have been focused. Nonetheless, it is also important to consider false positive detections joining segments which do not belong to the same tree in the reference mask. Note that this can be simply achieved via (6.20) by switching the roles of the predicted segmentation and the reference mask. Therefore, we define two loss terms, one that seeks to minimise the amount of missing segments that induce disjoint trees, according to the reference mask:

$$L_{tsens}(\mathbf{P},\mathbf{Y};r_M) = \frac{\left\| \sum_{r\in\mathbf{r}} \varepsilon_r \cdot (\mathbf{P}_{C(r)} - \mathbf{P}) \cdot \mathbf{Y}_s \right\|_1}{\|\mathbf{Y}_s\|_1} \tag{6.22}$$

and a second one, promoting that no false connections are introduced, again having into account the reference mask:

$$L_{tprec}(\mathbf{P},\mathbf{Y};r_M) = \frac{\left\| \sum_{r\in\mathbf{r}} \varepsilon_r \cdot (\mathbf{Y}_{C(r)} - \mathbf{Y}) \cdot \mathbf{P}_s \right\|_1}{\|\mathbf{P}_s\|_1} \tag{6.23}$$

where the denominators are introduced for normalising purposes only and not considered during the loss gradient calculation and, for simplicity of notation, $\mathbf{r} = r_M, r_{M-1}, \ldots, 1$. Both terms are combined into our proposed topological loss:

$$L_{topo}(\mathbf{P},\mathbf{Y};\alpha,r_M) = \alpha \cdot L_{tsens}(\mathbf{P},\mathbf{Y};r_M) + (1-\alpha) \cdot L_{tprec}(\mathbf{P},\mathbf{Y};r_M) \tag{6.24}$$

where $\alpha \in [0,1]$ determines the relative importance of each term.

In order to easily integrate the morphological operators into the model training procedure, we consider the neural dilation and erosion layers introduced in [198]. Note that, instead of having to design multiple neural layers to perform morphological closings with SE of different radius, we can simply cascade neural layers implementing morphological operations with a SE of radius 1.

This is possible since the following holds:

$$\mathbf{A}_{D(r)} = \mathbf{A}^{r}_{D(1)} \tag{6.25}$$

$$\mathbf{A}_{C(r)} = \left(\mathbf{A}_{D(r)}\right)^{r}_{E(1)} \tag{6.26}$$

where $\mathbf{A}^{r}_{M(1)}$ denotes r-consecutive uses of a morphological operator M with a SE with radius 1 in segmentation mask A. Algorithm 1 provides a pseudo-algorithm concerning the computation of the proposed loss.

Algorithm 1: Proposed topological loss function

input : ground truth y, prediction p, max radius r, weight α
output: proposed loss function L
$ratios \leftarrow [\,]$;
for $i \leftarrow 1$ **to** $r-1$ **do**
  $ratios.append\left(\frac{4+2i}{1+2i}\right)$;
end
$p\_s \leftarrow Skeleton(p)$;
$y\_s \leftarrow Skeleton(y)$;
$L\_tsens \leftarrow \mathbf{0}$;
$L\_tprec \leftarrow \mathbf{0}$;
$w \leftarrow 1$;
$w_{prev} \leftarrow 0$;

for $i \leftarrow r$ **to** 1 **do**
  $p\_tmp \leftarrow p$;
  $y\_tmp \leftarrow y$;
  for $j \leftarrow 1$ **to** i **do**
    $p\_tmp \leftarrow Dilation(p\_tmp)$;
    $y\_tmp \leftarrow Dilation(y\_tmp)$;
  end
  for $j \leftarrow 1$ **to** i **do**
    $p\_tmp \leftarrow Erosion(p\_tmp)$;
    $y\_tmp \leftarrow Erosion(y\_tmp)$;
  end
  $L\_tsens \leftarrow L\_tsens + (w - w\_prev) \cdot (p\_tmp - p) \cdot y\_s$;
  $L\_tprec \leftarrow L\_tprec + (w - w\_prev) \cdot (y\_tmp - y) \cdot p\_s$;
  $w\_prev \leftarrow w$;
  $w \leftarrow w \cdot ratios[i-1]$;
end

$L_{topo} \leftarrow \alpha \cdot \frac{\|L\_tsens\|_1}{\|y\_s\|_1} + (1-\alpha) \cdot \frac{\|L\_tprec\|_1}{\|p\_s\|_1}$;

6.3.4 Experiments and Discussion

To assess the contribute of the proposed topological loss, L_{topo} (6.24), in the scenario of blood vessel segmentation, we analysed how the performance of a U-Net model [23] varies according to the loss function that is minimised during training.

The first two baseline loss functions considered are the typically used soft Dice loss, L_{dice}, and the BCE loss, L_{bce}:

$$L_{dice}(\mathbf{P},\mathbf{Y}) = 1 - 2 \cdot \frac{\sum_i \mathbf{P}_i \cdot \mathbf{Y}_i}{\sum_i \mathbf{P}_i^2 + \mathbf{Y}_i^2} \tag{6.27}$$

$$L_{bce}(\mathbf{P},\mathbf{Y};\alpha) = -\frac{1}{N}\sum_{i=1}^{N}\Big(\alpha \cdot \mathbf{Y}_i \cdot \log(\mathbf{P}_i) + (1-\alpha)\cdot(1-\mathbf{Y}_i)\cdot\log(1-\mathbf{P}_i)\Big) \tag{6.28}$$

where N is the total number of pixels in the image, and $\alpha \in [0,1]$ determines the relative weight of each class (blood vessel and background).

We also consider the FL [183]:

$$\begin{aligned} L_{focal}(\mathbf{P},\mathbf{Y};\alpha,\gamma) = -\frac{1}{N}\sum_{i=1}^{N}\Big(&\alpha \cdot \mathbf{Y}_i \cdot (1-\mathbf{P}_i)^{\gamma} \cdot \log(\mathbf{P}_i) + \\ &(1-\alpha)\cdot(1-\mathbf{Y}_i)\cdot\mathbf{P}_i^{\gamma}\cdot\log(1-\mathbf{P}_i)\Big) \end{aligned} \tag{6.29}$$

with γ controlling the non-linear relation between the magnitude of the error and the weight.

Given the results achieved by the inclusion of the clDice loss in [197], we also test the loss proposed in [197] as a baseline:

$$L_{cldice}(\mathbf{P},\mathbf{Y};\alpha) = \alpha \cdot L_{dice}(\mathbf{P},\mathbf{Y}) + (1-\alpha)\cdot(1 - clDice(\mathbf{P},\mathbf{Y})) \tag{6.30}$$

where $clDice(\mathbf{P},\mathbf{Y})$ is the criterion defined in (6.9), and α determines the relative weight given to each loss component. A loss mimicking the ideas behind L_{cldice}, but using the BCE criterion instead, was also regarded in the experiments:

$$\begin{aligned} L_{clbce}(\mathbf{P},\mathbf{Y};\alpha;\beta) = -\frac{1}{N}\sum_{i=1}^{N}\Big(&(a\cdot\mathbf{Y}_i + \beta\cdot\mathbf{Y}_{s_i})\cdot\log(\mathbf{P}_i) + \\ &\big((1-\alpha)+(1-\beta)\cdot\mathbf{P}_{s_i}\big)\cdot(1-\mathbf{Y}_i)\cdot\log(1-\mathbf{P}_i)\Big) \end{aligned} \tag{6.31}$$

with β controlling the compromise between false positive and false negative centreline detections.

Having described the losses that will be used as benchmark in the experiments, the proposed loss function is now detailed. Since we recognise the relevance of emphasising the centrelines of the vessels, we let our proposed loss build upon L_{cldice} (6.30) and L_{clbce} (6.31), and extend them by including an additional term pertaining to the loss term we proposed in subsection 6.3.3:

$$L_{propdice}(\mathbf{P},\mathbf{Y};\alpha_1,\alpha_2,c,r) = L_{cldice}(\mathbf{P},\mathbf{Y};\alpha_1) + c \cdot L_{topo}(\mathbf{P},\mathbf{Y};\alpha_2,r) \tag{6.32}$$

$$L_{propbce}(\mathbf{P},\mathbf{Y};\alpha_1,\alpha_2,\beta,c,r) = L_{clbce}(\mathbf{P},\mathbf{Y};\alpha_1,\beta) + c \cdot L_{topo}(\mathbf{P},\mathbf{Y};\alpha_2,r) \tag{6.33}$$

where c sets the relevance of the proposed topological term in the overall loss, α_2 controls the compromise between L_{tsens} and L_{tprec}, and r defines the maximum radius to be considered in the involved closing operations.

6.3.4.1 Datasets and Metrics

Regarding the data used in the experiments, in addition to the retinal vessel segmentation benchmarks commonly used (DRIVE [38], STARE [39], and CHASEDB1 [40]), we have also considered a dataset containing coronary angiograms [46].

Concerning metrics and similarity indices to assess the performance of the different models, we consider not only the typically used metrics (AUC, accuracy, sensitivity, and specificity), but also the clDice score [197], and both variants of the proposed approximate topological similarity index (6.15), $\tilde{m}_H$ and $\tilde{m}_F$. The Monte Carlo approximation was performed with $n = 1000$.

6.3.4.2 Implementation details

From the 20 images comprising the training set of DRIVE, 4 were set aside for validation purposes. The original test set remained unchanged and was used for that stage. Regarding STARE, from a total of 20 images, 3 and 5 were reserved for, respectively, validation and testing phases. Concerning CHASE, from the available 28 images, 4 and 10 were used for the validation and test steps. Finally, for the CORONARY dataset, which contains 134 coronary angiograms, we considered 20 and 30 images for validating and testing, respectively, the trained models. Throughout training, small patches of dimensions 128×128 were fed to the model, whereas, during validation and testing, the entire images were given. Data augmentation during training comprised flipping and rotation transformations, and the addition of an intensity bias.

The U-Net model [23] was implemented according to its original description and 100 batches containing 8 patches each were fed every epoch. The Adam optimiser [184] was used to update the model parameters, with an initial learning rate of $1e{-}4$. The loss in the validation set was measured every 25 epochs. Every 50 epochs the learning rate was decreased by a factor of 0.1, being the training process terminated after 200 epochs. We kept the set of parameters leading to a minimum loss over the evaluations performed in the validation set. To regularise the weights

of the model, excluding the bias of neurons, we considered L_2-regularisation with a coefficient of 1e−5.

Two different configurations of L_{bce} were considered, one where an equal weight was given to both classes ($\alpha = 0.5$, unweighted), L_{bceu}, and another one weighting significantly more errors in blood vessel pixels ($\alpha = 0.7$, weighted), L_{bcew}. This weight has been found appropriate in previous experiments. Concerning L_{focal}, we consider a single parametrisation, $\alpha = 0.7, \gamma = 2$. L_{cldice} was configured following [197], by setting $\alpha = 0.5$. In our proposed variant, $L_{propdice}$, we also consider $\alpha_1 = 0.5$, then we set $r = 10$, $c = 0.1$, and test different values for α_2. The centreline aware BCE variants, L_{clbceu}, and L_{clbcew} were parametrised with $\beta = 0.5$, and α of, respectively, 0.5 and 0.7. Our proposed variants, $L_{propbceu}$ and $L_{propbcew}$, consider $\beta = 0.5$, $r = 10$, $c = 0.1$ and, respectively, $\alpha_1 = 0.5$ and $\alpha_1 = 0.7$. Different values are tested for α_2.

To disregard randomness involved in training using a GPU and other processes such as batch generation, we have performed deterministic training and evaluation. To further highlight the differences between the losses, we have run 2 times each of them, by picking 2 different seeds, guaranteeing that among experiments with the same seed, everything was constant except the loss function being optimised. An NVIDIA GeForce RTX 2080 Ti GPU was used to conduct the experiments.

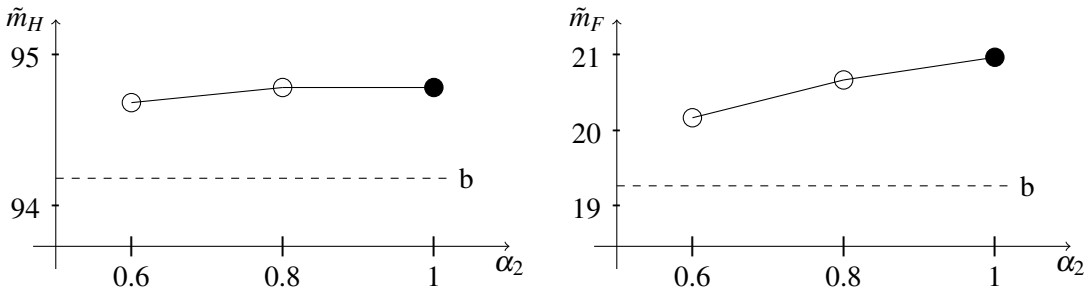
6.3.4.3 Results and discussion

The average value of the performance obtained when minimising the different losses, concerning the metrics specified in subsection 6.3.4.1, is shown in Table 6.2.

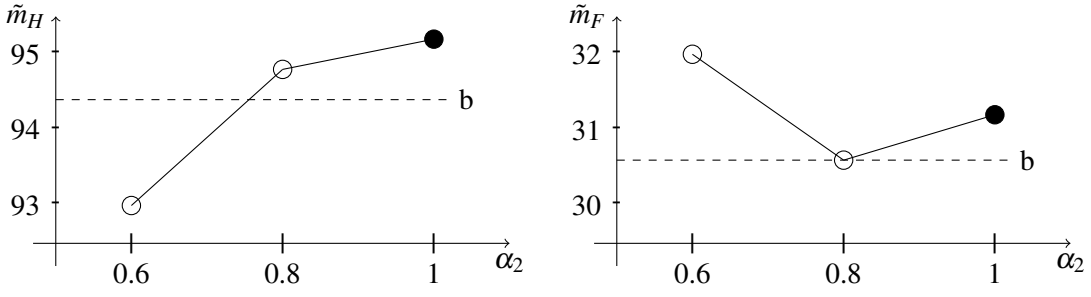
For each loss that includes the proposed cost function and for each dataset, we show only the results for the parametrisation achieving larger average value of $\tilde{m}_H$ and $\tilde{m}_F$. The best parametrisations are highlighted with a filled circle in the graphics concerning the ablation studies. Figures 6.11, 6.12, and 6.13 depict the ablation studies for $L_{propdice}, L_{propbceu}$, and $L_{propbcew}$, respectively. It is possible to conclude that the ideal value for α_2 varies according with the dataset and the loss that is being considered. The larger the α_2 value, the larger the focus on errors inducing disjoint trees over those merging separate trees. Even though a pattern is not clear for $L_{propdice}$, $L_{propbceu}$ and $L_{propbcew}$ seem to benefit from, respectively, larger and smaller α_2 values. The weighted BCE loss already weights significantly blood vessel locations, such that it seems plausible that a big focus on errors inducing disjoint trees would hurt the equilibrium between these errors and those merging distinct trees. Following the same rationale, models based on the unweighted BCE typically have difficulty in effectively capturing some of the vessel branches; therefore, it is likely that they benefit more from the use of large α_2 values. Despite this variability regarding the optimal α_2, the ablation studies show that, for a particular family of models, it is possible to find α values which lead to improvements in all the datasets when compared with the centreline-aware baselines. By analysing the graphics, it is also possible to conclude that there is not a direct relation between $\tilde{m}_H$ and $\tilde{m}_F$, such that it is possible to have models which improve one of them while having their performance decreased in the other one. This shows that, even though our proposed general similarity index (6.12) has fixed properties related to the graph of

Table 6.2: Performance of the models, in percentage, averaged over 2 runs. AUC, *acc*, *sen*, *spe* stand for, respectively, area under the roc curve, accuracy, sensitivity, and specificity. clDice is the metric proposed in [197], and $\tilde{m}_H$ and $\tilde{m}_F$ are the two approximate topological similarity indices presented in subsection 6.2.1. The larger all of these indicators, the better the performance of the model. The best obtained performance for each indicator, model family, and considered database is highlighted in bold. The best model for a given database and indicator is underlined.

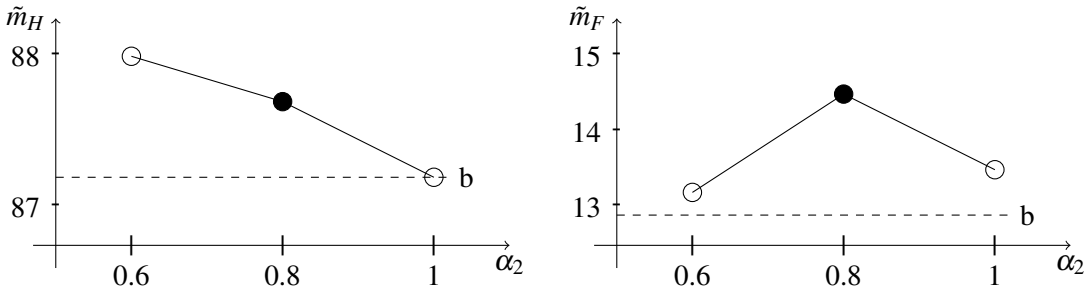
		L_{dice}	L_{cldice}	$L_{propdice}$	L_{bceu}	L_{clbceu}	$L_{propbceu}$	L_{bcew}	L_{clbcew}	$L_{propbcew}$	L_{focal}
DRIVE	AUC	**97.5**	95.8	96.0	**97.8**	97.6	97.6	**97.6**	**97.6**	97.4	**97.8**
	acc	**95.2**	94.6	93.5	**95.5**	95.2	94.0	**94.8**	94.4	93.6	95.0
	sen	83.8	86.0	**88.5**	80.0	80.2	**88.5**	85.3	87.0	**88.2**	85.2
	spe	**96.9**	95.8	94.3	**97.8**	97.5	94.8	**96.2**	95.6	94.4	96.4
	clDice	82.2	84.4	**85.2**	81.4	81.6	**84.6**	82.6	83.0	**83.4**	82.8
	$\tilde{m}_H$	93.0	94.2	**94.8**	93.0	92.8	**94.3**	93.0	92.8	**94.6**	93.1
	$\tilde{m}_F$	18.8	19.3	**21.0**	18.9	17.0	**19.4**	18.3	17.6	**20.4**	18.1
STARE	AUC	**98.6**	96.6	96.8	**98.6**	**98.6**	**98.6**	**98.8**	**98.8**	98.7	98.5
	acc	**97.0**	96.8	96.6	**97.0**	**97.0**	96.4	**96.6**	96.5	95.8	96.2
	sen	84.8	84.0	**86.3**	81.7	83.3	**89.2**	88.8	89.8	**92.0**	85.4
	spe	**98.2**	98.0	97.6	**98.6**	98.4	97.2	**97.4**	97.2	96.1	97.4
	clDice	87.3	87.7	**88.6**	86.7	87.5	**88.4**	87.6	87.7	**88.4**	84.7
	$\tilde{m}_H$	94.6	94.4	**95.2**	94.4	94.7	**95.7**	94.6	95.2	**95.6**	91.0
	$\tilde{m}_F$	29.4	30.6	**31.2**	30.8	30.3	**34.6**	31.0	34.0	**35.6**	25.1
CHASEDB1	AUC	**97.2**	94.4	94.8	97.8	**98.0**	97.8	**97.3**	96.8	96.8	97.0
	acc	**95.5**	95.4	94.6	**96.2**	**96.2**	95.8	**94.8**	94.1	93.6	94.2
	sen	80.9	80.5	**85.0**	80.6	82.3	**82.4**	**85.0**	83.8	84.3	83.6
	spe	**97.0**	96.8	95.5	**97.8**	97.6	97.2	**95.8**	95.2	94.6	95.4
	clDice	78.0	80.0	**80.6**	80.8	**81.9**	80.8	**77.2**	73.8	76.3	74.0
	$\tilde{m}_H$	85.6	87.2	**87.7**	89.6	89.4	**90.1**	85.0	82.4	**85.9**	82.1
	$\tilde{m}_F$	12.5	12.9	**14.5**	15.9	15.0	**17.0**	11.6	9.2	**14.2**	8.9
CORONARY	AUC	**98.4**	96.9	96.8	98.8	**98.9**	98.8	99.0	**99.1**	98.9	99.0
	acc	**97.6**	97.4	97.0	**97.7**	**97.7**	96.8	97.1	**97.2**	96.4	96.6
	sen	81.8	84.0	**87.0**	80.2	79.6	**90.0**	89.6	90.4	**92.6**	87.4
	spe	**98.5**	98.2	97.6	98.7	**98.8**	97.3	**97.6**	**97.6**	96.6	97.2
	clDice	84.6	**85.6**	85.4	84.0	**84.2**	84.1	83.4	84.2	**84.8**	78.4
	$\tilde{m}_H$	89.3	90.1	**90.6**	89.4	89.8	**90.4**	89.4	90.4	**91.1**	84.8
	$\tilde{m}_F$	39.3	42.0	**43.2**	39.2	40.4	**42.8**	41.0	**42.6**	41.9	40.2



