Segmentation automatique du système vasculaire pour améliorer le système d'aide à la décision basé sur l'IA pour l'artériopathie oblitérante des membres inférieurs

Lisa GUZZI

CENTRE INRIA D'UNIVERSITÉ CÔTE D'AZUR, Équipe EPIONE

Thèse dirigée par Hervé DELINGETTE/co-dirigée par Juliette RAFFORT

Soutenue le 5 Décembre 2025

Présentée en vue de l'obtention du grade de DOCTEUR EN AUTOMATIQUE TRAITEMENT DU SIGNAL ET DES IMAGES d'UNIVERSITÉ CÔTE D'AZUR.

Devant le jury composé de :

Bjoern MENZE	Universität Zürich	Rapporteur
Isabelle BLOCH	Sorbonne Université	Rapporteure
Antoine VACAVANT	Institut Pascal, Université Clermont Auvergne	Examinateur
Hervé DELINGETTE	Centre INRIA d'Université Côte d'Azur	Directeur de thèse
Juliette RAFFORT	CHU de Nice	Co-directrice de thèse
Maria A. ZULUAGA	EURECOM, Sophia Antipolis	Co-encadrante

SEGMENTATION AUTOMATIQUE DU SYSTÈME VASCULAIRE POUR AMÉLIORER LE SYSTÈME D'AIDE À LA DÉCISION BASÉ SUR L'IA POUR L'ARTÉRIOPATHIE OBLITÉRANTE DES MEMBRES INFÉRIEURS

Automatic segmentation of the vascular system to enhance AI-based decision support system for peripheral artery disease

Lisa GUZZI

Jury :

Rapporteurs:
Bjoern MENZE, Professeur, Universität Zürich
Isabelle BLOCH, Professeure des universités, Sorbonne Université

Examinateur:
Antoine VACAVANT, Professeur des universités, Institut Pascal, Université Clermont Auvergne

Directeur de thèse:
Hervé DELINGETTE, Directeur de recherche, Centre INRIA d'Université Côte d'Azur

Co-Directrice de thèse:
Juliette RAFFORT, Praticienne hospitalière en biologie médicale, CHU de Nice

Co-encadrante de thèse:
Maria A. ZULUAGA, Professeure, EURECOM, Sophia Antipolis

Résumé

L'artériopathie oblitérante des membres inférieurs (AOMI) touche plus de 230 millions de personnes dans le monde. Généralement causée par l'athérosclérose, elle se caractérise par un rétrécissement ou une occlusion des artères des membres inférieurs. En raison de sa nature souvent asymptomatique, l'AOMI est fréquemment diagnostiquée à un stade avancé, ce qui augmente le risque de complications cardiovasculaires et d'amputation. Ainsi, l'AOMI est associée à une forte mortalité et morbidité, représentant un problème majeur de santé publique. L'angiographie par tomodensitométrie (CTA) est couramment utilisée pour évaluer les lésions artérielles et l'anatomie vasculaire, et pour guider les stratégies de revascularisation. Cependant, l'analyse manuelle des CTA est chronophage et dépendante de l'opérateur, ce qui souligne le besoin d'outils automatisés pour aider la prise de décision clinique. Cette thèse vise à développer un système basé sur l'intelligence artificielle pour l'évaluation complète et automatisée de l'AOMI, facilitant la planification de traitements personnalisés. Le principal défi abordé concerne la segmentation des petites artères tortueuses avec de fréquentes occlusions, ainsi que des plaques de calcification et des stents artériels. La segmentation de ces structures dans les membres inférieurs doit permettre l'extraction de caractéristiques anatomiques et pathologiques précises afin de fournir aux cliniciens les informations nécessaires pour guider la planification préopératoire de l'AOMI. Pour relever ce défi de segmentation, nous proposons deux contributions méthodologiques et une application clinique : (1) SoftMorph est une méthode pour convertir toute opération morphologique binaire en un équivalent probabiliste dérivable, permettant son intégration dans les réseaux de neurones soit comme couche finale, soit au sein de la fonction de coût. Les filtres probabilistes sont définis comme l'espérance du filtre binaire sur l'ensemble des configurations binaires possibles et exprimés sous la forme d'un polynôme multilinéaire dérivé de sa table de vérité. Pour les cas complexes, des approximations sont obtenues via des opérateurs quasi-probabilistes en appliquant divers opérateurs de logique floue à l'expression booléenne définissant l'opération morphologique, tout en préservant la complexité du filtre original. Les expériences ont démontré des améliorations en termes de préservation topologique pour la segmentation des structures tubulaires. (2) Regional Hausdorff Distance losses : une famille de fonctions de coût développée pour améliorer la précision des bords de structures segmentées, particulièrement pertinente pour certaines lésions pathologiques. La méthode repose sur une fonction de distance dérivable basée sur un opérateur d'érosion, permettant un calcul dérivable des distances régionales maximales, modifiées et moyennes de Hausdorff. Ces fonctions de coût ont atteint de bonnes performances sans nécessiter de fonctions auxiliaires pour l'entraînement des réseaux de segmentation pour plusieurs modalités. (3) Enfin, ces innovations sont appliquées dans un contexte clinique pour la segmentation des artères des membres inférieurs, des stents et des plaques de calcification. PADSET, une base de données interne de CTA de patients atteints d'AOMI, a été constituée et annotée afin de fournir des masques de référence