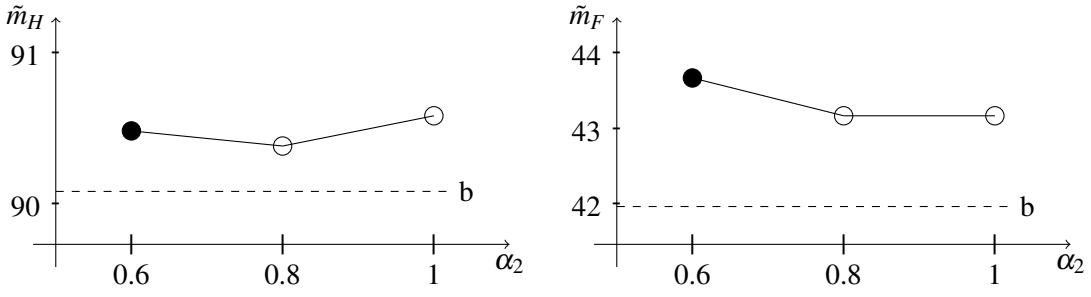
(a) DRIVE



(b) STARE

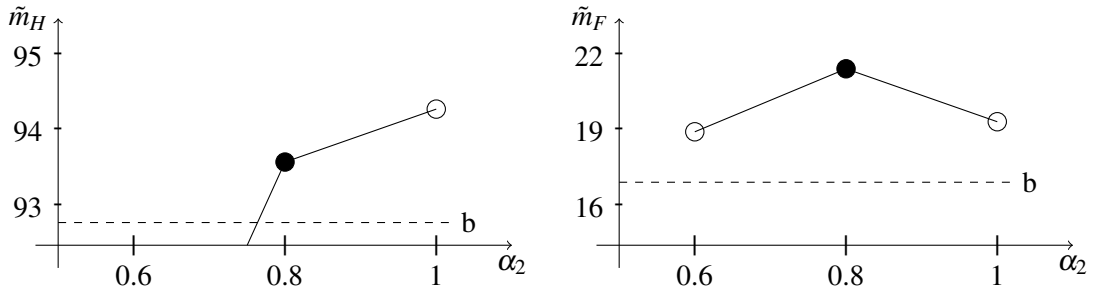


(c) CHASE

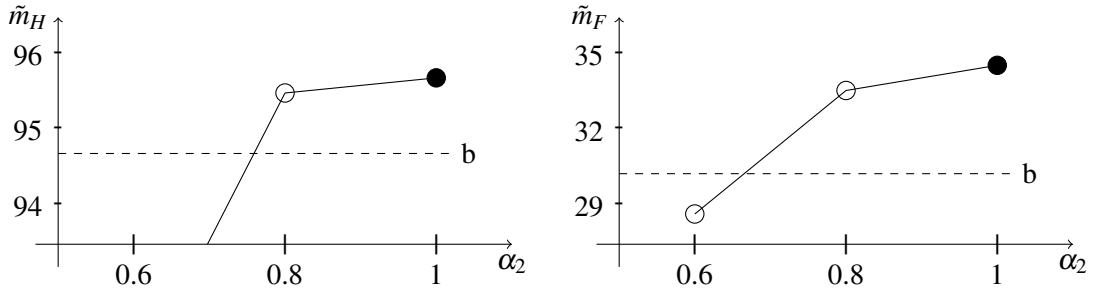


(d) CORONARY

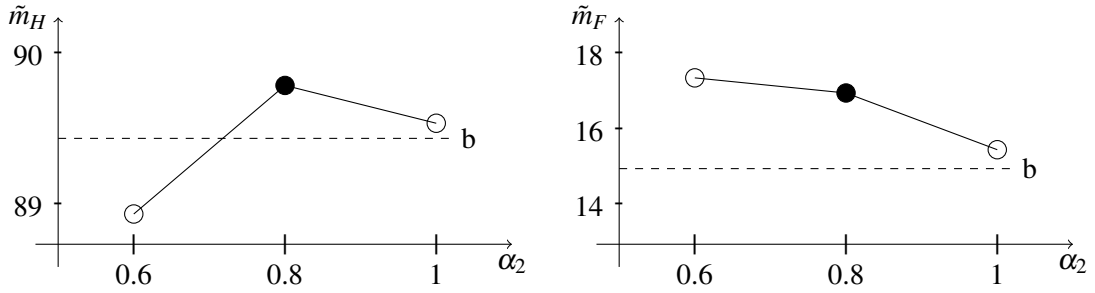
Figure 6.11: Ablation studies concerning the proposed model $L_{propdice}(\mathbf{P},\mathbf{Y};0.5,\alpha_2,0.1,10)$. The baseline performance of $L_{cldice}(\mathbf{P},\mathbf{Y};0.5)$ is represented by the dashed line.



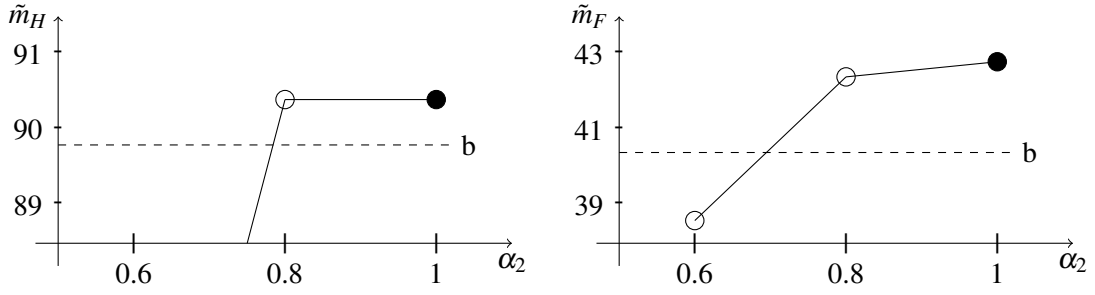
(a) DRIVE



(b) STARE

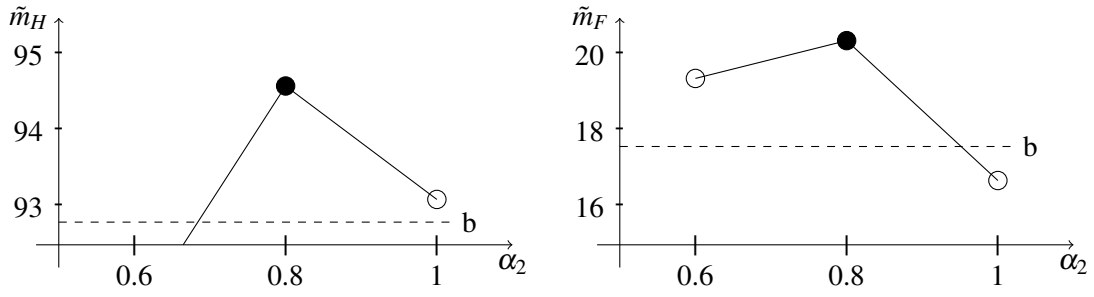


(c) CHASE

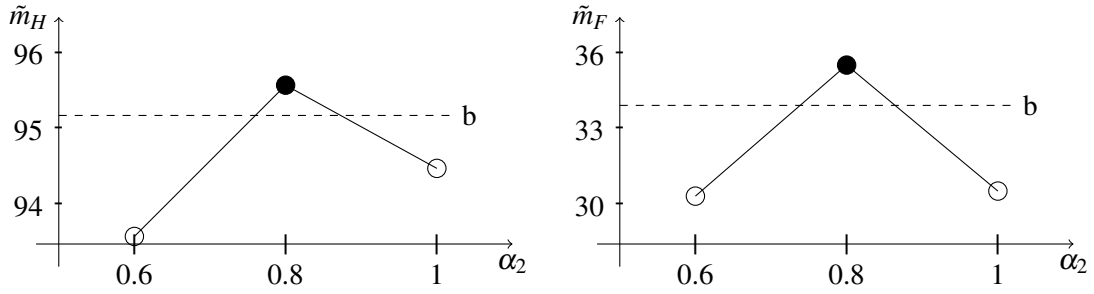


(d) CORONARY

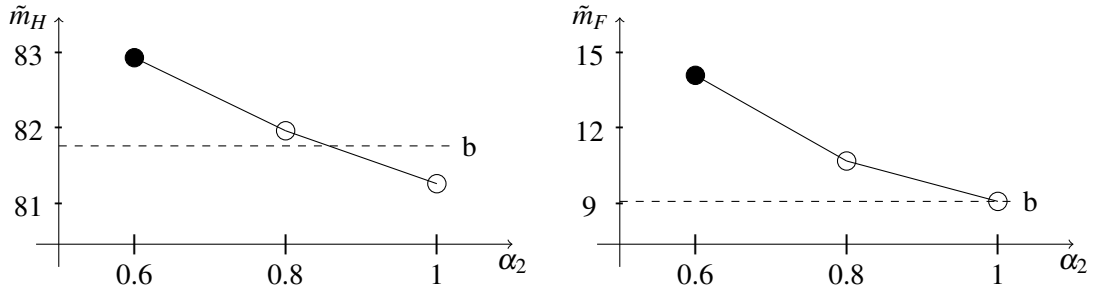
Figure 6.12: Ablation studies concerning the proposed model $L_{propbce}(\mathbf{P},\mathbf{Y};0.5,\alpha_2,0.5,0.1,10)$. The baseline performance of $L_{clbce}(\mathbf{P},\mathbf{Y};0.5,0.5)$ is represented by the dashed line.



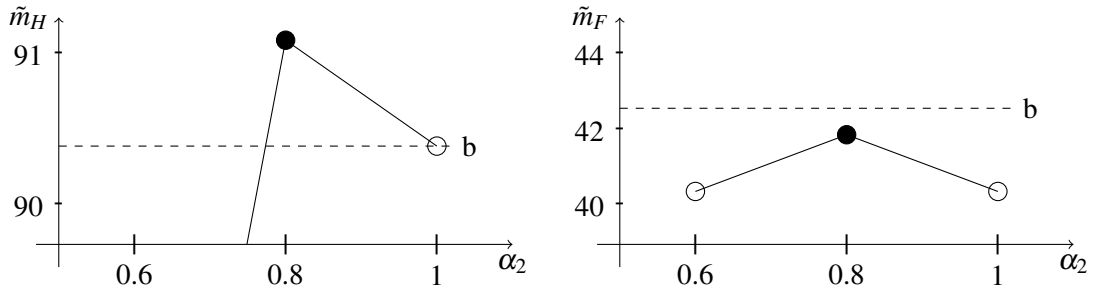
(a) DRIVE



(b) STARE



(c) CHASE



(d) CORONARY

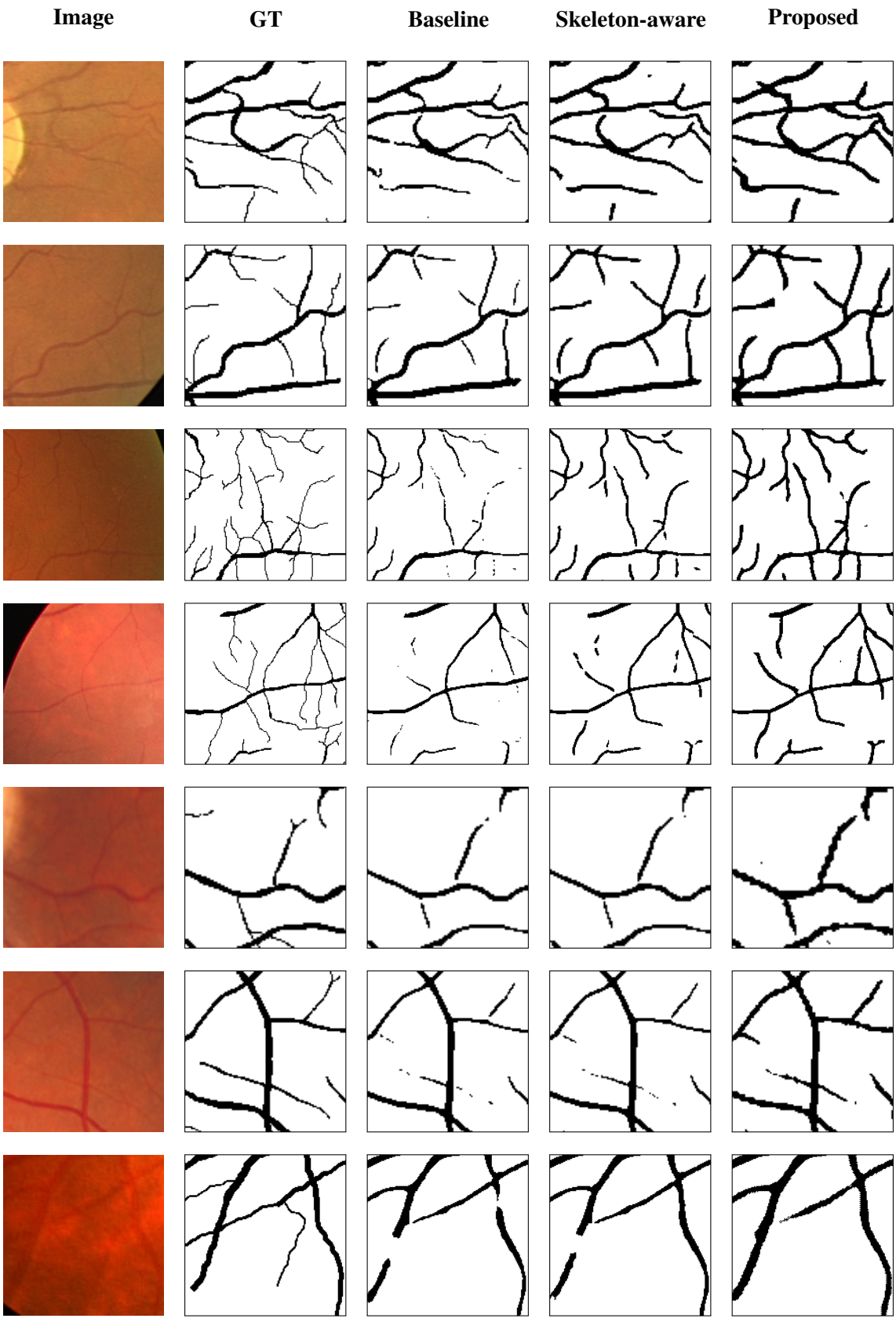
Figure 6.13: Ablation studies concerning the proposed model $L_{propbce}(\mathbf{P},\mathbf{Y};0.7,\alpha_2,0.5,0.1,10)$. The baseline performance of $L_{clbce}(\mathbf{P},\mathbf{Y};0.7,0.5)$ is represented by the dashed line.

the vascular trees, the considered function $f(P_{i,j}, \mathbf{A})$ also plays a role on the properties that are highlighted during the benchmark.

Analysing the patterns found for each of the loss function families, and starting with the Dice one, L_{dice}, L_{cldice}, and $L_{propdice}$, the impact of the soft clDice loss in the AUC is evident, since the models that include its minimisation have a significant decrease in the AUC metric. The accuracy was also slightly higher in the L_{dice} experiments; however, we believe this is not a crucial metric for blood vessel segmentation, as class imbalance exists, and models achieving better compromises between sensitivity and specificity tend to have lower accuracies. This trend is observable in the obtained results, since L_{cldice} and, especially, $L_{propdice}$, achieve higher sensitivity at the cost of decreased specificity, thus having lower accuracies. Regarding the proposed topological similarity indices, the runs minimising $L_{propdice}$ were the best performing ones. One interesting finding was that the inclusion of the topological term L_{topo} in $L_{propdice}$ also lead frequently to the increase of the clDice metric.

Concerning the BCE-based families, focusing the centrelines had only a negligible impact in the AUC. The remaining patterns follow the trends already discussed for the Dice family, with the proposed variants being once again the ones producing segmentations which are better topology-wise, according to the proposed similarity indices. An exception occurred in the Bcew family, where the proposed extension did not perform strictly better in the CORONARY database. The FL considered in the experiments, an extension of BCEw, did not achieve any benefit regarding the topological benchmarks.

The inferior performance of models belonging to the Dice and Bcew families in the CHASE dataset, when comparing with the ones from the Bceu family, was due to converging problems in some of the experimental runs. Having this in mind, and according to all of the obtained results, the best model to pick would likely be the one trained with $L_{propbce}(\mathbf{P}, \mathbf{Y}; 0.5, 1, 0.5, 0.1, 10)$. This model systematically performed better than the baselines topological-wise without disturbing significantly the pixel-wise metrics such as the AUC. Figure 6.14 shows example segmentations that can be obtained when learning a U-Net [23] with the different loss functions. For each image, we show the outputs of the family of models that achieved the best performance (average between $\tilde{m}_H$ and $\tilde{m}_F$) on its dataset. For example, regarding the images coming from the CORONARY dataset, where the best performing model was the one trained with $L_{propdice}$, we show images for L_{dice}, L_{cldice}, and $L_{propdice}$. From the qualitative assessment of the visual results, it is possible to verify that the proposed loss term helped the model to produce segmentations with less errors that lead to disjoint trees, without introducing a significant number of errors that join distinct trees. The skeleton-aware baselines seem to have a tendency to overestimate the calibre of narrow segments, an effect that seems to increase even a bit more when extending with the proposed loss term. This behaviour may introduce an error concerning the calibre estimation of narrow blood vessels and should be further investigated in the future.



(Continues on the next page)

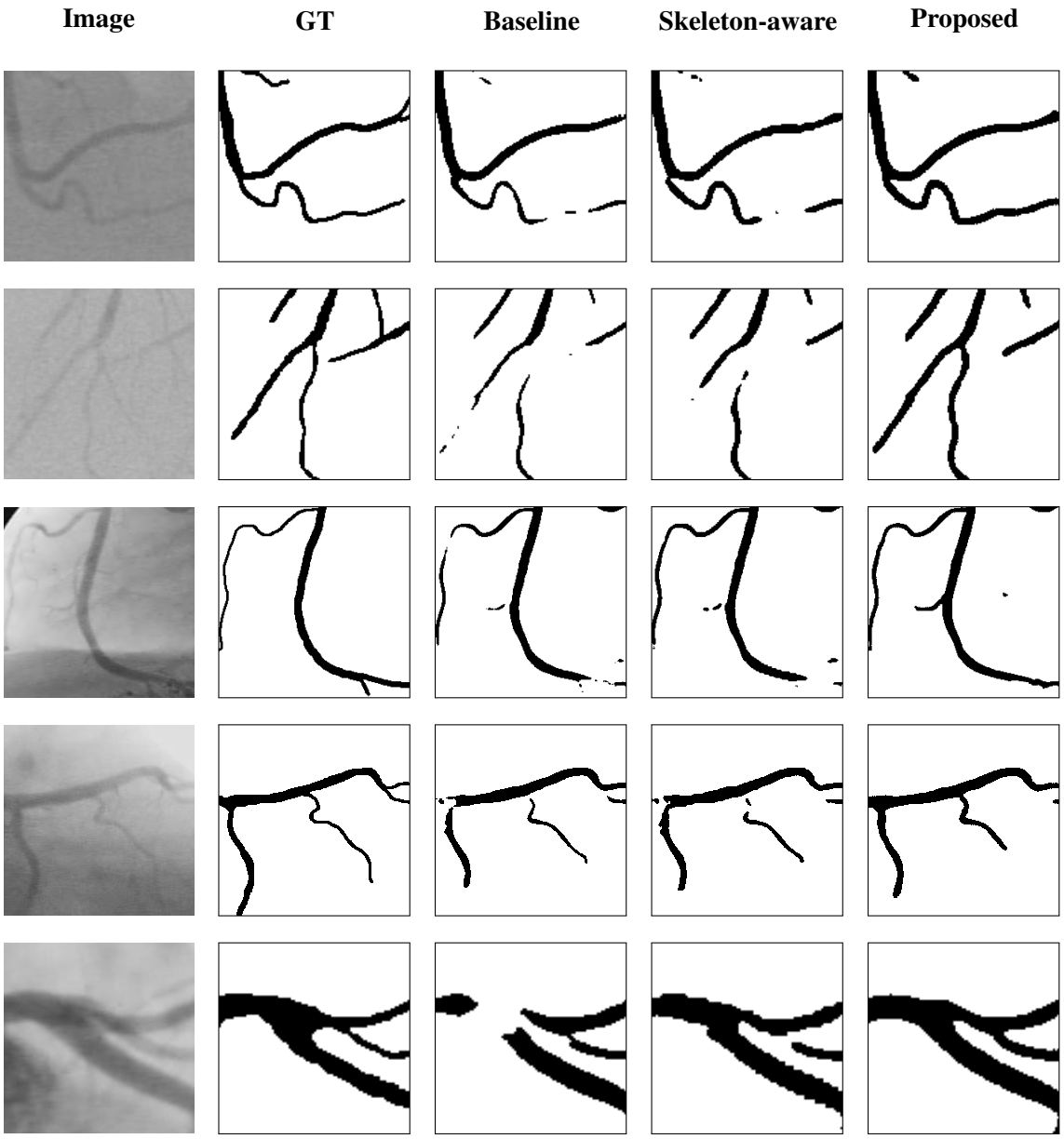


Figure 6.14: Example image patches, their ground truth, and segmentations achieved when minimising the typical baseline loss functions (L_{dice}, L_{bce}), the centreline-aware baseline loss functions (L_{cldice}, L_{clbce}), and the extension of the latter with the proposed loss function ($L_{propdice}, L_{propbce}$). Segmentations from the best performing model in the particular dataset are shown. Coloured patches are from retinal images while the grey-scale ones come from the images depicting the coronary tree. Notice that the masks are inverted for visualisation purposes.