pour chaque structure. Ensuite, des méthodes de segmentation automatique basées sur l'apprentissage profond sont explorées, intégrant les deux contributions techniques précédentes pour obtenir une segmentation automatique de haute performance sur le jeu de données PADSET. De plus, l'approche permet l'identification automatique des principales branches artérielles afin d'extraire la localisation précise des caractéristiques cliniquement pertinentes. L'outil automatique réduit la variabilité inter-observateurs et contribue à aider la planification préchirurgicale dans le traitement de l'AOMI.

Mots-clés : Segmentation d'images médicales, Apprentissage profond, Artères des membres inférieurs, Artériopathie oblitérante des membres inférieurs, Angiographie par tomodensitométrie, Calcifications, Stents.

Abstract

Peripheral artery disease (PAD) affects over 230 million people worldwide. Generally caused by atherosclerosis, it is characterized by the narrowing or occlusion of the arteries in the lower limbs. Owing to its frequent asymptomatic presentation, PAD is often diagnosed at advanced stages, increasing the risk of cardiovascular complications and amputation. As such, PAD is associated with high mortality and morbidity, representing a major public health concern. Computed Tomography Angiography (CTA) is commonly used to assess arterial lesions and anatomy, guiding revascularization strategies. However, current manual analysis of CTA is time-consuming and operator-dependent, underscoring the need for automated tools to support clinical decision-making. This thesis aims to develop an AI-based system for the comprehensive and automated assessment of PAD, facilitating personalized treatment planning. The main challenge addressed is the segmentation of small, tortuous arteries with frequent occlusions, as well as calcification plaques and arterial stents. The segmentation of these structures in the lower limbs should enable the extraction of precise anatomical and pathological features to provide clinicians with the necessary information to guide the preoperative planning for PAD. To overcome this segmentation challenge, we propose two methodological contributions and one clinical application. (1) SoftMorph is a framework that converts any binary morphological operation into a differentiable probabilistic counterpart, enabling its integration into neural networks either as a final layer or within the loss function. Probabilistic filters are defined as the expectation of the binary filter over all possible binary configurations and expressed as a multi-linear polynomial derived from its truth table. For intractable cases, approximations are obtained via quasi-probabilistic operators by applying various fuzzy logic operators to convert the Boolean expression defining the morphological operation, preserving the original filter's complexity. Experiments demonstrated improvements in topological preservation for the segmentation of tubular structures. (2) Regional Hausdorff Distance losses are developed, a family of loss functions to improve boundary precision in segmented structures, particularly relevant in pathological contexts. The method relies on a fully differentiable erosion-based distance function to produce differentiable computation of the maximum, modified, and averaged regional Hausdorff Distances. These loss functions achieved state-of-the-art performance without requiring any auxiliary losses for the training of segmentation networks across multiple modalities. (3) Finally, these innovations are applied in a real clinical context for the segmentation of lower-limb arteries, stents and calcification plaques. PADSET, an

in-house CTA dataset of PAD patients, is curated and annotated to provide ground truth masks of each structure. Then, deep-learning-based automatic segmentation methods are explored, along with the application of the two previous technical contributions to achieve high-performance automatic segmentation on the PADSET dataset. Additionally, the approach involves automatically identifying key arterial branches to extract the precise locations of clinically relevant features. The automatic tool reduces inter-observer variability and supports pre-surgical planning for the treatment of PAD.

Keywords: Medical image segmentation, Deep learning, Lower limb arteries, Peripheral artery disease, Computed tomography angiography, Calcifications, Stents.

Acknowledgement

I would like to express my deepest gratitude to my three supervisors, to whom I owe a great deal. I thank Juliette, co-director of this thesis, for initiating and designing the PhD project and for her constant availability throughout these years. I sincerely thank Maria, co-supervisor of this thesis, for welcoming me into a second laboratory and research team. This dual affiliation greatly enriched my PhD experience, both scientifically and personally, and her guidance, advice, and support were particularly valuable during key moments of this work. I am especially grateful to Hervé, PhD director, for his supervision throughout this thesis. His trust, encouragement, and positive attitude provided the ideal balance between guidance and independence, allowing me to grow in confidence and maturity as a researcher.

I extend my thanks to Fabien for bringing a strong clinical perspective to this work, giving it a concrete and applied dimension. I also thank the medical doctors Gilles and Andrea for their valuable help and expertise in the clinical aspects of this thesis.

I express my sincere gratitude to the members of my thesis jury, Bjoern Menze, Isabelle Bloch, and Antoine Vacavant, for accepting to evaluate this work. I thank them for the careful reading of the manuscript, the detailed reports, and the insightful questions and discussions during the defense.

I would also like to thank my colleagues and friends from both of my research teams, as having two laboratories was a great privilege during my PhD. At Eurecom, I thank Daniele, Luisa, Francesco, Eleonora, Valentin, and especially Hava and Vincenzo for their support, friendship, and the many moments shared both inside and outside the lab. At EPIONE, I thank the permanent staff, all the students and engineers for the stimulating work environment. I would like to thank the many colleagues I met throughout these years for their help and advice, including Jairo, Olivier, Peppe, Amel, Nicolas, Maëlis, Rafael, Hari, Guillaume, and many others who contributed, directly or indirectly, to this experience. I am particularly grateful to Tom, with whom I shared the entire PhD journey, for his constant support and kindness. Special thanks go to Camilla for her friendship

and support during the final year of the PhD, and to my office mate Seb for his daily support, advice, and generosity.

I am grateful to my friends outside academia for their unwavering support over the years, in particular Aurélia, Laurène, Célie, and Eliott for attending the defense, as well as to Tania, Cindie, and Mathilde for their long-standing encouragement and support. I also thank all my other friends for their presence and shared moments throughout this journey.

I am deeply grateful to my family for their patience, unconditional love, and constant support, including my siblings, parents, and step-parents. I owe my greatest thanks to my mother, whose sacrifices, support, and unwavering belief made this journey possible.

Finally, I thank Riccardo for his important contribution, especially during this final year, for his support, encouragement, many tiramisus, and help throughout this work.

I dedicate this work to the memory of my grandparents, who always believed in me.

Financial Support

This work has been supported by the French government, through the 3IA Côte d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002 and ANR-23-IACL-0001. The authors are grateful to the OPAL infrastructure from Université Côte d'Azur. Experiments presented in this paper were carried out using the Grid'5000 testbed, supported by a scientific interest group hosted by Inria and including CNRS, RENATER and several Universities as well as other organizations (see https://www.grid5000.fr).

Contents

1 Introduction

Contents

1.1 Thesis objectives

Peripheral artery disease (PAD) is a prevalent cardiovascular condition affecting an estimated quarter billion people worldwide [Kullo, 2016; Nordanstig, 2024; Song, 2019]. It is characterized by the narrowing (stenosis) or complete occlusion (thrombosis) of the arteries below the hips, leading to reduced blood flow and oxygen delivery to the lower limbs (ischemia). The obstruction is generally due to the formation of plaques related to atherosclerosis that thicken arterial walls and reduce lumen diameter [Kullo, 2016; Song, 2019].