6.3.5 Summary

In this Section, we described a loss term we have proposed for penalising two types of errors we deem as relevant in a topological loss: i) missing segments in the predicted segmentation that lead to disjoint trees (taking into account the topology of the reference mask); and ii) false positive segments in the predicted segmentation that join distinct trees (again, according to a reference mask). To do so, the proposed loss makes use of the morphological closing operation. In addition to the typical metrics that are used to evaluate blood vessel segmentations (AUC, accuracy, sensitivity, and specificity), the experiments considered the recently proposed clDice metric [197], which focuses blood vessel centrelines, and the approximated topological similarity indices presented in Section 6.2. The goal was to assess how useful the proposed loss term (6.24) is to promote segmentation models to produce masks that are better in terms of topological properties. The results of the experiments showed that this novel loss was indeed capable of improving the topological properties of the produced masks, relatively to the reference ones.

6.4 Main contributions and final considerations

Despite the breakthrough that has been happening due to deep learning and the continuous increase in the performance of these algorithms, it is still somewhat likely that we can distinguish segmentations produced by a human expert from those coming from a computer routine. The differences come mostly from the very strong prior knowledge that an expert human has regarding the concept of a blood vessel tree, and how difficult it is to encode those high level concepts in a complex model while expecting it to have the generalisation capability of a human. A type of error that continues to affect the performance of deep neural networks is the non-detection of particularly challenging blood vessel segments (poor local contrast, brightness, overlapping structures, among other reasons), leading to disjoint trees and, therefore, having a great impact in the overall graph of the blood vessel tree. There is also the possibility of detecting false segments and erroneously join trees. We designate these as topological errors and one of the main motivations of this thesis was to conduct research in order to find methodologies which could alleviate this problem. Increasing the capability of deep neural networks to deal with these topological errors has been gaining awareness lately [192; 197; 199].

Our first approach to this problem was from the model architecture point of view. By modelling errors affecting segmentation masks as noise, we have considered a probabilistic auto-encoding model to learn the most likely local blood vessel tree topology given a corrupted one. We have shown how this interpretation relates to a DVAE [195]. The conducted experiments shown that this design successfully restored some of the topological errors, as concluded by the decrease of infeasible paths and simultaneous growth of correct paths. These benefits were not achieved when replacing the probabilistic refiner with a second U-shaped network of similar complexity.

One of the things that became apparent in this first work, and also in the research conducted by other authors [192; 197], was the lack of a single unified metric or similarity index which could

be used to assess the topological properties of the segmented blood vessel trees. This has motivated us to derive novel similarity indices having properties which we deem as obligatory for assessing topological coherence: i) to clearly highlight the errors that have impact on the vascular tree graphs; ii) to further penalise missing segments inducing disjoint trees and false segments wrongly joining trees; and iii) topological errors in the main vascular tree branches should have larger impact. Additionally, our first work ended requiring increased model complexity to overcome some of the topological errors. Therefore, our second approach to tackle the problem at hand targeted the derivation of a loss whose minimisation promoted learning models which were less affected by topological errors. Since there already exist loss functions in the state-of-the-art promoting blood vessel centreline consistency (which, despite focusing the graph structure of the vascular tree, do not particularly penalise the errors that most contribute to graph changes), we focused on the penalisation of the errors described in point ii) above (which are the ones inducing the biggest graph changes). To find these errors, we developed a framework involving the morphological closing operation, and proposed a normalisation function based on the length of the error. Our experiments have shown that including this topological loss term during model training leads to obtaining masks which are closer to the reference ones topological-wise, according to an approximation of the proposed topological similarity indices.

We hope that the research conducted in this topic brings more awareness to the need of, not only improving the resilience to these topological errors in future approaches, but also reporting how good a given methodology is topological-wise. We stress that, with the continuous evolution methodologies will keep facing, only reporting the typical metrics (AUC, accuracy, sensitivity, and specificity) will be less sufficient as time goes by, since many properties of the produced segmentations will be overlooked in that case.

Chapter 7

Sparse Multi-Bending Snakes

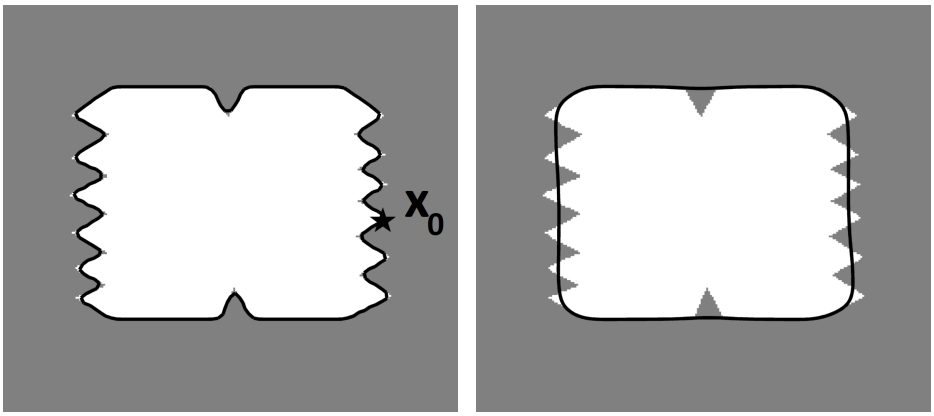
The content of this Chapter is based on the following work:

- R. J. Araújo, K. Fernandes, and J. S. Cardoso, "Sparse Multi-Bending Snakes", *IEEE Transactions on Image Processing*, 2019.

ACMs are one of the most emblematic algorithms of computer vision. Their strong theoretical foundations and high user interoperability turned them into a reference approach for object segmentation and tracking tasks. A brief overview on how these models segment objects and how they can be categorised according to their representation (explicit/parametric vs implicit) and guiding features (edge- vs region-based) can be found in subsection 3.1.6.

In this Chapter, we build upon traditional parametric ACMs (tACMs from now on). Since the introduction of ACMs [109], different contributions have been showing mechanisms to overcome some of their limitations. To deal with ill defined forces far away from edges and poor convergence to concavities, balloon forces [200] were introduced. Nonetheless, this solution requires prior knowledge related to which direction should the force be acting and weak edges are easily overpassed due to the pressure, leading to erroneous boundaries when portions of the object are not well defined. Another method addressing these issues was later proposed, establishing a new force field [201], known as Gradient Vector Flow (GVF). Other force fields have been proposed meanwhile [202]. Further improvement of robustness to noise and initialisation was achieved with the proposal of a decoupled active contour [203], where alternate steps of external energy minimisation and prior-based smoothing are taken.

The state-of-the-art ACMs do not allow, however, to find the correct boundary of objects having variable dynamics along the contour, where some of them occur due to the presence of artefacts. Consider, for instance, the object shown in Figure 7.1. Suppose that the concavities present in its top and bottom boundaries are due to an artefact (noise, object overlapping, image acquisition issue, or related problems), such that the true boundary of the object is smooth at those regions. A tACM is unable to produce the desired result, as the rigidity is constant along the entire snake. High rigidity coefficients produce a smooth boundary that cannot fit correctly any contour dynamics, while low rigidity makes the snake model the regions we want to disregard. Figure 7.1



(a) (b)

Figure 7.1: Synthetic object displaying several dynamics along the boundary. Contours obtained using (a) a tACM with rigidity of 0 and using (b) a tACM with rigidity of 10. $\mathbf{x}_0$ is represented in the left image as a black star and the contour evolves clockwise.

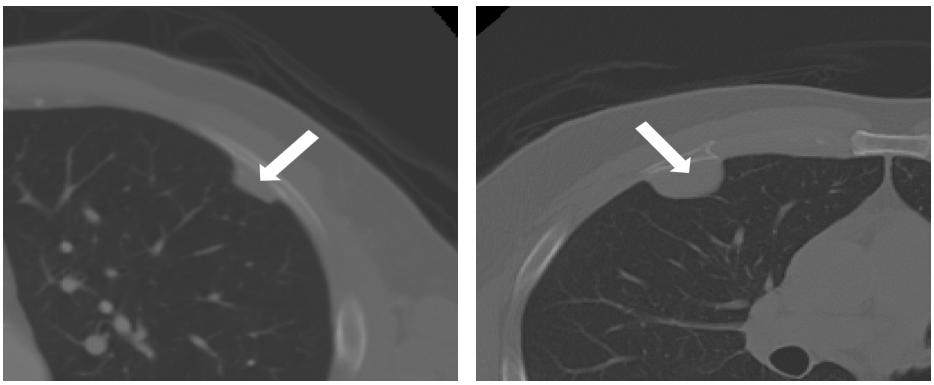


Figure 7.2: Examples of juxta-pleural nodules in CT images, indicated by white arrows.

presents the contours obtained with those models. An example of a real scenario where this issue arises is the segmentation of the lung in CT images. The inclusion of nodules on the boundary of the lung, which are commonly referred to as juxta-pleural nodules, is challenging given that they respond differently than healthy tissue to the imaging acquisition procedure (see Figure 7.2). This makes these nodules highly susceptible to be discarded in the lung segmentation step, putting at risk the nodule detection and characterisation steps that follow. There is strong evidence that tumour growth and metastasis relies on angiogenesis, the process where new blood vessels emerge from pre-existing ones [204; 205]. The analysis of blood vessel patterns near tumours, which are typically very distinct from the ones appearing in healthy tissues, is relevant for understanding the tumour mechanisms and therapeutic guidance [17; 18]. Hence, both nodule detection and local blood vessel analysis play an important role in these studies.

The problem of finding the true object boundary has been targeted in the past, however the methodologies are specifically tailored for a given application [206; 207] or they do not allow for different behavior along the contour [208; 209]. Recently, an orientation-lifted Finsler minimum

path approach has been developed and extended to closed contour detection [209], allowing to overcome the common shortcut problem by penalising the variation of the tangent vector along the curve. Even though such approach was effective at finding the true smooth boundary, it does not solve the problem mentioned before, as it is not capable of achieving heterogeneous behavior along the contour. In this Chapter, we discuss a novel parametric ACM that allows having a finite number of contiguous regions with different bending properties. We designate such method as Sparse Multi-Bending (SMB) snake, since rigidity transitions between consecutive contour points are sparse, leading to contiguous regions of equal bending resistance.

7.1 Parametric active contour models

Traditional snakes, as introduced in [109], are represented as a parametric curve $\mathbf{x}(s) = \big(x(s), y(s)\big)$, $s \in [0,1]$. The energy of a given snake configuration is obtained by evaluating an energy function along s:

$$\begin{aligned} E_{snake} &= \int_0^1 E_{snake}\big(\mathbf{x}(s)\big) \quad ds \\ &= \int_0^1 E_{int}\big(\mathbf{x}(s)\big) + E_{ext}\big(\mathbf{x}(s)\big) \quad ds \end{aligned} \tag{7.1}$$

where E_{int} is the energy due to topological constraints of the snake, and E_{ext} is an energy map where image features of interest, usually edges, have lower energy. Additionally, a term related to user input may also be considered, where attracting and/or repulsive forces are manually set.

The internal energy E_{int} is expressed as:

$$E_{int} = \frac{1}{2}\big(\alpha(s)|\mathbf{x}'(s)|^2 + \beta(s)|\mathbf{x}''(s)|^2\big) \tag{7.2}$$

where $\mathbf{x}'(s)$ and $\mathbf{x}''(s)$ are the first and second-order derivatives of $\mathbf{x}(s)$, respectively. The first term penalises the growth of the snake, making it behave as a membrane, while the second penalises bending, making it act like a thin plate. The coefficients $\alpha(s) \in \mathbb{R}_{\geq 0}$ and $\beta(s) \in \mathbb{R}_{\geq 0}$ control the relevance of those properties along the snake.

The external energy E_{ext} defines the features that attract the snake. Although different functions have been designed, in this work we consider one that associates lower energy to stronger edges. Then, the external energy may be given by:

$$E_{ext} = -\big|\nabla(G_\sigma * I)\big|^2 \tag{7.3}$$

where $G_\sigma * I$ is a smoothed version of image I obtained by performing a convolution with a Gaussian having a standard deviation of σ, and ∇ denotes the gradient operation.

A snake lying in a minimum of (7.1) must satisfy the Euler-Lagrange equation:

$$\alpha \mathbf{x}'' - \beta \mathbf{x}'''' - \nabla E_{ext}(\mathbf{x}) = 0 \tag{7.4}$$

that can be equivalently expressed as a force balance:

$$\mathbf{f}_{int}(\mathbf{x}) + \mathbf{f}_{ext}(\mathbf{x}) = 0 \tag{7.5}$$

where $\mathbf{f}_{int}(\mathbf{x}) = \alpha \mathbf{x}'' - \beta \mathbf{x}''''$ is the force due to the internal constraints, and $\mathbf{f}_{ext}(\mathbf{x}) = -\nabla E_{ext}(\mathbf{x})$ is the force that pulls the contour to the desired image features.

By treating $\mathbf{x}(s)$ as a function of time t, we can iteratively update the contour using a gradient descent scheme:

$$\frac{\partial \mathbf{x}(s,t)}{\partial t} = \alpha \mathbf{x}''(s,t) - \beta \mathbf{x}''''(s,t) + \mathbf{f}_{ext}(\mathbf{x}(s,t)) \tag{7.6}$$

The solution to (7.4) is found when (7.6) reaches a steady state. Numerical schemes perform a discretisation of $\mathbf{x}(s)$ into n contour points, $\mathbf{x}_i = \{x_i, y_i\}$, $i = 1, \ldots, n$, allowing to solve (7.6) in a discrete grid, such as a 2D image. The discrete version of (7.1) is given by:

$$E_{snake} = \sum_{i}^{n} \frac{1}{2} \left(\alpha_i |\mathbf{x}_i'|^2 + \beta_i |\mathbf{x}_i''|^2 \right) + E_{ext}(\mathbf{x}_i) \tag{7.7}$$

However, these ACMs suffer from strong dependency to the initialisation and are unable to model strong concavities. The GVF snake, as introduced in [201], is one of the fundamental methods dealing with these issues. It uses a different external force, represented by the vector field $\mathbf{v}(x,y) = [u(x,y), v(x,y)]$ that minimises the following function:

$$E = \int \int \mu \left(u_x^2 + u_y^2 + v_x^2 + v_y^2 \right) + \left| \nabla f \right|^2 \left| \mathbf{v} - \nabla f \right|^2 dxdy \tag{7.8}$$

where $f(x,y) = -E_{ext}$ and μ is a parameter controlling the smoothness of the GVF field. This function combines two interesting properties: (1) when $|\nabla f|$ is high, (7.8) is minimised by setting $\mathbf{v} = \nabla f$, such that the force field is kept similar to the gradient of the edge map at regions where strong edges exist; (2) when $|\nabla f|$ is low, the first term of (7.8) dominates and the energy is minimised when the partial derivatives are small, which means that the force field is smoothed at homogeneous regions. These properties allow extending the range of the force field and improving the convergence to concavities.

Even though (7.7) allows having a parameterised contour with variable stretching and bending resistances, for the best of our knowledge, no scheme for the automatic optimisation of such properties exists. tACMs set α_i and $\beta_i, i = 1, \ldots, n$ to constant values, such that identical properties exist across the entire contour. In Section 7.2, we present a novel snake model that automatically finds proper rigidity distribution, $\boldsymbol{\beta}$, along the contour, being able to accurately segment objects whose boundary displays dynamics that result from the presence of artefacts.

7.2 Sparse multi-bending snake

The SMB snake is a novel parametric ACM, since it is able to automatically tune the bending resistance along the contour. Our motivation is to have a flexible model that is able to accurately fit the true contour of objects that, due to noise or other artefacts, have a locally distorted boundary. A synthetic example was provided in Figure 7.1. This problem amounts to the necessity of having ACMs that are able to fit certain dynamics while neglecting others.

7.2.1 Energy definition

The desired properties for our ACM are two-fold. First, it should allow to accurately model highly dynamic regions. Second, it should be flexible enough in order to smooth regions with dynamical behavior in favor of fitting others that present higher dynamics. Here, we propose a novel energy function that allows achieving such behavior.

The energy of a parametric ACM having constant stretching and heterogeneous bending resistances is given by a particular case of (7.7):

$$E = \sum_{i}^{n} \frac{1}{2} \left(\alpha |\mathbf{x}_i'|^2 + \beta_i |\mathbf{x}_i''|^2 \right) + E_{ext}(\mathbf{x}_i) \tag{7.9}$$

Given that $\beta_i \in \mathbb{R}_{\geq 0}$ and $|\mathbf{x}_i''|^2 \in \mathbb{R}_{\geq 0}$, the minimisation of the energy (7.9) trivially induces configurations where $\beta_i = 0, i = 1, \ldots, n$. Therefore, constraints are required to produce more interesting models in terms of applicability.

In our proposed framework, we restrict the optimisation space by forcing the final solution to have an average rigidity value of $\overline{\beta}$. This may be thought as defining a budget that has to be distributed over the entire contour. In this particular case, the total budget is $n\overline{\beta}$, where n is the number of discrete points parametrising the contour. Even then, the existence of a point $\mathbf{x}_j$ with curvature $|\mathbf{x}_j''|^2 \approx 0$ would be enough to allocate the entire budget to β_j. Thus, a second constraint sets an upper bound M on the budget that might be given to a single point.

These constraints already allow obtaining non-trivial models whose rigidity is heterogeneously distributed along the contour, in such a way that low rigidity exists in regions of higher dynamics. However, noise and other artefacts may lead to capturing dynamical behaviours that we do not want to model. Thus, we seek solutions with sparse rigidity variation that divide the contour into a small number of regions having heterogeneous bending properties, in order to fit the true dynamics of a contour while ignoring the contribution of noise and other artefacts. We address this issue by modifying the energy function to promote sparse rigidity transitions along the contour, for the sake of inducing contiguous regions with different bending properties. Such goal is achieved by considering a L_0 norm term that penalises the number of transitions in $\boldsymbol{\beta}$, leading to our proposed energy function:

$$E_{SMB} = \sum_{i}^{n} \frac{1}{2} \left(\alpha |\mathbf{x}_i'|^2 + \beta_i |\mathbf{x}_i''|^2 + \lambda \left\| \beta_i - \beta_{i-1} \right\|_0 \right) + E_{ext}(\mathbf{x}_i) \tag{7.10}$$

where $\beta_i \in [0, M]$, $\sum_i \beta_i = n\overline{\beta}$, $\lambda \in \mathbb{R}_{\geq 0}$ is a regularisation constant that controls the sparsity inducing term, and $\|\cdot\|_0$ denotes the L_0 norm, which in this particular setting is 1 every time that the rigidity parameter β changes between consecutive points in the contour, and 0 otherwise. Note that we are considering closed parametric contours, such that $\beta_0 \equiv \beta_n$. This holds for the remaining expressions presented in this Chapter, although we omit it for simplicity of notation.