PAD represents a major public health concern as the disease is highly prevalent and is associated with high rates of morbidity and mortality [Gerhard-Herman, 2017; Hirsch, 2001; Hirsch, 2006]. Despite its elevated risks of cardiovascular mortality and amputation, PAD remains underdiagnosed and underestimated [Criqui, 2015]. It is often diagnosed at an advanced stage of the disease due to low patient awareness and partly because early stages are often asymptomatic or manifest with atypical symptoms [McDermott, 2015]. In addition, several studies have suggested that many patients may be undertreated, pointing to the need to improve the use of evidence-based recommended therapies in patients with PAD [Lee, 2019]. It is therefore crucial to detect PAD at an early stage, establish an accurate diagnosis, and implement appropriate treatment to effectively manage the disease and prevent complications.

Diagnosis typically combines a clinical examination, functional tests, and the identification of arterial lesions with medical imaging. Among imaging modalities, computed tomography angiography (CTA) is widely used as vascular surgeons rely on precise 3D

visualization of arterial anatomy and lesion characteristics to plan interventions [Committee, 2015]. The extent and exact location of occlusions, calcifications, and prior stents, as well as the overall arterial anatomy of the patient, strongly influence the revascularization strategy [Nordanstig, 2024]. However, manual analysis of lower extremity CTA is time-consuming, operator-dependent, and prone to variability. Thus, the automatic segmentation of the lower-limb arteries, along with the detection of calcified plaque and stents, could significantly improve pre-surgical planning. The automatic measurements could help retrieve useful information, for example, to calculate standardized staging scores (e.g., GLASS [Wijnand, 2021], TASCII score [Norgren, 2007]) from the imaging, supporting consistent and personalized treatment decisions.

However, PAD presents unique segmentation challenges: arteries are often narrowed or occluded, distal branches can be small and tortuous, and heavy calcifications and surgical stents cause imaging artifacts, degrading conventional segmentation performance. Vessel segmentation algorithms may falsely interrupt vessels at occlusions or bleed into high-contrast plaque. Clinically, missing a side branch or mis-measuring a lesion can misguide revascularization decisions. The segmentation of clinical anomalies such as calcifications and stents is also challenging due to their very small size and heterogeneous presentation.

Deep learning (DL), and more importantly, convolutional neural networks (CNNs) have transformed 3D medical image segmentation [Anaya-Isaza, 2021; Sermesant, 2021]. It has been used successfully to segment different types of arteries and calcification plaques [Hilbert, 2022; Bagheri Rajeoni, 2023; Lareyre, 2021; Isensee, 2021], for example, to segment carotid arteries and calcified plaques on CTA [Zhu, 2022]. Zhou et al. [Ouyang, 2024] additionally proposed a general CTA model for multiple anatomical structures and lesions, and outperformed prior methods on all evaluated structures, notably including coronary arteries, aorta, and lower limb arteries. While these studies show the potential of CNNs to segment vascular structures on CTA, most focus on major vessels such as the aorta, coronaries, carotids or on classification tasks, but none of them fully segment the total lower-limb arterial tree, typically stopping at the iliac or femoral level. To further enhance the segmentation of fine tubular structures, loss functions and architectures have been developed to improve, for example, the topology of the segmentation, particularly relevant for vessels [Shit, 2021; Zhang, 2023; Weng, 2023; Stucki, 2023; Qiu, 2023; Clough, 2020]. To our knowledge, these advances have yet to be fully exploited for the entire lower limb vasculature in PAD patients.

The main goal of this thesis is to leverage recent advances in DL-based segmentation to develop an automatic tool for the segmentation of the lower-limb arteries from CTA scans. The anatomical and clinical features extracted from the segmentation will support clinicians by accelerating pre-operative planning, and reinforcing decision-making con-

fidence in selecting the appropriate revascularization strategy for PAD, and ultimately, improving patient outcomes.

To meet this aim, the thesis is organized into three main objectives, progressing from methodological development to clinical application as follows:

- **Objective 1:** The segmentation of the target structures in this thesis involves addressing two distinct challenges. First, the arteries present as tortuous tubular structures that require precise delineation to capture their complex geometry. Second, the calcifications and stents appear as small, disconnected, and heterogeneous components, which add complexity due to their size, shape variability, and image appearance. Therefore, the initial objective is to develop and refine methodological innovations in CNN-based segmentation that specifically target these challenges, improving accuracy, robustness, and generalization for both tubular and discrete structures.

- **Objective 2:** The second objective focuses on translating these methodological advances into practical application within a real clinical setting. This involves applying the proposed segmentation methods to CTA datasets from patients with PAD, which often present with diverse pathological characteristics. Specifically, the targeted structures include the entire lower-limb arterial tree from the abdominal aorta to the fibular and posterior tibial arteries, together with the identification of each distinct branch, calcification plaques and endoprosthetic stents. This step aims to demonstrate the efficacy and adaptability of the segmentation framework in handling clinically relevant, complex vascular anatomies and lesions.

- **Objective 3:** Finally, the third objective is to utilize the fully automated segmentation outputs to extract detailed anatomical and clinical features that are critical for personalized pre-surgical planning. These features will enable clinicians to make informed decisions about the type and strategy of revascularization for PAD patients, ultimately leading to improved treatment outcomes and patient care.

1.2 Thesis outline and Contributions

Table 1.1.: Overview of the main contributions and their relation to the thesis objectives.

	Chapter 4	Chapter 5	Chapter 6	Chapter 7
O1: Methodological advances in segmentation	✓	✓		
O2: Application to clinical PAD dataset			✓	✓
O3: Feature extraction and clinical analysis				✓

Table 1.1 indicates how each contribution chapter addresses the corresponding objectives previously defined.

The foundations of this work are presented in **Part I**:

Chapter 2. Peripheral Artery Disease. This chapter presents the clinical background of PAD, with an emphasis on aspects directly relevant to this thesis. It reviews the anatomy of the lower-limb arterial system and the characteristics of its main pathological manifestations. It highlights the role of CTA in diagnosis and surgical planning in how imaging findings influence the choice of revascularization strategy. The chapter also provides essential reference material for understanding the vascular structures, lesions, and implanted devices addressed in later sections of this work.

Chapter 3. Deep Learning-based Segmentation for Peripheral Artery Disease. This chapter reviews the current state-of-the-art in DL for medical image segmentation, with a focus on techniques and applications relevant to PAD. After a brief introduction to CNNs in the context of segmentation, we survey existing work on the segmentation of lower-limb arteries, calcification plaques, and endoprosthetic stents, including approaches targeting similar small or topologically complex structures in other vascular territories. Particular attention is given to methods designed for topological preservation and the accurate delineation of fine or heterogeneous components. The chapter concludes by identifying key limitations and challenges in the existing literature, thereby motivating the methodological developments presented in this thesis.

With the theoretical and methodological foundations established, **Part II** details the methodological contributions developed in this work for image segmentation:

Chapter 4. SoftMorph: Differentiable Probablistic Morphological Operators for Image Segmentation. This chapter introduces SoftMorph, a novel framework to convert any binary morphological operation into a differentiable probabilistic equivalent that can be integrated into CNNs, either as a final layer or within the loss function. Probabilistic morphological filters are defined as the expectation of the binary filter over the probability of generating each possible binary configuration and expressed as a multi-linear polynomial deduced from the truth table. For intractable truth tables, they are approximated by a family of quasi-probabilistic operators, using various fuzzy logics to directly convert the Boolean expression defining the morphological operation into a differentiable expression, maintaining the complexity of the original binary filter. Its application on multiple segmentation tasks leads to state-of-the-art performance in terms of topological preservation for tubular structures.