The term of (7.10) accounting for the curvature is responsible for pulling lower rigidity to points with higher curvature, while the L_0 norm imposes that the solutions are sparse, that is, a limited amount of different bending regions is allowed to exist. In the end, an adequate choice of λ allows obtaining a sparse solution where the most bending-resisting segments are those where the frequency of high curvatures is lower. When $\lambda \to \infty$, the energy function will not allow any transition to occur in $\boldsymbol{\beta}$, such that the minimum of (7.10) amounts to setting $\beta_i = \overline{\beta}$, $i = 1, \ldots, n$. Then, the SMB snake would behave as a tACM with $\beta = \overline{\beta}$. The other extreme case, when $\lambda = 0$, is not interesting in terms of applicability. Such setting freely introduces transitions in $\boldsymbol{\beta}$, such as to minimise the curvature related term of (7.10), only restricted by the constraints. The ratio $M/\overline{\beta}$ controls the proportion of contour points that are allowed to have low β. It is expected that for high ratios, a large portion of points have very low β at the expense of assigning high β to a point where the curvature is low. This particular case produces final contours very similar to the ones obtained when using a tACM with low β. A more interesting scenario arises when $\lambda \in]0, \infty[$. An adequate parametrisation allows modelling the curvature of some regions while smoothing the fluctuations of others. In addition, the parametrisation of the SMB snake naturally allows for a significant flexibility in terms of the number of different bending regions and their respective sizes. The joint optimisation problem considered in the proposed SMB snake comes as follows:

$$\begin{aligned} \underset{\mathbf{x}_i, \beta_i}{\arg\min} \quad & \sum_{i=1}^{n} \frac{1}{2}\left(\alpha|\mathbf{x}_i'|^2 + \beta_i|\mathbf{x}_i''|^2 + \lambda\left\|\beta_i - \beta_{i-1}\right\|_0\right) + E_{ext}(\mathbf{x}_i) \\ \text{subject to} \quad & 0 \leq \beta_i \leq M, \quad i = 1, \ldots, n, \\ & \frac{1}{n}\sum_i \beta_i = \overline{\beta} \end{aligned} \tag{7.11}$$

7.2.2 Optimisation framework

The joint optimisation of the contour $\mathbf{x}$ and the rigidity parameters $\boldsymbol{\beta}$, as expressed in (7.11), is a complex problem. Nonetheless, we propose to tackle it by considering alternate steps of partial optimisation: optimisation of $\mathbf{x}$ given rigidity parameters $\boldsymbol{\beta}$, and optimisation of $\boldsymbol{\beta}$ given contour $\mathbf{x}$, until convergence. The first step amounts to evolve $\mathbf{x}$ for a given $\boldsymbol{\beta}$, using the traditional discrete schemes. The second one is now addressed, as we present a method to iteratively optimise $\boldsymbol{\beta}$ given $\mathbf{x}$.

For simplicity of notation, let $\boldsymbol{k}$ be a vector representing the curvature along the parametrised contour $\mathbf{x}$, such that the curvature at $\mathbf{x}_i$, given by $|\mathbf{x}_i''|^2$, is denoted by k_i. The optimisation of $\boldsymbol{\beta}$

given $\mathbf{x}$ is deduced from (7.11) by only taking into account the terms that depend on the rigidity distribution:

$$\begin{aligned} \underset{\beta_i}{\arg\min} \quad & \sum_{i=1}^{n} \beta_i k_i + \lambda \left\| \beta_i - \beta_{i-1} \right\|_0 \\ \text{subject to} \quad & 0 \leq \beta_i \leq M, \quad i = 1, \ldots, n, \\ & \frac{1}{n} \sum_i \beta_i = \overline{\beta} \end{aligned} \tag{7.12}$$

Finding the solution to (7.12) is not trivial due to the inclusion of the L_0 norm. It makes the energy function non-differentiable, such that typical gradient-based descent methods cannot be used. Here, we propose a group optimisation strategy based on pairwise coordinate descent to minimise (7.12).

Consider any division of the contour $\mathbf{x}$ into two contiguous regions, R_1 and R_2, such that $R_1 \cap R_2 = \emptyset$ and $R_1 \cup R_2 = \mathbf{x}$. An example of such a division is provided in Figure 7.3. By (7.12),

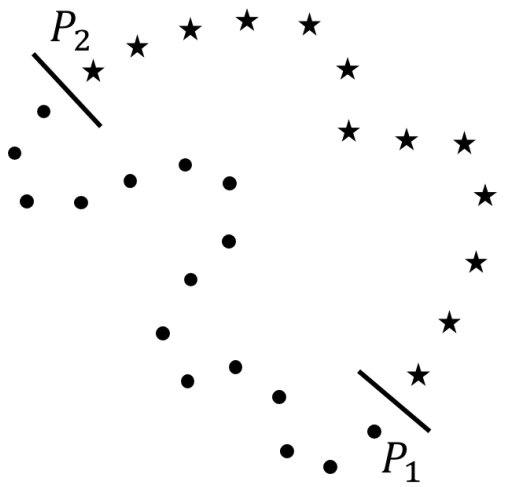


Figure 7.3: Possible division of a snake into two different regions, R_1 and R_2. P_1 and P_2 are the indices of the first contour points of regions R_1 and R_2, respectively, considering clockwise order.

we know that $\sum_i \beta_i = n\overline{\beta}$. Thus, the following trivially holds:

$$\gamma - \gamma + \sum_{i \in R_1} \beta_i + \sum_{i \in R_2} \beta_i = n\overline{\beta} \tag{7.13}$$

which is equivalent to:

$$\sum_{i \in R_1} \left(\beta_i + \frac{\gamma}{n_1} \right) + \sum_{i \in R_2} \left(\beta_i - \frac{\gamma}{n_2} \right) = n\overline{\beta} \tag{7.14}$$

where γ is a variable representing a perturbation to the rigidity distribution, and n_1 and n_2 are the number of points in R_1 and R_2, respectively. Given a feasible $\boldsymbol{\beta}$ distribution, and a particular

division of the contour into two regions, our group optimisation strategy updates the rigidity β_i of each contour point $\mathbf{x}_i$ by δ_i, according to the region it belongs to:

$$\delta_i = \begin{cases} \gamma/n_1 & \text{if } i \in R_1 \\ -\gamma/n_2 & \text{if } i \in R_2 \end{cases}, \; i = 1, \ldots, n \tag{7.15}$$

Then, for a given division of the contour, we are interested in finding the value of γ that leads to a minimisation of (7.12). Such is achieved by transforming (7.12) into an expression that depends on γ:

$$\begin{aligned} \arg\min_{\gamma} \quad & \sum_{i=1}^{n} \Big(\beta_i + \delta_i\Big) k_i + \lambda \Big\| \beta_i + \delta_i - \beta_{i-1} - \delta_{i-1} \Big\|_0 \\ \text{subject to} \quad & 0 \leq \beta_i + \delta_i \leq M, \quad i = 1, \ldots, n \end{aligned} \tag{7.16}$$

By keeping only the terms dependent on γ, (7.16) is simplified to:

$$\begin{aligned} \arg\min_{\gamma} \quad & \gamma \left(\frac{1}{n_1} \sum_{i \in R_1} k_i - \frac{1}{n_2} \sum_{i \in R_2} k_i \right) + \\ & \lambda \left(\left\| \beta_{P_1} + \frac{\gamma}{n_1} - \beta_{P_1 - 1} + \frac{\gamma}{n_2} \right\|_0 + \left\| \beta_{P_2} - \frac{\gamma}{n_2} - \beta_{P_2 - 1} - \frac{\gamma}{n_1} \right\|_0 \right) \\ \text{subject to} \quad & 0 \leq \beta_i + \frac{\gamma}{n_1} \leq M, \quad i \in R_1, \\ & 0 \leq \beta_i - \frac{\gamma}{n_2} \leq M, \quad i \in R_2 \end{aligned} \tag{7.17}$$

where P_1 and P_2 are the first contour points of R_1 and R_2, respectively, considering clockwise order.

Let $\beta^m_{i \in R_j}$ and $\beta^M_{i \in R_j}$ denote, respectively, the minimum and maximum β values found among the points of region R_j. The constraints in (7.17) can be equivalently written as the following set of conditions:

$$\left\{ \gamma \geq \max\Big(-n_1 \beta^m_{i \in R_1}, n_2 \big(\beta^M_{i \in R_2} - M\big)\Big), \;\; \gamma \leq \min\Big(n_1 \big(M - \beta^M_{i \in R_1}\big), n_2 \beta^m_{i \in R_2}\Big) \right\} \tag{7.18}$$

The set of conditions in (7.18) defines a minimum and maximum limit for γ, which we designate as γ^m and γ^M, respectively. A further reduction of the search space can be achieved, by carefully analysing the shape of the terms in the goal function of (7.17). The first term is linear on γ, being the slope dictated by the relation between the mean region curvatures, $\frac{1}{n_1}\sum_{i \in R_1} k_i$ and $\frac{1}{n_2}\sum_{i \in R_2} k_i$. The contribution of this term to the energy (7.17), as a function of γ, is presented in Figure 7.4. When the slope is positive, this term pushes the optimum γ towards γ^m. Otherwise, when the slope is negative, it drives the optimum γ to γ^M. Note that, for the particular case where the regions have equal mean curvature, the slope is 0, such that any γ leads to the same cost. This means that this term either contributes to one minimum, γ^m or γ^M, or to none. The second term is a sum of two L_0 norms weighted by the regularisation term λ. A L_0 norm contributes in a very particular manner to the energy (see Figure 7.5). It generates a local minimum at the

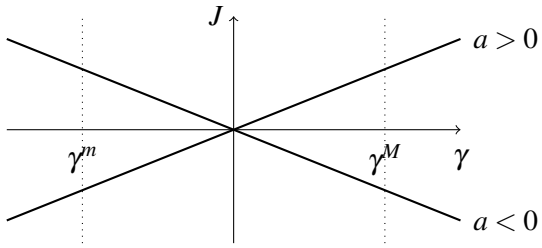


Figure 7.4: Cost J of the first term of (7.17), as a function of γ. a is the difference between the mean curvatures of R_1 and R_2.

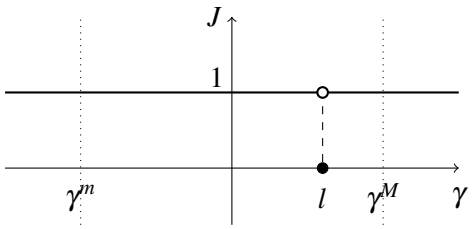


Figure 7.5: Cost J of a single L_0 norm of the second term of (7.17), as a function of γ. l is given by $\frac{n_1 n_2}{n_1+n_2}(\beta_{P_1-1}-\beta_{P_1})$ for the first L_0 norm and $\frac{n_1 n_2}{n_1+n_2}(\beta_{P_2}-\beta_{P_2-1})$ for the second one.

value of γ that reduces its argument to 0. Hence, the first L_0 norm induces a local minimum at $\gamma = \frac{n_1 n_2}{n_1+n_2}(\beta_{P_1-1}-\beta_{P_1})$, while the second is responsible for one at $\gamma = \frac{n_1 n_2}{n_1+n_2}(\beta_{P_2}-\beta_{P_2-1})$.

Given that (7.17) is simply given by the addition of these functions, it amounts to find which of the following values of γ minimises (7.17):

$$\left\{ \gamma^m,\ \gamma^M,\ \frac{n_1 n_2}{n_1+n_2}(\beta_{P_1-1}-\beta_{P_1}),\ \frac{n_1 n_2}{n_1+n_2}(\beta_{P_2}-\beta_{P_2-1}) \right\} \tag{7.19}$$

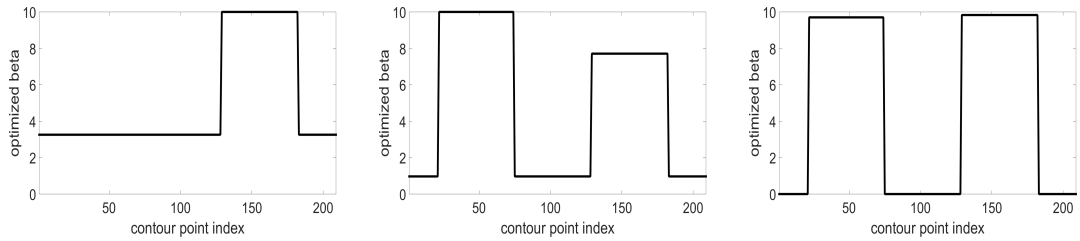
After finding the optimum γ for a given division, $\boldsymbol{\beta}$ may be updated accordingly:

$$\beta_i(\tau+1) = \begin{cases} \beta_i(\tau)+\frac{\gamma}{n_1} & \text{if } i \in R_1 \\ \beta_i(\tau)-\frac{\gamma}{n_2} & \text{if } i \in R_2 \end{cases}, \ i = 1,\ldots,n \tag{7.20}$$

where $\tau \in \mathbb{Z}^+$ denotes the current iteration of the optimisation procedure. Essentially, we propose an iterative procedure (see Algorithm 2), where, at iteration τ, we consider all the possible divisions of the contour into two contiguous regions, and for each, we find the optimum γ. The division that is effectively considered at iteration τ is the one that minimises (7.16). We repeat such process until convergence or, in other words, until no additional perturbation further decreases our energy function. Note that, even though we only consider a division into two regions at each time, the iterative nature of the optimisation procedure allows to incrementally find new heterogeneous bending regions. Figure 7.6 exemplifies this, by showing how the rigidity evolves with τ, when $\mathbf{x}$ is the contour shown in the left image of Figure 7.1, $\lambda = 5$, $\overline{\beta} = 5$, and $M = 10$.

Algorithm 2: Optimisation of rigidity distribution $\boldsymbol{\beta}$ given contour $\mathbf{x}$.

input : rigidity distribution $\boldsymbol{\beta}$, contour $\mathbf{x}$
output: optimised rigidity distribution $\widehat{\boldsymbol{\beta}}$
$\boldsymbol{\beta}^{it} \leftarrow \boldsymbol{\beta}$
$\mathbf{k} \leftarrow$ curvature along $\mathbf{x}$
$J \leftarrow \text{LossFunction}(\boldsymbol{\beta}, \mathbf{k})$ // (7.12)
do
  $\widehat{\boldsymbol{\beta}} \leftarrow \boldsymbol{\beta}^{it}$
  for $(P_1, P_2),\ 1 \leq P_1 < P_2 \leq n$ **do**
    $\gamma \leftarrow \text{GroupOptimisation}(\widehat{\boldsymbol{\beta}}, \mathbf{k}, P_1, P_2)$ // (7.16)
    $\boldsymbol{\beta}^{config} \leftarrow \text{UpdateBeta}\ (\widehat{\boldsymbol{\beta}}, \mathbf{k})$ // (7.20)
    $cost \leftarrow \text{LossFunction}(\boldsymbol{\beta}^{config}, \mathbf{k})$ // (7.12)
    if $cost < J$ **then**
      $\boldsymbol{\beta}^{it} = \boldsymbol{\beta}^{config}$
      $J \leftarrow cost$
    end
  end
while $\boldsymbol{\beta}^{it} \neq \widehat{\boldsymbol{\beta}}$;
return $\widehat{\boldsymbol{\beta}}$



(a) (b) (c)

Figure 7.6: Rigidity distribution evolution with the number of iterations of the proposed optimisation algorithm, when applied to the contour shown in Figure 7.1a, with $\lambda = 5$, $\overline{\beta} = 5$, and $M = 10$: (a) after 1 iteration; (b) 2 iterations; and (c) convergence.

7.3 Experimental Results

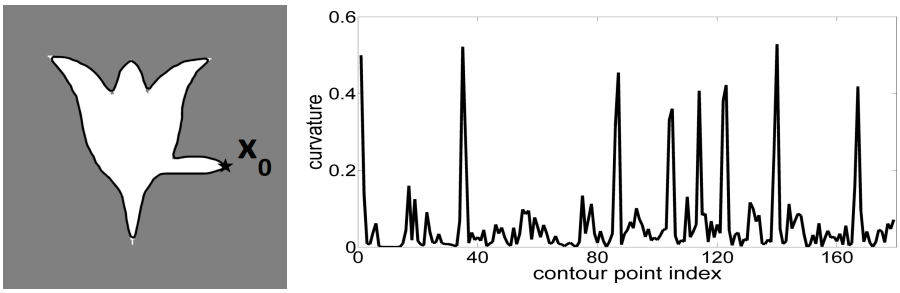
In this Section, we demonstrate how the SMB snake design allows to address different scenarios. First, we perform experiments on synthetic images just to show the adaptability of the model to different user requirements. Afterwards, we apply the SMB snake to a real application that benefits from having a multi-bending ACM, the lung segmentation in CT images. Finally, we report some findings related to initialisation and computational efficiency. A Matlab implementation of our algorithm is available at `https://github.com/rjtaraujo/smb-snake`. Additional experiments regarding the delineation of hands in hand gesture images can be found in the published paper.

In the experiments described in subsections 7.3.1 and 7.3.2, we initialised the contour $\mathbf{x}(t_0)$ at

the boundary of a dilated version of the ground truth of the object. A square kernel of side 13 was used. The number of discrete points was set to a fraction of the perimeter of that initial contour. The fraction was $1/3$ for synthetic images, and $1/5$ for the lung ones, as the lung boundary can be accurately modelled with fewer contour points. For all tested parametric models, including ours, we set $\alpha = 0$, such that we do not penalise stretching forces. Additionally, the image related forces were obtained after 100 iterations of GVF (7.8). Regarding our model, the rigidity across $\mathbf{x}(t_0)$ was set to $\beta(t_0), \forall i$. After letting the snake converge, at time t_1, we optimised $\boldsymbol{\beta}$ given the curvature $\boldsymbol{k}$ along the contour $\mathbf{x}(t_1)$. The curvature at point $\mathbf{x}_i$ was estimated by the inverse of the radius of the circle fitted using $\mathbf{x}_i$ and its neighbours $\mathbf{x}_{i-1}$ and $\mathbf{x}_{i+1}$. Finally, we let the snake converge again, according to the new rigidity distribution, leading to the final contour $\mathbf{x}(t_2)$. In our experiments, no further improvement was obtained by undergoing additional steps of rigidity distribution optimisation.

7.3.1 Synthetic images

Let us consider first a binary object resembling the shape of a flower, as shown in Figure 7.7. Three regions induce high curvature along the boundary of the object, namely the petal (top), sepal

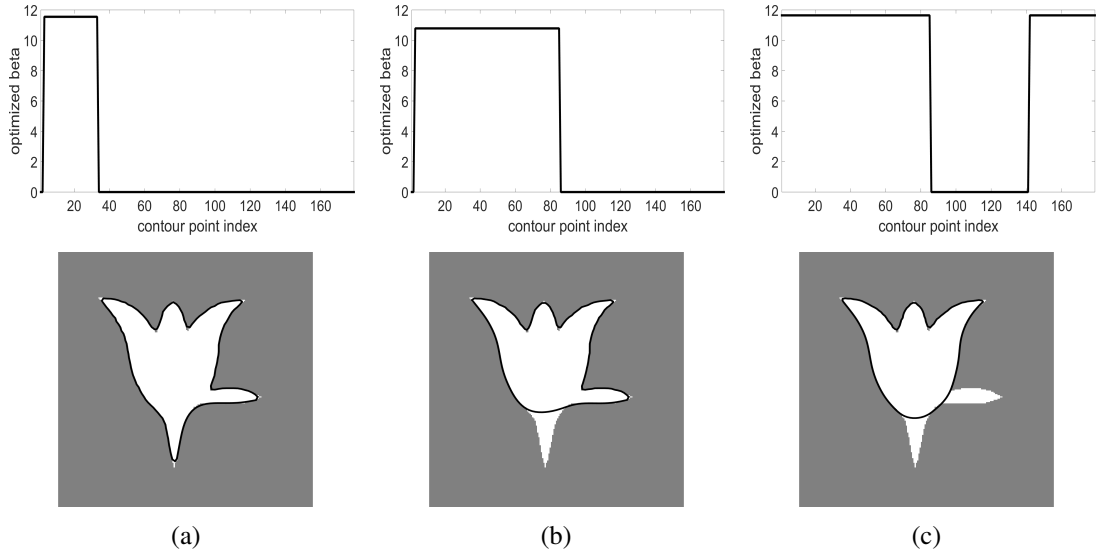


(a) (b)

Figure 7.7: Synthetic image and curvature along the contour: (a) image and contour $\mathbf{x}(t_1)$ in black, with $\mathbf{x}_0$ represented as a black star and considering that the contour evolves clockwise; and (b) curvature along the contour.