Chapter 5. Regional Hausdorff Distance Losses for Medical Image Segmentation. This chapter addresses the limitations of conventional overlap-based metrics such as the Dice coefficient by focusing on the Hausdorff Distance, which provides a more sensitive measure of boundary errors. Despite its clinical relevance, directly optimizing the HD within CNNs remains challenging due to the non-differentiability of standard distance

transform algorithms. We propose a novel family of regional Hausdorff Distance loss functions, specifically the maximum, modified, and average variants, that overcome these limitations without the need for auxiliary losses. Our approach is based on a fully differentiable erosion-based distance function implemented using the previously presented Softmorph framework, enabling accurate and stable approximation of signed, unsigned, or positive distance maps. Validation on multiple public medical image segmentation datasets demonstrates that these loss functions achieve competitive improvements in distance-based metrics while maintaining high overlap accuracy.

The technical innovations developed are applied and evaluated within a practical clinical context for PAD in **Part III**:

Chapter 6. PADSET: A Private CTA Dataset for Peripheral Artery Disease. This chapter introduces PADSET, a dedicated dataset curated specifically for this thesis, consisting of 196 lower-limb CTA scans collected at the Hospital of Nice. It details the dataset's key characteristics and the annotation protocol developed to generate high-quality ground truth masks for arteries, calcification plaques, stents, and individual arterial branches. The chapter also discusses the assessment of inter-annotator variability and the strategies implemented to enhance annotation consistency and reliability. Annotation was performed in close collaboration with clinical experts, ensuring both medical accuracy and technical rigor.

Chapter 7. Automatic Segmentation of lower-limb Arteries on CTA for Pre-surgical Planning of Peripheral Artery Disease. This chapter presents the application of the methodological innovations developed earlier to the curated PADSET dataset. It details the implementation and evaluation of automatic segmentation methods for the comprehensive delineation of the entire lower-limb arterial tree, the identification of each arterial branch, as well as the detection of calcification plaques and endoprosthetic stents. Comparative analyses with established segmentation techniques are provided to highlight the improvements achieved. Furthermore, the chapter describes the extraction of clinically relevant features from the segmentation outputs, accompanied by statistical analysis to characterize the dataset and support clinical insights.

Finally, this thesis is concluded in Chapter 8:

Chapter 8. Conclusion. This chapter summarizes the key findings of the thesis, discusses their implications, and outlines future perspectives from both methodological and clinical viewpoints. It also highlights potential real-world applications and the broader impact of this work.

1.3 Publications

The contributions of this manuscript led to the following publications and submissions in conferences and peer-reviewed journals.

1.3.1 International Conference Proceedings

- *Differentiable Soft Morphological Filters for Medical Image Segmentation* [Guzzi, 2024] Guzzi L, Zuluaga M A, Lareyre F, Di Lorenzo G, Goffart S, Chierici A, Raffort J, Delingette H. *The 27th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2024, Marrakech, Morroco.*

- *Regional Hausdorff Distance Losses for Medical Image Segmentation* [Guzzi, 2025b] Guzzi L, Zuluaga M A, Taiello R, Lareyre F, Di Lorenzo G, Goffart S, Chierici A, Raffort J, Delingette H. *The 16th International Workshop on Machine Learning in Medical Imaging (MLMI 2025), held in conjunction with MICCAI, 2025, Daejeon, Republic of Korea.*

- *Automatic Segmentation of lower-limb Arteries on CTA for Pre-surgical Planning of Peripheral Artery Disease* [Guzzi, 2025a] Guzzi L, Zuluaga M A, Lareyre F, Di Lorenzo G, Goffart S, Chierici A, Raffort J, Delingette H. *The Fourth Workshop on Applications of Medical Artificial Intelligence (AMAI 2025), held in conjunction with MICCAI, 2025, Daejeon, Republic of Korea.*

1.3.2 National Conference Proceedings

- *Application du modèle nnUNet pour la segmentation automatique du réseau artériel des membres inférieurs dans le cadre de l'artériopathie oblitérante des membres inférieurs* Guzzi L, Di Lorenzo G, Lareyre F, Goffart S, Zuluaga M A, Delingette H, Raffort J. Accepted for a poster presentation at Colloque Français d'Intelligence Artificielle en Imagerie Biomédicale (IABM), 2024, Grenoble, France.

- *SoftMorph: Differentiable Morphological Filters for Medical Image Segmentation* Guzzi L, Zuluaga M A, Lareyre F, Di Lorenzo G, Goffart S, Chierici A, Raffort J, Delingette H. Accepted for a poster presentation at Sophia Summit 2024, Sophia Antipolis, France.