(right) and stem (bottom) regions. The SMB snake parametrisation allows modelling portions of the mentioned dynamics. For instance, by setting $\beta(t_0) = 0$, we first obtain a detailed contour of the flower (black line imposed in Figure 7.7a). The curvature along the contour is represented in Figure 7.7b.

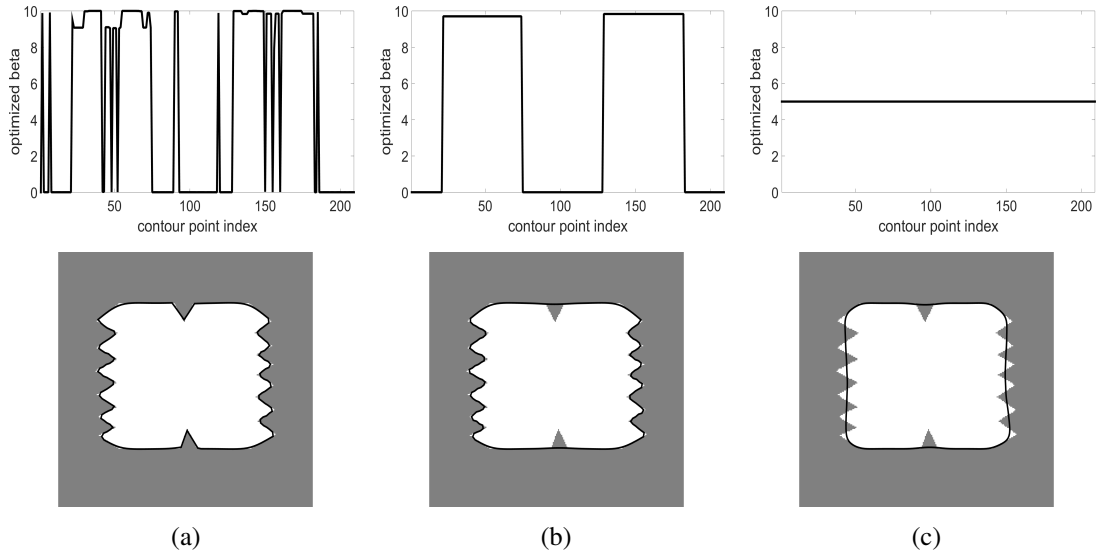
As can be seen, the stronger dynamic appears in the petal region, such that this is the first region that the SMB snake tries to distinguish. However, an adequate parametrisation may also append adjacent dynamics, such as the sepal one. Figure 7.8 illustrates the flexibility of the proposed model, where we set $\beta(t_0) = 0$, $\lambda = 5$, $M = 12$ and show the effect of varying $\overline{\beta}$. The results show that by decreasing the ratio $M/\overline{\beta}$, we induce configurations with a smaller portion of the contour having low rigidity, as discussed in Section 7.2.



(a) (b) (c)

Figure 7.8: $\boldsymbol{\beta}$ distribution (top) and SMB snake result (bottom) for $\beta(t_0) = 0$, $\lambda = 5$, $M = 12$ and varying $\overline{\beta}$: (a) 2, (b) 5, and (c) 8.

It may also be of interest to use parametrisations that induce a partition of the contour into more than two different bending resisting regions. Returning to the example provided in Figure 7.1, the objective was to fit the lateral dynamics while neglecting the remaining ones. Figure 7.9 shows how the optimised rigidity distribution varies with λ and demonstrates that the SMB snake is able to accomplish its purpose for an adequate regularisation value, which in this case was any λ between 1.4 and 10.3.



(a) (b) (c)

Figure 7.9: $\boldsymbol{\beta}$ distribution (top) and SMB snake result (bottom) for $\beta(t_0) = 0$, $\overline{\beta} = 5$, $M = 10$ and varying λ: (a) 0, (b) 5, and (c) 20.

7.3.2 Lung CT images

The SMB snake is able to naturally handle the challenge presented in the beginning of this Chapter, concerning the inclusion of juxta-pleural nodules during lung segmentation. Throughout the CT images, the lungs frequently exhibit a medial region with intermediate to high dynamics, while the lateral region is generally smooth. Hence, a SMB snake capable of modeling the lung contour as two different segments seems ideal to include juxta-pleural nodules that exist in the lateral region of the lung.

To conduct our experiments, we used 406 lung CT images from the LIDC-IDRI database [210; 211; 212], of which 208 include juxta-pleural nodules. Manual annotations of the entire lung regions and of the individual juxta-pleural nodules were made by an expert. We compare the results of our model to the GVF snake [201], and two implicit ACMs, the Distance Regularised Level Set Evolution (DRLSE) method [113], and the Selective Binary and Gaussian Filtering Regularised Level Set (SBGFRLS) [213]. The edge-based application of DRLSE is ruled by the following gradient flow:

$$\frac{\partial \phi}{\partial t} = aR(\phi) + bL(\phi) + cA(\phi) \tag{7.21}$$

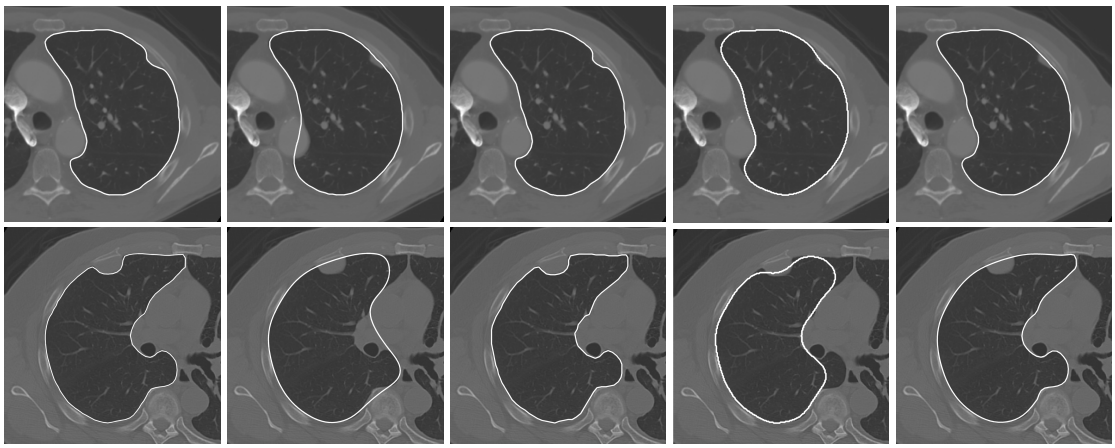
where ϕ is a level set function that represents the contour implicitly as the zero-level set, $R(\phi)$ is associated with a term that penalises the difference of ϕ to a signed distance function, $L(\phi)$ is related to the line integral of the edge indicator function along the zero level set of ϕ, $A(\phi)$ to the weighted area of the region inside zero level set of ϕ, and $a \in \mathbb{R}_{>0}$, $b \in \mathbb{R}_{>0}$, $c \in \mathbb{R}$ control the relevance of those terms, respectively. The sign of c dictates whether the zero level set expands or contracts.

On the other hand, the region-based SBGFRLS evolution is given by:

$$\frac{\partial \phi}{\partial t} = spf(I) \cdot d|\nabla \phi| \tag{7.22}$$

where $spf(I)$ is a region-based signed pressure function, making the contour shrink when outside the region of interest and expand when inside. The second term, modulated by a constant d, increases the propagation speed. Regarding the implementation of this level set formulation, after each evolution step, a Gaussian filter is used to regularise the level set function, such that its standard deviation σ dictates the amount of regularisation of the contour.

Regarding the GVF snake, we tested different values of β. As for the parametrisation of the DRLSE model, we set $a = 0.2$, $c = 3$ and tested different configurations of b. With respect to the SBGFRLS model, we set $d = 1$ and tuned σ. Finally, for our model, we empirically set $\beta(t_0) = 5$, $\lambda = 1$, $M = 17$, and varied the value of $\overline{\beta}$. Table 7.1 summarises the performance of these models with respect to three different metrics: the percentage of juxta-pleural nodule area that is included in the segmentations, and the Jaccard index and F1-score accounting for the entire lung area segmentation. Visual results related to the examples provided in Figure 7.2 are presented in Figure 7.10.



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

Figure 7.10: Example contours of the lung area, obtained using (a) a GVF snake with $\beta = 0$; (b) a GVF snake with $\beta = 20$; (c) the DRLSE model with $b = 50$; (d) the SBGFRLS model with $\sigma = 10$; and (e) the proposed SMB snake with $\overline{\beta} = 10$.

Table 7.1: Evaluation of the lung segmentations obtained with different ACMs. Comparison is made in terms of the mean juxta-pleural nodule area that is included and also in terms of the complete lung area, using the Jaccard index and the F1-score.

Model		**% nodule area segmentation**	**J**	**F1**
GVF	$\beta = 0$	24.1	0.963	0.981
	$\beta = 1$	60.7	0.958	0.979
	$\beta = 5$	76.6	0.944	0.971
	$\beta = 10$	81.1	0.913	0.964
	$\beta = 20$	83.8	0.908	0.951
DRLSE	$b = 10$	22.6	0.970	0.985
	$b = 50$	26.8	0.973	0.986
	$b = 100$	26.7	0.971	0.985
SBGFRLS	$\sigma = 1$	11.9	0.968	0.983
	$\sigma = 5$	62.4	0.957	0.978
	$\sigma = 10$	71.0	0.924	0.959
Proposed	$\overline{\beta} = 8$	67.2	0.958	0.978
	$\overline{\beta} = 10$	80.0	0.956	0.977
	$\overline{\beta} = 12$	81.4	0.951	0.975

The results show that, as we increase β in the GVF snake, we are able to include a higher percentage of nodule area, at the cost of decreasing the accuracy at other dynamic regions, as concluded from the decrease in the Jaccard index and F1-score. This was expected since more topological restricted snakes induce smoother contours, allowing them to ignore the concavities that exist due to the juxta-pleural nodules. However, they also fail at returning a detailed lung

boundary in highly curved regions. The DRLSE method produced results similar to a GVF snake with zero rigidity, and varying the parametrisation did not significantly change the behaviour of the model. Regarding the SBGFRLS model, the increase of regularisation allowed including more nodule area, as expected; however, the loss of lung boundary detail also increased significantly. The SMB snake was able to achieve the advantages of both low and high bending resisting tACMs, combining a proper inclusion of juxta-pleural nodules and accurate lung boundary modelling. It proved to be the model achieving the best compromise. Even then, occasionally, juxta-pleural nodules may appear in the medial region of the lung or in its vicinity, and be wrongly interpreted as part of the natural dynamic of the lung (see Figure 7.11).

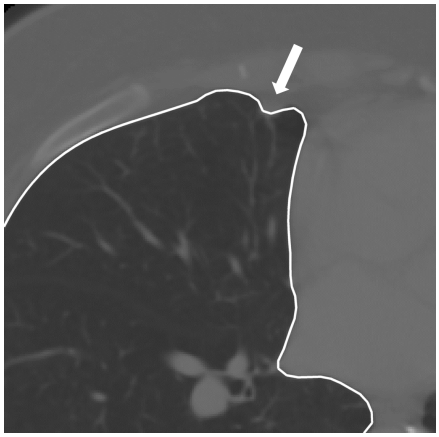


Figure 7.11: An example juxta-pleural nodule that has been missed by our proposed framework due to its inclusion in the region of higher dynamics.

7.3.3 Impact of intensity inhomogeneity

Intensity inhomogeneity, either due to non-uniform illumination or the nature of the image, introduces a challenge to the evolution of ACMs, as it leads to significant changes of the gradient and region statistics. Here, taking the image in Figure 7.1, we artificially generate images that are affected by intensity inhomogeneity. We analysed how the proposed methodology behaved in such scenario, in comparison with the edge-based DRLSE method described in section 7.3.2 and a region-based level set formulation that is robust in the presence of intensity inhomogeneity, the Locally Statistical Active Contour Model (LSACM) [214]. Visual results[1] are provided in Figure 7.12. Regarding the proposed model, the number of discrete points was set to half the perimeter of the initial contour, $\beta(t_0) = 0$ (such that a GVF snake with no rigidity first fits all of the object dynamics), $\lambda = 10$, $\overline{\beta} = 10$, and $M = 25$. As for the DRLSE, $a = 1$, $b = 10$, and $c = 4$ were used. Finally, concerning the LSACM, we used the implementation provided in [215] and set the parameter controlling the size of the constant kernel to 20.

[1] In our experiments concerning intensity inhomogeneity, the SBGFRLS was clearly worst than the LSACM; thus, we only show results for the latter. Despite this, the SBGFRLS achieved better compromises regarding segmented nodule area and lung boundary detail in the lung experiment; thus, we selected it as representative of region-based level set methods in subsection 7.3.2.

The GVF snake was capable of converging to the dynamics of the object in the first three scenarios of intensity inhomogeneity, including the case where the shadow of the object was present (third row of Figure 7.12). This is a result of smoothing and propagating the most relevant gradient vectors, which exist between the object and the background. The DRLSE method failed in this particular case, as the edge between the shadow and the background was strong enough to prevent further evolution of the contour. However, in the fourth case, where we mimic object overlapping, the gradient vectors between the overlapped object and the background are sufficiently strong to prevent the GVF snake from converging only to the dynamics of the target object. Our methodology, by taking the result of the GVF snake and according to its parametrisation, was capable of breaking the contour into a set of contiguous regions with different bending properties, as shown before. The LSACM successfully fitted the dynamics of the object in all the images, even though it does not possess the flexibility of the SMB snake and cannot replicate its results when desired.

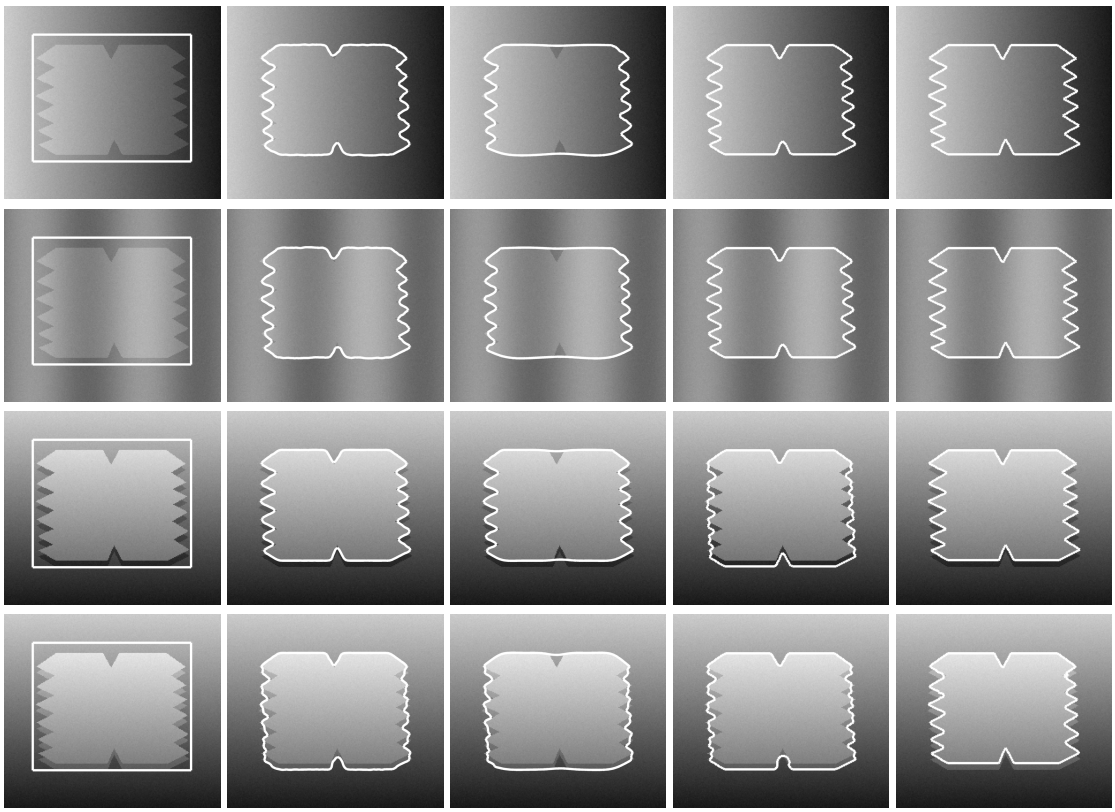


Figure 7.12: Impact of intensity inhomogeneity in the evolution of different ACMs: (a) initial contours; the results obtained using (b) the GVF snake, (c) our proposed methodology, (d) the DRLSE level set, and (e) the LSACM.

7.3.4 Impact of contour initialisation

The evolution of parametric ACMs is largely influenced by the initialisation of the curve, especially when local minima exist due to noise and other objects. This is the case of the lung scenario we described before, which justifies that we analyse here the effect of varying the initialisation in the performance of some of the parametric ACMs that were considered for that application: GVF snake with $\beta = 0$, with $\beta = 20$, and the SMB snake with $\overline{\beta} = 10$. We now consider initialising with dilations of the ground truth using kernels of varying width, more precisely, 0 (no dilation), 13, 26, and 39. We also consider the same dilations affected by local distortions obtained by adding noise drawn from a normal distribution $\mathcal{N}(\mu = 0, \sigma^2 = 100)$. Finally, we also test evolving the ACMs from an ellipse fitted to the object, and from the bounding box of the object. Let these scenarios be represented by D_0, D_{13}, D_{26}, D_{39}, DN_0, DN_{13}, DN_{26}, DN_{39}, ELL, and $BBOX$, respectively. Figure 7.13 presents a graphical comparison of how the performance of the ACMs was affected by the initialisation, with respect to the overall lung segmentation quality, according to the Jaccard index.

The performance of the considered ACMs decreased when increasing the distance between the initial curve and the target lung boundary. This was expected due to the influence of the non-target lung in the evolution of the contour. The added noise did not have such a significant impact, but it ended up affecting slightly more the GVF with zero rigidity. As that model does not penalise bending, any distortion of the contour that makes it approximate the boundary of the non-target lung results in a region that becomes trapped. Again, initialising at the bounding box of the target lung frequently induced the contour to be trapped at the boundary of the non-target lung, leading to poor performance. The initialisation as an ellipse fitted to the target lung did not severely suffer from this since it better encapsulated the object of interest. The initialisation dependence of our model is related to the parametric ACM used to first fit the dynamics of the object (t_1), before optimising the rigidity along the contour. To demonstrate the properties of our framework, we used the GVF snake with rigidity $\beta(t_0)$ in that first step. A more sophisticated parametric ACM, in the sense of being robust to noise and initialisation, could be used instead; however, such detail is behind the scope of this work.

7.3.5 Time efficiency

The SMB snake includes an intermediate step where a new rigidity distribution is found, therefore it has an increased computational cost associated. Even though we introduced an exhaustive search algorithm for optimising the rigidity distribution, note that for each possible division of the contour into two, there is an analytical solution, only requiring the evaluation of the cost function at four points. In fact, the optimisation procedure only took, on average, 0.014 s for each considered image of the lung database, using a naive implementation in C++ and an Intel Core i7-6700 CPU 3.40GHz in a setup with 16.0 GB of RAM. For the sake of comparison, 100 iterations of tACM evolution, which are not enough for convergence in most applications, take 0.5 s in a Matlab implementation [216]. Thus, our rigidity optimisation procedure takes less time than 4 iterations

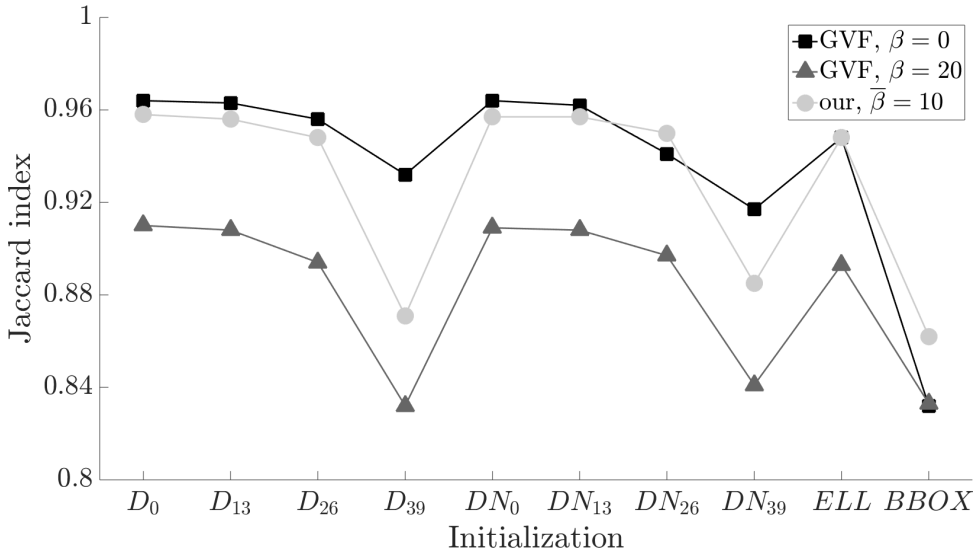


Figure 7.13: Performance of parametric ACMs when only considering the segmentation of the total area of the target lung, according to curve initialisation.

of tACM evolution, showing that its computational cost is practically negligible, especially when we also take into consideration the cost of calculating the image energy and the field of forces. We believe that most applications benefiting from our ACM fall into the scenarios included in this Chapter, in the sense that only a small number of regions should have distinct bending properties, thus requiring few iterations of the optimisation algorithm until convergence.

7.4 Main contributions and final considerations

This Chapter addressed the challenge of including juxta-pleural nodules when using automated algorithms for the segmentation of the lung area. This task is typically performed before lung nodule detection and analysis, such that it is very important to not discard any nodule during the segmentation step. The analysis of nodules commonly takes into account features such as their size and shape, nonetheless it may also involve assessing local blood vessel patterns for a more complete study, and guiding and evaluating therapeutic treatments [17; 18].

A novel parametric ACM that allows to divide the snake into a set of contiguous regions with different rigidity properties was presented. The proposed energy function allows the user to control the amount of such regions, since it incorporates a regularisation term that penalises the number of transitions in the $\boldsymbol{\beta}$ distribution along the contour. A group optimisation strategy was also presented, which may be used to optimise $\boldsymbol{\beta}$, given a contour with points $\mathbf{x}_i, i = 1, \ldots, n$.

In addition to experiments with synthetic images, we tested the proposed model in the real application of lung area segmentation in CT images. We showed how our model achieved a result that is not possible when using other explicit and implicit ACMs. The SMB snake was able to accurately follow the lung curvature while including most of the juxta-pleural nodules. Other ACMs can only achieve one of those desired properties. Parametrisations of low rigidity tend to

not include the nodule area, while the ones with high rigidity fail at retrieving the detailed lung boundary.

Our main contributions to the ACM literature are the following: (1) proposal of a snake model that is able to automatically fit different dynamics along the boundary of the object; (2) design of a novel energy function that induces few transitions in the rigidity coefficients along the contour, and consequently the existence of a small number of contiguous regions; (3) derivation of a pairwise coordinate descent strategy to optimise the proposed energy function.

Regarding future work, given the relevance of level set based approaches in the current framework of ACMs, it would be interesting to have an implicit ACM, which naturally handles topology changes and is more robust to initialisation, benefiting from the flexibility that characterises the SMB snake. Additionally, an extension to 3D would also be highly desirable.

Chapter 8

Conclusions and Future work

Blood vessels are the structures responsible for carrying blood throughout the different tissues of the body, enabling the exchange of molecules between this fluid and cells. The imaging and analysis of blood vessels is required by clinicians in several scenarios: suspicion of abnormalities in their structure, such as aneurysms and stenosis, both of which may lead to life-threatening events; to obtain cues on some systemic diseases, such as hypertension and diabetes; to better characterise certain tumours; and to determine surgery eligibility and perform pre-operative planning. Yet, blood vessel analysis is a very time consuming and repetitive task, an issue that becomes even more apparent in screening programmes, where a large volume of data is collected in a short span of time. Not surprisingly, the task of blood vessel segmentation became a hot topic of research in the computer vision community, with the goal of identifying the data pixels belonging to blood vessels. This map may facilitate the analysis by clinicians or be fed as input to an automated analysis and characterisation module. Nowadays, the blood vessel segmentation literature is already vast and, even though great progress has been achieved, there are still challenges that require attention. This thesis discusses some of the areas where we believe the literature is lacking and presents the outcomes of our research.

The evolution of medicine and clinical practice pushes forward the field of computer vision, since many new applications, challenges and needs arise. An area where medicine has been evolving significantly in the last years is autologous-based breast reconstruction. Through microsurgery techniques, it became possible to extract tissue from the belly without significantly disturbing the abdominal wall. Even then, this procedure requires a careful pre-operative planning, which is dependent on the analysis of the DIEPs, the blood vessels vascularising the anterior abdominal wall, by a radiological team. This is a requirement since the extracted flap must contain adequate blood vessel branches for the re-anastomosis with the blood vessels of the chest, in order to ensure the proper vascularisation of the new breast. The radiologists commonly resort to a CTA or MRA to acquire images of these blood vessels and then transmit the findings to the surgeons, which will subsequently design the plan for the extraction of the flap. The process of localising and characterising the DIEPs is very time-consuming and prone to subjectivity, which increases the uncertainty during the assessment by the surgeons. In a collaboration with the Breast Unit

of the Champalimaud Foundation, we have developed a computer vision approach for the semi-automatic extraction of the DIEPs in CTAs. To the best of our knowledge, we were the first targeting the automated extraction of these blood vessels. The main challenges are the small calibre of the vessel segments (1-3mm, which in a typical CTA scan corresponds to 1-5 pixels only), the low SNR even after the injection of contrast, especially in their intramuscular portion, and the existence of nearby structures which are also highlighted by the imaging procedure, like the abdominal muscle. The proposed methodology starts by finding the anterior fascia of the abdominal muscle in order to transform the problem of extracting the DIEPs into two different ones, each of them having their own set of challenges: i) extraction of the subcutaneous course of the DIEPs from manually given points until the fascia, where the major difficulty is to find the correct course when the vessel evolves nearby or even adjacent to the fascia; ii) extraction of the intramuscular course from the point where the subcutaneous tracker ends until the location where the respective DIEA perforates the posterior lamella of the abdominal muscle, a very challenging task since the blood vessel signal is disguised in the muscle, and frequently only some portions of the course are observable. For the subcutaneous part, we designed a centreline tracker which makes use of Frangi's vesselness data to decrease the influence of the 3D plane-like fascia layer. Concerning the intramuscular course, a minimum cost path approach was employed and the usefulness of cost functions combining Frangi's vesselness and intensity information was assessed. The calibre of each DIEP was estimated during its subcutaneous tracking by fitting Gaussian functions to intensity profiles from the cross section of the blood vessel. The experiments compared the findings of the proposed method with those retrieved manually by radiologists, and the following observations were made: a statistical significant difference was verified in calibre estimation, but not when considering only the perforators which ended being included in the flap; having as reference the measurements taken during the surgery, the automated method showed smaller median error in larger perforators yet larger median error in the smaller ones; regarding the location where the perforators pierce the anterior fascia, the automated method and the manual analysis were only statistically different with respect to the vertical component, even though the effect size was small and did not have impact in practice. A time reduction of about 2 h per patient was estimated when using the proposed methodology. These promising results support our conviction that the developed CAD algorithm is capable of supporting the clinicians in this challenging application, leading to a more objective and faster analysis. A prototype software providing a simple interface to conduct our proposed methodology was created and, in the future, it will be tested at the Breast Unit of the Champalimaud Foundation, in order to better assess its usability and suitability, and also further validating the algorithm in a more clinical context. The further automation of report generation and blood vessel extraction will be considered. Concerning the latter, preliminary experiments have shown that, by using vessel enhanced data and hysteresis threshold, the subcutaneous portion of the DIEPs can be extracted with a lot of detail, even allowing to better assess how they communicate between them, a relevant information to have into account during the preoperative planning. Augmented reality systems would also be interesting to consider in this scenario, given that the simultaneous visualisation of the data and the extracted DIEP tree would

give a better perception to the surgeons of the 3D arrangement of the structures of interest.

Another specific scenario which was targeted during this thesis was the segmentation of retinal vessels in fundus photography images. This scenario has been widely studied, not only due to its impact in healthcare, but also because there exist several databases with annotations which allowed to establish proper benchmarks and creating machine learning models. At the time of our research, the state-of-the-art supervised deep learning approaches had already achieved human-like performance regarding metrics such as the AUC and accuracy, nonetheless they were slow during prediction. This was caused by the need of splitting an image into several patches and performing classification for each of them. One of the main necessities for the automated analysis of retinal fundus photos arises during screening programmes, where it is likely that a large number of images will be acquired in a short amount of time. Therefore, it is important to decrease the time spent in the inference process as much as possible. We have designed a convenient and efficient FCN, which allows to run inference on images of different sizes at a single step and avoids time-consuming pre- and post-processing steps. A discussion regarding the use of batch normalisation and dropout, and the overall network design was contemplated. The proposed methodology was able to achieve competitive performance when comparing with the state-of-the-art algorithms while being very fast during inference. It took us less than 5 s in average to make a prediction (Nvidia GeForce GTX 1080 Ti GPU), whereas the state-of-the-art patch classification based approach required, on average, 92 s (Nvidia GTX Titan GPU).

The algorithms discussed in the two preceding paragraphs are application-specific, since they were designed having a particular use case in mind. This does not mean that, for example, the blood vessel centreline extraction methods employed in the DIEP case would not work in other applications, but they would likely be sub-optimal at most, as their design was heavily influenced by particularities of the use case. It comes as no surprise that there is a tendency to use anatomical prior knowledge and heuristics when targeting a very particular use case of blood vessel segmentation. Even then, the relevance of studying methodologies that generalise well to new data is clear. Whenever a new application emerges, such a generalisable algorithm could be promptly used without forcing developers to re-adapt hyper-parameters or, worst, requiring new datasets to re-train the networks, especially if annotations were required. Deep learning has been breaking boundaries in supervised scenarios, yet the learned models tend to perform poorly when they are given differently distributed data, such as the data coming from a different imaging technique or even another blood vessel tree. In this thesis, we explored the combination of supervised deep learning and an unsupervised framework that is known for its generalisation capabilities in blood vessel enhancement, the eigenvalue analysis of the Hessian matrix at multiple scales. By defining expressions based on prior knowledge that respond to the eigenvalues characteristic of tubular structures, a vessel probability is obtained when using these methods. Our motivation was to learn a more optimal vesselness measure, making use of the available annotated data. This can also be seen as a network regularisation mechanism, since instead of letting it learn a mapping from the images themselves, the input is already a richer and more compact representation of local structure. The experiments showed that a shallow network implementing a deep vesselness measure

was able to, not only surpass traditional vesselness metrics, but also generalise better than a Unet model. Despite the promising results, we would like to explore domain adaptation in the future. It is one of the hot topics in machine learning at the moment and it is likely to provide further gains, since it allows to find more complex manifolds representing differently distributed data without the constraint of relying on a simpler prior representation of the data.

In clinical practice, the blood vessel segmentation step is followed by the analysis and characterisation of the extracted tree, which could also be automated or not. Keeping an accurate graph-like structure of the tree is very relevant in this characterisation step, as failing to do so can lead to the misinterpretation of some aspects of the network, or even losing some sub-trees when using automated analysis methods. Even then, the topological similarity between the segmentations and the ground truth are typically overlooked during the benchmarking of proposed blood vessel segmentation algorithms. This may be due to the lack of proper metrics in the literature to assess this property. Moreover, this issue extends to how machine learning models are learned, which is typically by minimising loss functions which only account for pixel-wise error and do not reliably enforce that the graph structure is kept. During this thesis, we conducted research to find proper ways to reduce these limitations, hopefully increasing awareness to the relevance of the topic. Starting with the evaluation of topological properties, we have designed a general similarity index which is based on the paths that are possible to traverse in a tree and having the following properties: i) errors originating disjoint trees or merging different trees are weighted more than those at blood vessel terminations; ii) errors at the main branches of the blood vessel tree are likely to be more penalised. Properties i) and ii) are responsible for penalising more the errors having larger impact on the global graph, making this similarity index distinct from the metrics available in the literature. Two particular designs of the proposed similarity index were presented. Concerning the challenge of learning models which better promote topological consistency, two approaches were proposed. The first one comprises the design of a model architecture where the errors affecting a first segmentation are interpreted as a noise process, which we aim to eliminate through the use of a probabilistic auto-encoding model. Our experiments showed that this architecture produces segmentations which are topologically more coherent than having both segmentation and refinement steps implemented as Unet models, with a total number of parameters similar to the proposed network. The second approach relied on the design of a novel loss function whose minimisation promotes learning models which are better regarding topological coherences. State-of-the-art losses promoting blood vessel centreline consistency (which, despite focusing the graph structure of the vascular tree, do not particularly penalise the errors that most contribute to graph modifications) were extended in order to further penalise the errors mentioned previously in property i). That was accomplished by a loss involving the morphological closing operation and a normalisation function taking into account the length of the error. Experiments showed that the inclusion of the novel loss term allowed a Unet model to reach states that output segmentations which are closer to the reference ones topology-wise, according to the proposed topological similarity indices. We hope that the research conducted in this topic brings more awareness to the need of, not only improving the resilience to these topological errors in future

approaches, but also reporting how good a given methodology is topology-wise. We stress that, with the continuous evolution which methodologies will keep facing, only reporting the typical metrics (AUC, accuracy, sensitivity, and specificity) will be less sufficient as time goes by, since many properties of the produced segmentations will be overlooked in that case. We expect that the development of learning procedures that better capture the topological properties of blood vessel trees will significantly boost the generalisation and robustness of the learned models.

Finally, the research conducted during this thesis includes a fundamental contribution to the ACM literature. ACMs are very commonly used in a myriad of computer vision applications, including biomedical ones. Despite the large number of existing ACMs, to the best of our knowledge, all of them have homogeneous stretching and bending properties along the contour. This is not ideal when an object displays several dynamics and some of them result from noise or other artefacts and, therefore, should not be fitted. One example of this is the segmentation of the lungs in CT images, where a state-of-the-art ACM cannot follow the natural dynamics of these organs and, at the same time, include in the segmentation the nodules lying in their peripheral region. The inclusion of these nodules is essential for the steps that follow, nodule detection and characterisation and, possibly, the analysis of the surrounding blood vessel patterns to gather more insight. A novel parametric ACM, the SMB snake, was proposed to address this limitation of the literature, by allowing to have contours with a finite number of contiguous regions with different bending properties. To make this possible, a novel energy function inducing sparse transitions (L_0 norm) in the rigidity coefficients along the contour was designed. To optimise our novel energy function, a pairwise coordinate descent strategy was derived. The conducted experiments have shown how an adequate parametrisation of the SMB snake is able to achieve contours that state-of-the-art models cannot. In addition to the scenario of lung segmentation, where our model was the one dealing better with the trade-off between accurate delineation and peripheral nodule inclusion, we have also demonstrated its properties in experiments concerning synthetic and hand gesture images.

With the outcomes of this thesis, we believe that: (i) AI-based blood vessel segmentation was shown to still be far from mimicking the rationale a human employs when solving the task, such that there are still several open challenges to address; (ii) we have paved the way for the employment of computer vision in the preoperative planning of DIEP flaps; (iii) the array of methodologies available for clinicians who deal with blood vessel segmentation and/or analysis was increased, with the attenuation of some of the limitations of the literature; (iv) the relevance of topics which have not been very focused during blood vessel segmentation research, such as the topological properties of the vascular trees, was stressed, hopefully raising their awareness; and (v) the applicability of ACMs was extended with the proposal of the novel SMB snake.

Bibliography

[1] R.J. Hodes. Aging hearts and arteries: a scientific quest. *National Institute of Aging: Bethesda, MA, USA*, 1990.

[2] American Heart Association. *What is an Aneurysm?*, (accessed May 6, 2020). https://www.heart.org/en/health-topics/aortic-aneurysm/what-is-an-aneurysm.

[3] I. Kronzon and P.A. Tunick. Aortic atherosclerosis disease and stroke. *Circulation*, 114(1):63–75, 2016.

[4] S. Dalager-Pedersen, H.B. Ravn, and E. Falk. Atherosclerosis and acute coronary events. *The American Journal of Cardiology*, 82(10):37–40, 1998.

[5] V.F. Tapson. Acute pulmonary embolism. *The New England Journal of Medicine*, 358:1037–1052, 2008.

[6] National Heart, Lung, and Blood Institute. *Atherosclerosis*, (accessed May 6, 2020). https://www.nhlbi.nih.gov/health-topics/atherosclerosis.

[7] R. Hannah and R. Max. Causes of death. *Our World in Data*, 2020. https://ourworldindata.org/causes-of-death.

[8] World Health Organization. *Cardiovascular diseases data and statistics*, (accessed May 6, 2020). www.euro.who.int/en/health-topics/noncommunicable-diseases/cardiovascular-diseases/data-and-statistics.

[9] S.S. Hayreh, B. Zimmerman, M.J. McCarthy, and P. Podhajsky. Systemic diseases associated with various types of retinal vein occlusion. *American Journal of Ophthalmology*, 131(1):61–77, 2001.

[10] L. Rousso and J. Sowka. Recognizing abnormal vasculature: a guide to following and educating patients who face this class of sight-threatening diagnoses. *Review of Optometry*, 154(1):82–87, 2017.

[11] T.Y. Wong, R. Klein, D.J. Couper, L.S. Cooper, E. Shahar, L.D. Hubbard, M.R. Wofford, and A.R. Sharrett. Retinal microvascular abnormalitites and incident stroke: the atherosclerosis risk in communities study. *The Lancet*, 358(9288):1134–1140, 2001.

[12] R.I. Lindley, J.J. Wang, M.-C. Wong, P. Mitchell, G. Liew, P. Hand, J. Wardlaw, D.A. De Silva, M. Baker, E. Rochtchina, C. Chen, G.J. Hankey, H.-M. Chang, V.S.C. Fung, L. Gomes, and T.Y. Wong. Retinal microvasculature in acute lacunar stroke: a cross-sectional study. *The Lancet Neurology*, 8(7):628–634, 2009.

[13] C. Y.-L. Cheung, M. K. Ikram, C. Chen, and T. Y. Wong. Imaging retina to study dementia and stroke. *Progress in Retinal and Eye Research*, 57:89–107, 2017.

[14] American Academy of Ophthalmology Retina/Vitreous Panel. Preferred practice pattern guidelines. diabetic retinopathy. 2019. https://www.aao.org/preferred-practice-pattern/diabetic-retinopathy-ppp.

[15] International Diabetes Federation. *IDF diabetes atlas 9th edition 2019*, (accessed May 21, 2020). https://diabetesatlas.org/en/.

[16] M.M. Nentwich and M.W. Ulbig. Diabetic retinopathy - ocular complications of diabetes mellitus. *World Journal of Diabetes*, 6(3):489–499, 2015.

[17] C. Viallard and B. Larrivée. Tumor angiogenesis and vascular normalization: alternative therapeutic targets. *Angiogenesis*, 20:409–426, 2017.

[18] K. Hida, N. Maishi, and Y. Hida. *Tumor blood vessels as targets for cancer therapy*, pages 41–56. 2019.

[19] F. Lee. *Liver Cancer New Research*. Nova Science Publishers, 2006.

[20] W.M. Rozen, N.D. Houseman, and M.W. Ashton. The absent inferior epigastric artery: a unique anomaly and implications for deep inferior epigastric artery perforator flaps. *Journal of Reconstructive Microsurgery*, 25(5):289–293, 2009.

[21] D.E. Janhofer, C. Lakhiani, P.J. Kim, C. Akbari, I. Naz, E.A. Tefera, C. Attinger, and K.K. Evans. The utility of preoperative arteriography for free flap planning in patients with chronic lower extremity wounds. *Plastic and Reconstructive Surgery*, 143(2):604–613, 2019.

[22] A. Cina, M. Salgarello, L. Barone-Adesi, P. Rinaldi, and L. Bonomo. Planning breast reconstruction with deep inferior epigastric artery perforating vessels: multidetector ct angiography versus color doppler us. *Radiology*, 255(3):979–987, 2010.

[23] O. Ronneberger, P. Fischer, and T. Brox. U-net: convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. Wells, and A. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, volume 9351 of *Lecture Notes in Computer Science*. Springer, Cham, 2015.

[24] R.J. Araújo. Computer aided detection of perforating arteries in ct angiography, 2016.

[25] R. Drake, A.W. Vogl, and A.W.M. Mitchell. *Gray's anatomy for students e-book*. Elsevier Health Sciences, 2009.

[26] T.J. Phillips, D.L. Stella, W.M. Rozen, M. Ashton, and G.I. Taylor. Abdominal wall ct angiography: a detailed account of a newly established preoperative imaging technique. *Radiology*, 249(1):32–44, 2008.

[27] D. Prada, A. Harris, G. Guidoboni, L. Rowe, A.C. Verticchio-Vercellin, and S. Mathew. Vascular anatomy and physiology of the eye. In G. Guidoboni, A. Harris, and R. Sacco, editors, *Ocular Fluid Dynamics*, pages 23–45. Birkh auser, Cham, 2019.

[28] M. Sit and A.V. Levin. Direct ophthalmoscopy in pediatric emergency care. *Pediatric emergency care*, 17(3):199–204, 2001.

[29] Z. Zielicka, S. Hull, and N. Davies. Direct versus indirect ophthalmoscopy: medical student assessment of retinal pathology. *Acta Ophthalmologica*, 91, 2013.

[30] G. Liew, P. Mitchell, J.J. Wang, and T.Y. Wong. Fundoscopy: to dilate or not to dilate? *BMJ*, 2006.

[31] L. Mottow-Lippa. Ophthalmology in the medical school curriculum: reestablishing our value and effecting change. *Ophthalmology*, 116(7):1235–1236, 2009.

[32] V. Biousse, B.B. Bruce, and N.J. Newman. Ophthalmoscopy in the 21st century: the 2017 h. houston merritt lecture. *Neurology*, 90(4):167–175, 2018.

[33] P. Nishtha, P. Huang, J. Lee, P.A. Keane, T.S. Chuan, A. Richhariya, S. Teoh, T.H. Lim, and R. Agrawal. Fundus photography in the 21st century - a review of recent technological advances and their implications for worldwide healthcare. *Telemedicine and e-Health*, 22(3):198–208, 2016.

[34] S.M. Saleem, L.R. Pasquale, P.A. Sidoti, and J.C. Tsai. Virtual ophthalmology: telemedicine in a covid-19 era. *American Journal of Ophthalmology*, 216:237–242, 2020.

[35] R.F. Spaide, J.M. Klancnik, and M.J. Cooney. Retinal vascular layers imaged by fluorescein angiography and optical coherence tomography angiography. *JAMA Ophthalmology*, 133(1):45–50, 2015.

[36] M.A. Klufas, N.A. Yannuzzi, C.E. Pang, S. Srinivas, S.R. Sadda, K. B. Freund, and S. Kiss. Feasibility and clinical utility of ultra-widefield indocyanine green angiography. *Retina*, 35(3):508–520, 2015.

[37] T.E. De Carlo, A. Romano, N.K. Waheed, and J.S. Duker. A review of optical coherence tomography angiography (octa). *International Journal of Retina and Vitreous*, 1(1):5, 2015.

[38] J. Staal, M.D. Abràmoff, and M. Niemeijer. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, 2004.

[39] A. Hoover, V. Kouznetsova, and M. Goldbaum. Locating blood vessels in retinal images by piece-wise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19(3):203–210, 2000.

[40] C.G. Owen, A.R. Rudnicka, R. Mullen, S.A. Barman, D. Monekosse, P.H. Whincup, J. Ng, and C. Paterson. Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (caiar) program. *Investigative ophthalmology & visual science*, 50(5):2004–2010, 2009.

[41] P.J. Sparrow, N. Merchant, Y.L. Provost, D.J. Doyle, E.T. Nguyen, and N.S. Paul. Ct and mr imaging findings in patients with acquired heart disease at risk for sudden cardiac death. *Radiographics*, 29(3):805–823, 2009.

[42] E. Maffei, S. Seitun, A.I. Guaricci, and F. Cademartiri. Chest pain: coronary ct in the er. *The British Journal of Radiology*, 89(1061):20150954, 2016.

[43] R.A. Byrne, G.W. Stone, J. Ormiston, and A. Kastrati. Coronary balloon angioplasty, stents, and scaffolds. *The Lancet*, 390(10096):781–792, 2017.

[44] G. Stefanini and S. Windecker. Can coronary computed tomography angiography replace invasive angiography? *Circulation*, 131(4):418–426, 2015.

[45] S. Achenbach and W.G. Daniel. Cardiac imaging in the patient with chest pain: coronary ct angiography. *Heart*, 96(15):1241–1246, 2010.

[46] F. Cervantes-Sanchez, I. Cruz-Aceves, A. Hernandez-Aguirre, M.A. Hernandez-Gonzalez, and S.E. Solorio-Meza. Automatic segmentation of coronary arteries in x-ray angiograms using multiscale analysis and artificial neural networks. *Applied Sciences*, 9(24):5507, 2019.

[47] R. Ohkuma, R. Mohan, P.A. Baltodano, M.J. Lacayo, J.M. Broyles, E.B. Schneider, M. Yamazaki, D.S. Cooney, M.A. Manahan, and G.D. Rosson. Abdominally based free flap planning in breast reconstruction with computed tomographic angiography: systematic review and meta-analysis. *Plastic and Reconstructive Surgery*, 133(3):483–494, 2014.

[48] S. Aubry, J. Pauchot, A. Kastler, O. Laurent, Y. Tropet, and M. Runge. Preoperative imaging in the planning of deep inferior epigastric artery perforator flap surgery. *Skeletal Radiology*, 42:319–327, 2012.

[49] A. Cina, L. Barone-Adesi, P. Rinaldi, A. Cipriani, M. Salgarello, R. Masetti, and L. Bonomo. Planning deep inferior epigastric perforator flaps for breast reconstruction: a comparison between multidetector computed tomography and magnetic resonance angiography. *European Radiology*, 23(8):2333–2343, 2013.

[50] A.T. Mohan and M. Saint-Cyr. Advances in imaging technologies for planning breast reconstruction. *Gland Surgery*, 5(2):242–254, 2016.

[51] Mayo Clinic. *Chest X-rays*, (accessed December 18, 2020). https://www.mayoclinic.org/tests-procedures/chest-x-rays/about/pac-20393494.

[52] American Lung Association. *CT Scan*, (accessed December 18, 2020). https://www.lung.org/lung-health-diseases/lung-procedures-and-tests/ct-scan.

[53] A.S. Bhalla, A. Das, P. Naranje, A. Irodi, V. Raj, and A. Goyal. Imaging protocols for ct chest: a recommendation. *The Indian Journal of Radiology & Imaging*, 29(3):236–246, 2019.

[54] T. van der Hulle, C.E.A. Dronkers, F.A. Klok, and M.V. Huisman. Recent developments in the diagnosis and treatment of pulmonary embolism. *Journal of Internal Medicine*, 279(1):16–29, 2016.

[55] M. Yazdani, C.T. Lau, J.K. Lempel, R. Yadav, El-Sherief A.H., J.T. Azok, and R.D. Renapurkar. Historical evolution of imaging techniques for the evaluation of pulmonary embolism. *Radiographics*, 35(4):1245–1262, 2015.

[56] R.D. Rudyanto, S. Kerkstra, E.M. van Rikxoort, C. Fetita, P.-Y. Brillet, C. Lefevre, W. Xue, and B. van Ginneken. Comparing algorithms for automated vessel segmentation in computed tomography scans of the lung: the vessel12 study. *Medical Image Analysis*, 18(7):1217–1232, 2014.

[57] C. Kirbas and F.K.H. Quek. Vessel extraction techniques and algorithms: a survey. In *Third IEEE Symposium on Bioinformatics and Bioengineering*, pages 238–245. IEEE, 2003.

[58] D. Lesage, E.D. Angelini, I. Bloch, and G. Funka-Lea. A review of 3d vessel lumen segmentation techniques: models, features and extraction schemes. *Medical Image Analysis*, 13(6):819–845, 2009.

[59] M.M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, and S.A. Barman. Blood vessel segmentation methodologies in retinal images - a survey. *Computer Methods in and Programs in Biomedicine*, 108(1):407–433, 2012.

[60] S. Moccia, E. De Momi, S. El Hadji, and L.S. Mattos. Blood vessel segmentation algorithms - review of methods , datasets and evaluation metrics. *Computer Methods and Programs in Biomedicine*, 158:71–91, 2018.

[61] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Transactions on Medical Imaging*, 8(3):263–269, 1989.

[62] M. Al-Rawi, M. Qutaishat, and M. Arrar. An improved matched filter for blood vessel detection of digital retinal images. *Computers in Biology and Medicine*, 37(2):262–267, 2007.

[63] R. Poli and G. Valli. An algorithm for real-time vessel enhancement and detection. *Computer Methods and Programs in Biomedicine*, 52(1):1–22, 1997.

[64] L. Gang, O. Chutatape, and S.M. Krishnan. Detection and measurement of retinal vessels in fundus images using amplitude modified second-order gaussian filter. *IEEE Transactions on Biomedical Engineering*, 49(2):168–172, 2002.

[65] M. Sofka and C.V. Stewart. Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures. *IEEE Transactions on Medical Imaging*, 25(12):1531–1546, 2006.

[66] B. Zhang, L. Zhang, L. Zhang, and F. Karray. Retinal vessel extraction by matched filter with first-order derivative of gaussian. *Computers in Biology and Medicine*, 40(4):438–445, 2010.

[67] O. Chutatape, L. Zheng, and S.M. Krishnan. Retinal blood vessel detection and tracking by matched gaussian and kalman filters. In *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 6, pages 325–334. Springer, Cham, 1998.

[68] G. Azzopardi, N. Strisciuglio, M. Vento, and N. Petkov. Trainable cosfire filters for vessel delineation with application to retinal images. *Medical Image Analysis*, 19(1):46–57, 2015.

[69] O. Friman, M. Hindennach, C. Kuhnel, and H.-O. Peitgen. Multiple hypothesis template tracking of small 3d vessel structures. *Medical Image Analysis*, 14(2):160–171, 2010.

[70] Y. Wang, G. Ji, P. Lin, and E. Trucco. Retinal vessel segmentation using multiwavelet kernels and multiscale hierarchical decomposition. *Pattern Recognition*, 46(8):2117–2133, 2013.

[71] C.K. Chui and J.-a. Lian. A study of orthonormal multi-wavelets. *Applied Numerical Mathematics*, 20(3):273–298, 1996.

[72] J. Odstrcilik, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Cernosek, O. Svoboda, and E. Angelopoulou. Retinal vessel segmentation by improved matched filtering: Evaluation on a new high-resolution fundus image database. *IET Image Processing*, 7(4):373–383, 2013.

[73] C. Lorenz, I.-C. Carlsen, T.M. Buzug, C. Fassnacht, and J. Weese. Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2d and

3d medical images. In *Proceedings of the First Joint Conference on Computer Vision, Virtual Reality and Robotics in Medicine and Medical Robtics and Computer-Assisted Surgery*, pages 233–242. Springer, 1997.

[74] Y. Sato, S. Nakajima, N. Shiraga, H. Atsumi, S. Yoshida, T. Koller, G. Gerig, and R. Kikinis. Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Medical Image Analysis*, 2(2):143–168, 1998.

[75] A. Frangi, W. Niessen, K. Vincken, and M. Viergever. Multi-scale vessel enhancement filtering. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 130–137. Springer, 1998.

[76] K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y. Trousset. Model based detection of tubular structures in 3d images. *Computer Understanding*, 80(2):130–171, 2000.

[77] Q. Li, S. Sone, and K. Doi. Selective enhancement filters for nodules, vessels, and airway walls in two-and three-dimensional ct scans. *Medical Physics*, 30(8):2040–2051, 2003.

[78] M. Erdt, M. Raspe, and M. Suehling. Automatic hepatic vessel segmentation using graphics hardware. In *International Workshop on Medical Imaging and Augmented Reality*, pages 403–412. Springer, 2008.

[79] T. Jerman, F. Pernus, B. Likar, and Z. Spiclin. Enhancement of vascular structures in 3d and 2d angiographic images. *IEEE Transactions on Medical Imaging*, 35(9):2107–2118, 2016.

[80] T. Lindeberg. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):117–156, 1998.

[81] C. Zhou, H.-P. Chan, B. Sahiner, L.M. Hadjiiski, A. Chughtai, S. Patel, and J. Wei. Automatic multiscale enhancement and segmentation of pulmonary vessels in ct pulmonary angiography images for cad applications. *Medical Physics*, 34(12):4567–4577, 2007.

[82] R. Manniesing, M.A. Viergever, and W.J. Niessen. Vessel enhancing diffusion: A scale space representation of vessel structures. *Medical Image Analysis*, 10(6):815–825, 2006.

[83] B.D. Thackray and A.C. Nelson. Semi-automatic segmentation of vascular network images using a rotating structuring element (rose) with mathematical morphology and dual feature thresholding. *IEEE Transactions on Medical Imaging*, 12(3):385–392, 1993.

[84] F. Zana and J-C. Klein. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE Transactions on Image Processing*, 10(7):1010–1019, 2001.

[85] T. Walter and J.-C. Klein. Segmentation of color fundus images of the human retina: detection of the optic disc and the vascular tree using morphological techniques. In *International Symposium on Medical Data Analysis*, pages 282–287, 2001.

[86] A.M. Mendonça and A. Campilho. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE Transactions on Medical Imaging*, 25(9):1200–1213, 2006.

[87] M.A.T. Figueiredo and J.M.N. Leitão. A nonsmoothing approach to the estimation of vessel contours in angiograms. *IEEE Transactions on Medical Imaging*, 14(1):162–172, 1995.

[88] M.H. Wilkinson and M.A. Westenberg. Shape preserving filament enhancement filtering. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 770–777. Springer, 2001.

[89] E.R. Urbach and M.H. Wilkinson. Shape-only granulometries and greyscale shape filters. In *International Symposium on Mathematical Morphology*, pages 305–314, 2002.

[90] B. Naegel, N. Passat, and C. Ronse. Grey-level hit-or-miss transforms – part ii: Application to angiographic image processing. *Pattern Recognition*, 40(2):648–658, 2007.

[91] B. Bouraoui, C. Ronse, J. Baruthio, N. Passat, and P. Germain. 3d segmentation of coronary arteries based on advanced mathematical morphology techniques. *Computerized Medical Imaging and Graphics*, 34(5):377–387, 2010.

[92] A. Dufour, O. Tankyevych, B. Naegel, H. Talbot, C. Ronse, J. Baruthio, P. Dokládal, and N. Passat. Filtering and segmentation of 3d angiographic data: advances based on mathematical morphology. *Medical Image Analysis*, 17(2):147–164, 2013.

[93] Y. Sun. Automated identification of vessel contours in coronary arteriograms by an adaptive tracking algorithm. *IEEE Transactions on Medical Imaging*, 8(1):78–88, 1989.

[94] L. Zhou, M.S. Rzeszotarski, L.J. Singerman, and J.M. Chokreff. The detection and quantification of retinopathy using digital angiograms. *IEEE Transactions on Medical Imaging*, 13(4):619–626, 1994.

[95] R.E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME – Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[96] Y.A. Tolias and S.M. Panas. A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. *IEEE Transactions on Medical Imaging*, 17(2):263–273, 1998.

[97] S. Wörz and K. Rohr. Segmentation and quantification of human vessels using a 3-d cylindrical intensity model. *IEEE Transactions on Image Processing*, 16(8):1994–2004, 2007.

[98] Y. Yin, M. Adel, and S. Bourennane. Retinal vessel segmentation using a probabilistic tracking method. *Pattern Recognition*, 45(4):1235–1244, 2012.

[99] E. Bekkers, R. Duits, T. Berendschot, and B. ter Haar Romeny. A multiorientation analysis approach to retinal vessel tracking. *Journal of Mathematical Imaging and Vision*, 49(3):583–610, 2014.

[100] S.Y. Wan and W.E. Higgins. Symmetric region growing. *IEEE Transactions on Image processing*, 12(9):1007–1015, 2003.

[101] T. Boskamp, D. Rinck, F. Link, B. Kummerlen, G. Stamm, and P. Mildenberger. New vessel analysis tool for morphometric quantification and visualization of vessels in ct and mr imaging data sets. *Radiographics*, 24(1):287–297, 2004.

[102] M.E. Martínez-Pérez, A.D. Hughes, A.V. Stanton, S.A. Thom, A.A. Bharath, and K.H. Parker. Retinal blood vessel segmentation by means of scale-space analysis and region growing. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 90–97. Springer, 1999.

[103] M.E. Martínez-Pérez, A.D. Hughes, S.A. Thom, A.A. Bharath, and K.H. Parker. Segmentation of blood vessels from red-free and fluorescein retinal images. *Medical Image Analysis*, 11(1):47–61, 2007.

[104] M.A. Palomera-Pérez, M.E. Martínez-Pérez, H. Benítez-Pérez, and J.L. Ortega-Arjona. Parallel multiscale feature extraction and region growing: Application in retinal blood vessel detection. *IEEE Transactions on Information Technology in Biomedicine*, 14(2):500–506, 2010.

[105] C. Metz, M. Schaap, A. van der Giessen, T. van Walsum, and W.J. Niessen. Semi-automatic coronary artery centerline extraction in computed tomography angiography data. In *International Symposium in Biomedical Imaging: From Nano to Macro, 2007*, pages 856–859. IEEE, 2007.

[106] Y. Li, S. Zhou, J. Wu, X. Ma, and K. Peng. A novel method of vessel segmentation for x-ray coronary angiography images. In *Fourth International Conference on Computational and Information Sciences (ICCIS)*, pages 468–471. IEEE, 2012.

[107] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[108] Y.Q. Zhao, X.H. Wang, X.F. Wang, and F.Y. Shih. Retinal vessels segmentation based on level set and region growing. *Pattern Recognition*, 47(7):2437–2446, 2014.

[109] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

[110] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.

[111] T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[112] S. Osher and J.A. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.

[113] C. Li, C. Xu, C. Gui, and M.D. Fox. Distance regularized level set evolution and its application to image segmentation. *IEEE Transactions on Image Processing*, 19(12):3243–3254, 2010.

[114] K. Sum and P.Y. Cheung. Vessel extraction under non-uniform illumination: a level set approach. *IEEE Transactions on Biomedical Engineering*, 55(1):358–360, 2008.

[115] K. Sun, Z. Chen, and S. Jiang. Local morphology fitting active contour for automatic vascular segmentation. *IEEE Transactions on Biomedical Engineering*, 59(2):464–473, 2012.

[116] B. Al-Diri, A. Hunter, and D. Steel. An active contour model for segmenting and measuring retinal vessels. *IEEE Transactions on Medical Imaging*, 28(9):1488–1497, 2009.

[117] G. Läthén, J. Jonasson, and M. Borga. Blood vessel segmentation using multiscale quadrature filtering. *Pattern Recognition Letters*, 31(8):762–767, 2010.

[118] Y. Shang, R. Deklerck, E. Nyssen, A. Markova, J. de Mey, X. Yang, and K. Sun. Vascular active contour for vessel tree segmentation. *IEEE Transactions on Biomedical Engineering*, 58(4):1023–1032, 2011.

[119] Y. Zhao, L. Rada, K. Chen, S.P. Harding, and Y. Zheng. Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images. *IEEE Transactions on Medical Imaging*, 34(9):1797–1807, 2015.

[120] M. Barchiesi, S.H. Kang, T.M. Le, M. Morini, and M. Ponsiglione. A variational model for infinite perimeter segmentations based on lipschitz level set functions: denoising while keeping finely oscillatory boundaries. *Multiscale Modeling & Simulation*, 8(5):1715–1741, 2010.

[121] L.R. Ford and D.R. Fulkerson. Maximal flow through a network. *Canadian Journal of Mathematics*, 8(3):399–404, 1956.

[122] G. Slabaugh and G. Unal. Graph cuts segmentation using an elliptical shape prior. In *International Conference on Image Processing (ICIP)*, volume 2, pages II–1222. IEEE, 2005.

[123] M. Schaap, L. Neefjes, C. Metz, A. van der Giessen, A. Weustink, N. Mollet, J. Wentzel, T. van Walsum, and W.J. Niessen. Coronary lumen segmentation using graph cuts and robust kernel regression. In *International Conference on Information Processing in Medical Imaging 2009*, pages 528–539. Springer, 2009.

[124] S. Esneault, C. Lafon, and J.-L. Dillenseger. Liver vessels segmentation using a hybrid geometrical moments/graph cuts method. *IEEE Transactions on Biomedical Engineering*, 57(2):276–283, 2010.

[125] V. Pamulapati, B.J. Wood, and M.G. Linguraru. Intra-hepatic vessel segmentation and classification in multi-phase ct using optimized graph cuts. In *2011 IEEE International Symposium in Biomedical Imaging (ISBI): From Nano to Macro*, pages 1982–1985, 2011.

[126] Y. Zhao, Y. Liu, X. Wu, S.P. Harding, and Y. Zheng. Retinal vessel segmentation: an efficient graph cut approach with retinex and local phase. *Plos One*, 10(4), 2015.

[127] E.H. Land. Recent advances in retinex theory. *Vision Research*, 26(1):7–21, 1986.

[128] D.P. Kingma, S. Mohamed, D.J. Rezende, and M. Welling. Semi-supervised learning with deep generative models. In *Advances in Neural Information Processing Systems*, pages 3581–3589, 2014.

[129] Y. Wei, X. Liang, Y. Chen, X. Shen, M.-M. Cheng, J. Feng, Y. Zhao, and S. Yan. Stc: a simple to complex framework for weakly-supervised semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(11):2314–2320, 2016.

[130] Z. Jiang, H. Zhang, Y. Wang, and S.-B. Ko. Retinal blood vessel segmentation using fully convolutional network with transfer learning. *Computerized Medical Imaging and Graphics*, 68:1–15, 2018.

[131] S. Liu, A.J. Davison, and E. Johns. Self-supervised generalisation with meta auxiliary learning. *arXiv:1901.08933*, 2019.

[132] M. Javanmardi and T. Tasdizen. Domain adaptation for biomedical image segmentation using adversarial training. In *15th IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 554–558. IEEE, 2018.

[133] S. Abassi-Sureshjani, I. Smit-Ockelon, J. Zhang, and B.T.H. Romeny. Biologically-inspired supervised vasculature segmentation in slo retinal fundus images. In *International Conference Image Analysis and Recognition*, pages 325–334. Springer, Cham, 2015.

[134] M. Niemeijer, J. Staal, B. van Ginneken, M. Loog, and M.D. Abramoff. Comparative study of retinal vessel segmentation methods on a new publicly available database. In *Medical Imaging 2004: Image Processing*, volume 5370, pages 648–656. SPIE, 2004.

[135] J.V. Soares, J.J. Leandro, R.M. Cesar, H.F. Jelinek, and M.J. Cree. Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. *IEEE Transactions on Medical Imaging*, 25(9):1214–1222, 2006.

[136] E. Ricci and R. Perfetti. Retinal blood vessel segmentation using line operators and support vector classification. *IEEE Transactions on Medical Imaging*, 26(10):1357–1365, 2007.

[137] C.A. Lupascu, D. Tegolo, and E. Trucco. Fabc: Retinal vessel segmentation using adaboost. *IEEE Transactions on Information Technology in Biomedicine*, 14(5):1267–1274, 2010.

[138] D. Marín, A. Aquino, M.E. Gegúndez-Arias, and J.M. Bravo. A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. *IEEE Transactions on Medical Imaging*, 30(1):146–158, 2011.

[139] M.M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, and S.A. Barman. An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Transactions on Biomedical Engineering*, 59(9):2538–2548, 2012.

[140] S. Roychowdhury, D.D. Koozekanani, and K.K. Parhi. Blood vessel segmentation of fundus images by major vessel extraction and sub-image classification. *IEEE Journal of Biomedical and Health Informatics*, 19(3):1118–1128, 2015.

[141] N. Strisciuglio, G. Azzopardi, M. Vento, and N. Petkov. Supervised vessel delineation in retinal fundus images with the automatic selection of b-cosfire filters. *Machine Vision and Applications*, 27(8):1137–1149, 2016.

[142] J.I. Orlando, E. Prokofyeva, and M.B. Blaschko. A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images. *IEEE Transactions on Biomedical Engineering*, 64(1):16–27, 2017.

[143] P. Krähenbühl and V. Koltun. Efficient inference in fully connected crfs with gaussian edge potentials. In *Advances in Neural Information Processing Systems*, pages 109–117, 2012.

[144] T. Joachims, T. Finley, and C.N.J. Yu. Cutting-plane training of of structural svms. *Machine Learning*, 77(1):27–59, 2009.

[145] U.T.V. Nguyen, A. Bhuiyan, L.A.F. Park, and K. Ramamohanarao. An effective retinal blood vessel segmentation method using multi-scale line detection. *Pattern Recognition*, 46(3):703–715, 2013.

[146] J. Zhang, Y. Chen, E. Bekkers, M. Wang, B. Dashtbozorg, and B.M. ter Haar Romeny. Retinal vessel delineation using a brain-inspired wavelet transform and random forest. *Pattern Recognition*, 69:107–123, 2017.

[147] X. Wang, X. Jiang, and J. Ren. Blood vessel segmentation from a fundus image by a cascade classification framework. *Pattern Recognition*, 88:331–341, 2019.

[148] X.D. Jiang. Image detail-preserving filter for impulsive noise attenuation. *IEE Proceedings - Vision, Image and Signal Processing*, 150(3):179–185, 2003.

[149] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[150] A. Krizhevsky, I. Sutskever, and G.E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 1097–1105, 2012.

[151] A. Anwar. *Difference between AlexNet, VGGNet, ResNet, and Inception*, (accessed October 17, 2020). https://towardsdatascience.com/the-w3h-of-alexnet-vggnet-resnet-and-inception-7baaaecccc96.

[152] M. Melinščak, P. Prentašić, and S. Lončarić. Retinal vessel segmentation using deep neural networks. In *10th International Conference on Computer Vision Theory and Applications (VISAPP)*, 2015.

[153] P. Liskowski and K. Krawiec. Segmenting retinal blood vessels with deep neural networks. *IEEE Transactions on Medical Imaging*, 35(11):2369–2380, 2016.

[154] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang. A cross-modality learning approach for vessel segmentation in retinal images. *IEEE Transactions on Medical Imaging*, 35(1):109–118, 2016.

[155] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440. IEEE, 2015.

[156] A. Dasgupta and S. Singh. A fully convolutional neural network based structure prediction towards the retinal vessel segmentation. In *14th International Symposium on Biomedical Imaging (ISBI)*, pages 248–251. IEEE, 2017.

[157] Z. Feng, J. Yang, and L. Yao. Patch-based fully convolutional neural network with skip connections for retinal blood vessel segmentation. In *International Conference on Image Processing (ICIP)*, pages 1742–1746. IEEE, 2017.

[158] Y. Zhang and A.C.S. Chung. Deep supervision with additional labels for retinal vessel segmentation task. In A. Frangi, J. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger, editors, *Medical Image Computing and Computer Assisted Intervention - MICCAI*, volume 11071. Springer, Cham, 2018.

[159] Q. Jin, Z. Meng, T.D. Pham, Q. Chen, L. Wei, and R. Su. Dunet: a deformable network for retinal vessel segmentation. *Knowledge-based systems*, 178:149–162, 2019.

[160] A. Oliveira, S. Pereira, and C.A. Silva. Retinal vessel segmentation based on fully convolutional neural networks. *Expert Systems with Applications*, 112(1):229–242, 2018.

[161] H. Fu, Y. Xu, S. Lin, D.W.K. Wong, and J. Liu. Deepvessel: retinal vessel segmentation via deep learning and conditional random field. In S. Ourselin, L. Joskowicz, M. Sabuncu, G. Unal, and W. Wells, editors, *Medical Image Computing and Computer Assisted Intervention - MICCAI*, volume 9901. Springer, Cham, 2016.

[162] S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, C. Huang, and P.H.S. Torr. Conditional random fields as recurrent neural networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1529–1537, 2015.

[163] L.C. Neto, G.L.B. Ramalho, J.F.S.R. Neto, R.M.S. Veras, and F.N.S. Medeiros. An unsupervised coarse-to-fine algorithm for blood vessel segmentation in fundus images. *Expert Systems with Applications*, 78(15):182–192, 2017.

[164] Z. Fan, J. Lu, C. Wei, H. Huang, X. Cai, and X. Chen. A hierarchical image matting model for blood vessel segmentation in fundus images. *IEEE Transactions on Image Processing*, 28(5):2367–2377, 2019.

[165] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, and A. Jemal. Global cancer statistics 2018: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6):394–424, 2018.

[166] National Research Council. *Meeting psychosocial needs of women with breast cancer*. National Academies Press, 2004.

[167] O.D. Gentilini, M.J. Cardoso, and P. Poortmans. Less is more. breast conservation might be even better than mastectomy in eraly breast cancer patients. *The Breast*, 35:32–33, 2017.

[168] A.E. Dragun, B. Huang, T.C. Tucker, and W.J. Spanos. Increasing mastectomy rates among all age groups for early stage breast cancer: a 10-year study of surgical choice. *Breast Journal*, 18:318–325, 2012.

[169] U. Mahmood, A.L. Hanlon, M. Koshy, R. Buras, S. Chumsri, K.H. Tkaczuk, S.B. Cheston, W.F. Regine, and S.J. Feigenberg. Increasing national mastectomy rates for the treatment of early stage breast cancer. *Annals of Surgical Oncology*, 20:1436–1443, 2013.

[170] Breastcancer.org. *Is mastectomy right for you?*, (accessed December 12, 2020). https://www.breastcancer.org/treatment/surgery/mastectomy/who_for.

[171] C.J. Lange, N.D. Thimmappa, S.R. Boddu, S.P. Dutruel, M. Pei, Z. Farooq, A.H. Behzadi, Y. Wang, R. Zabih, and M.R. Prince. Automating perforator flap mra and cta reporting. *Journal of Digital Imaging*, 30:350–357, 2017.

[172] T. Gómez-Cía, P. Gacto-Sánchez, D. Sicilia, C. Suárez, B. Acha, C. Serrano, C. Parra, and J. De La Higuera. The virtual reality tool virsspa in planning diep microsurgical breast reconstruction. *International Journal of Computer Assisted Radiology and Surgery*, 4:375–382, 2009.

[173] W.M. Rozen, E. Garcia-Tutor, A. Alonso-Burgos, R. Acosta, F. Stillaert, J.L. Zubieta, M. Hamdi, I.S. Whitaker, and M.W. Ashton. Planning and optimising diep flaps with virtual surgery: the navarra experience. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 63:289–297, 2010.

[174] R.A. Araújo and H.P. Oliveira. Segmentation of the rectus abdominis muscle anterior fascia for the analysis of deep inferior epigastric perforators. In L. Alexandre, J. Salvador Sánchez, and J. Rodrigues, editors, *Pattern Recognition and Image Analysis*, pages 537–545, 2017.

[175] G. Agam, S.G. Armato, and C. Wu. Vessel tree reconstruction in thoracic ct scans with application to nodule detection. *IEEE Transactions on Medical Imaging*, 24:486–499, 2005.

[176] H.P. Oliveira, J.S. Cardoso, A.T. Magalhães, and M.J. Cardoso. A 3d low-cost solution for the aesthetic evaluation of breast cancer conservative treatment. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, 2(2):90–106, 2014.

[177] P.E. Hart, N.J. Nilsson, and B. Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968.

[178] C.T. Metz, M. Schaap, A.C. Weustink, N.R. Mollet, T. van Walsum, and W.J. Niessen. Coronary centerline extraction from ct coronary angiography images using a minimum cost path approach. *Medical Physics*, 36:5568–5579, 2009.

[179] C. Mavioso, R.J. Araújo, H.P. Oliveira, J.C. Anacleto, M.A. Vasconcelos, D. Pinto, P.F. Gouveia, C. Alves, F. Cardoso, J.S. Cardoso, and M.J. Cardoso. Automatic detection of perforators for microsurgical reconstruction. *The Breast*, 50:19–24, 2020.

[180] J. Cohen. *Statistical power analysis for the behavioral sciences*. Hillsdale, N.J.: Erlbaum Associates, 1998.

[181] J. Kanski and B. Bowling. *Clinical ophthalmology: a systematic approach*. Elsevier Health Sciences, 2011.

[182] D. Klonoff and D. Schwartz. An economic analysis of interventions for diabetes. *Diabetes Care*, 23(3):390–404, 2000.

[183] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2980–2988, 2017.

[184] D.P. Kingma and J. Ba. Adam: a method for stochastic optimization. *arXiv:1412.6980*, 2014.

[185] W. Fu, K. Breininger, T. Würfl, N. Ravikumar, R. Schaffert, and A. Maier. Frangi-net: a neural network approach to vessel segmentation. *arXiv:1711.03345*, 2017.

[186] S. Ioffe and C. Szegedy. Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv:1502.03167*, 2015.

[187] F. Yu and V. Koltun. Multi-scale context aggregation by dilated convolutions. *arXiv:1511.07122*, 2015.

[188] G. Hamarneh and P. Jassi. Vascusynth: simulating vascular trees for generating volumetric image data with ground truth segmentation and tree analysis. *Computerized Medical Imaging and Graphics*, 34(8):605–616, 2010.

[189] Y. Zhao, J. Xie, P. Su, Y. Zheng, Y. Liu, J. Cheng, and J. Liu. Retinal artery and vein classification via dominant sets clustering-based vascular topology estimation. In A. Frangi, J. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, volume 11071 of *Lecture Notes in Computer Science*, pages 56–64. Springer, Cham, 2018.

[190] A. BenTaieb and G. Hamarneh. Topology aware fully convolutional networks for histology gland segmentation. In S. Ourselin, L. Joskowicz, M. Sabuncu, G. Unal, and W. Wells, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, volume 9901 of *Lecture Notes in Computer Science*, pages 460–468. Springer, Cham, 2016.

[191] A. Mosinska, P. Marquez-Neila, M. Koziński, and P. Fua. Beyond the pixel-wise loss for topology-aware delineation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3136–3145, 2018.

[192] F. Uslu and A.A. Bharath. A multi-task network to detect junctions in retinal vasculature. In A. Frangi, J. Schnabel, C. Davatzikos, C. Alberola-López, and G. Fichtinger, editors, *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, volume 11071 of *Lecture Notes in Computer Science*, pages 92–100. Springer, Cham, 2018.

[193] D.P. Kingma, T. Salimans, R. Jozefowicz, X. Chen, I. Sutskever, and M. Welling. Improved variational inference with inverse autoregressive flow. In *Advances in Neural Information Processing Systems*, volume 29, pages 4743–4751, 2016.

[194] D.P. Kingma and M Welling. Auto-encoding variational bayes. In *Proceedings of International Conference on Learning Representations*, 2014.

[195] D.I.J. Im, S. Ahn, R. Memisevic, and Y. Bengio. Denoising criterion for variational auto-encoding framework. In *Thirty-First AAAI Conference on Artificial Intelligence*. AAAI Press, 2017.

[196] J.D. Wegner, J.A. Montoya-Zegarra, and K. Schindler. A higher-order crf model for road network extraction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1698–1705, 2013.

[197] S. Shit, J.C. Paetzold, A. Sekuboyina, I. Ezhov, A. Unger, A. Zhylka, J.P.W. Pluim, U. Bauer, and B.H. Menze. cldice – a novel topology-preserving loss function for tubular

structure segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16560–16569, 2021.

[198] F.Y. Shih, Y. Shen, and X. Zhong. Development of deep learning framework for mathematical morphology. *International Journal of Pattern Recognition and Artificial Intelligence*, 33(6):1954024, 2019.

[199] R. Xu, T. Liu, X. Ye, L. Lin, and Y.-W. Chen. Boosting connectivity in retinal vessel segmentation via a recursive semantics-guided network. In A. L. Martel et al., editor, *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, volume 12265 of *Lecture Notes in Computer Science*, pages 786–795. Springer, Cham, 2020.

[200] L.D. Cohen. On active contour models and balloons. *CVGIP: Image Understanding*, 53(2):211–218, 1991.

[201] C. Xu and J.L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on image processing*, 7(3):359–369, 1998.

[202] B. Li and S.T. Acton. Active contour external force using vector field convolution for image segmentation. *IEEE Transactions on Image Processing*, 16(8):2096–2106, 2007.

[203] A.K. Mishra, P.W. Fieguth, and D.A. Clausi. Decoupled active contour (dac) for boundary detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(2):310–324, 2011.

[204] J. Folkman. What is the evidence that tumors are angiogenesis dependent? *JNCI: Journal of the National Cancer Institute*, 82(1):4–7, 1990.

[205] N. Weidner. The importance of tumor angiogenesis: the evidence continues to grow. *American journal of clinical pathology*, 122(5):675–677, 2004.

[206] T. Guan, D. Zhou, and Y. Liu. Accurate segmentation of partially overlapping cervical cells based on dynamic sparse contour searching and gvf snake model. *IEEE journal of biomedical and health informatics*, 19(4):1494–1504, 2015.

[207] K. Li, Z. Lu, W. Liu, and J. Yin. Cytoplasm and nucleus segmentation in cervical smear images using radiating gvf snake. *Pattern Recognition*, 45(4):1255–1264, 2012.

[208] W. Aitfares, A. Herbulot, M. Devy, E.-H. Bouyakhf, and F. Regragui. A novel region-based active contour approach relying on local and global information. In *18th IEEE International Conference on Image Processing (ICIP)*, pages 1029–1032. IEEE, 2011.

[209] D. Chen, J.M. Mirebeau, and L.D. Cohen. Global minimum for a finsler elastica minimal path approach. *International Journal of Computer Vision*, 122(3):458–483, 2017.

[210] S.G. Armato, G. McLennan, L. Bidaut, M.F. McNitt-Gray, C.R. Meyer, A.P. Reeves, L.P. Clarke, et al. Data from lidc-idri. The cancer imaging archive., 2015.

[211] S.G. Armato, G. McLennan, L. Bidaut, M.F. McNitt-Gray, C.R. Meyer, A.P. Reeves, B. Zhao, D.R. Aberle, C.I. Henschke, E.A. Hoffman, et al. The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. *Medical physics*, 38(2):915–931, 2011.

[212] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, et al. The cancer imaging archive (tcia): maintaining and operating a public information repository. *Journal of digital imaging*, 26(6):1045–1057, 2013.

[213] K. Zhang, L. Zhang, H. Song, and W. Zhou. Active contours with selective local or global segmentation: a new formulation and level set method. *Image and Vision Computing*, 28(4):668–676, 2010.

[214] K. Zhang, L. Zhang, K.M. Lam, and D. Zhang. A level set approach to image segmentation with intensity inhomogeneity. *IEEE transactions on cybernetics*, 46(2):546–557, 2016.

[215] K. Zhang. Locally Statistical Active Contour Model code, 2-phase. Available: http://www4.comp.polyu.edu.hk/~cslzhang/LSACM/LSACM.htm (accessed 2018-08-18).

[216] D. Kroon. Snake: active contour - File Exchange - MATLAB Central. Available: https://www.mathworks.com/matlabcentral/fileexchange/28149-snake---active-contour (accessed 2016-09-27).