- *Des Filtres Morphologiques Dérivables pour la Segmentation d'Images Médicales* Guzzi L, Zuluaga M A, Lareyre F, Di Lorenzo G, Goffart S, Chierici A, Raffort J,

Delingette H. Accepted for a poster presentation at Colloque Français d'Intelligence Artificielle en Imagerie Biomédicale (IABM), 2025, Nice, France.

- *Stratégies d'apprentissage pour la prédiction du risque d'amputation chez les patients atteints d'artériopathie oblitérante des membres inférieurs* Goffart S, Hart O, Lareyre F, Guzzi L, Khashram M, Delingette H, Raffort J. Accepted for a poster presentation at Colloque Français d'Intelligence Artificielle en Imagerie Biomédicale (IABM), 2025, Nice, France.

1.3.3 Journals

- *Artificial Intelligence in Vascular Surgical Decision-Making* [Lareyre, 2023] Lareyre F, Yeung KK, Guzzi L, Di Lorenzo G, Chaudhuri L, Behrendt CA, Spanos K, Raffort J. Seminars in Vascular Surgery. 2023.

- *Artificial Intelligence Techniques for Prognostic and Diagnostic Assessments in Peripheral Artery Disease: A Scoping Review* [Goffart, 2025b] Goffart S, Delingette H, Chierici A, Guzzi L, Nasr B, Lareyre F, Raffort J. Angiology. 2025.

- *Artificial Intelligence to Enhance Future Clinical Trials in Vascular Surgery* [Goffart, 2025a] Goffart S, Chierici A, Guzzi L, Delingette H, Alouane A, Lareyre F, Raffort J. Annals of Vascular Surgery. 2025.

- *Deep Learning Strategies for Predicting Amputation Risk in Patients With Peripheral Artery Disease* [Goffart, 2025c] Goffart S, Hart O, Lareyre F, Guzzi L, Yeung KK, Delingette H, Khashram M, Raffort-Lareyre J. EJVES. 2025.

- *Imaging characterization of peripheral artery disease: a scoping review on current classifications and new insights brought by Artificial Intelligence* [Lareyre, 2025] Lareyre F, Guzzi L, Nasr B, Alouane A, Goffart S, Chierici A, Delingette H, Raffort J. In : EJVES Vascular Forum. 2025.

- *Applications of artificial intelligence in liver cancer: A scoping review* [Chierici, 2025] Chierici A, Lareyre F, Iannelli A, Salucki B, Goffart S, Guzzi L, Poggi E, Delingette H, Raffort J. Artificial Intelligence in Medicine. 2025.

- *SoftMorph: Differentiable Soft Morphological Operators for Image Analysis* Guzzi L, Zuluaga M A, Lareyre F, Di Lorenzo G, Goffart S, Chierici A, Raffort J, Delingette H. Under Review at Medical Image Analysis.

Part I

Pre-requisites

2 Peripheral Artery Disease

Contents

This chapter provides the medical and anatomical background necessary to understand the clinical challenges addressed in this thesis. It focuses on aspects of peripheral arterial disease that directly influence image analysis requirements, with particular emphasis on CTA as a diagnostic and surgical planning tool.

We first present the organization of the lower-limb arterial system and describe the principal vascular territories that will be referenced throughout this work. We then examine the main pathological processes affecting these vessels, including their clinical manifestations, associated risk factors, and epidemiological impact. Special attention is given to features that can be visualized in CTA, such as calcified atherosclerotic plaques and implanted endovascular devices (stents), as these structures play a central role in the segmentation tasks developed later.

Finally, the chapter discusses how imaging findings influence the choice of revascularization strategy and surgical approach, highlighting the specific interpretation challenges

posed by complex vascular anatomies, severe calcification, and the presence of stents. This context serves both as a reference for understanding the dataset and as a clinical framework for evaluating the practical utility of the proposed segmentation methods.

2.1 Anatomy of the Lower-Limb Arterial System

The lower limbs can be studied below the hips. The arterial anatomy in the lower limbs is illustrated in Fig. 2.1.

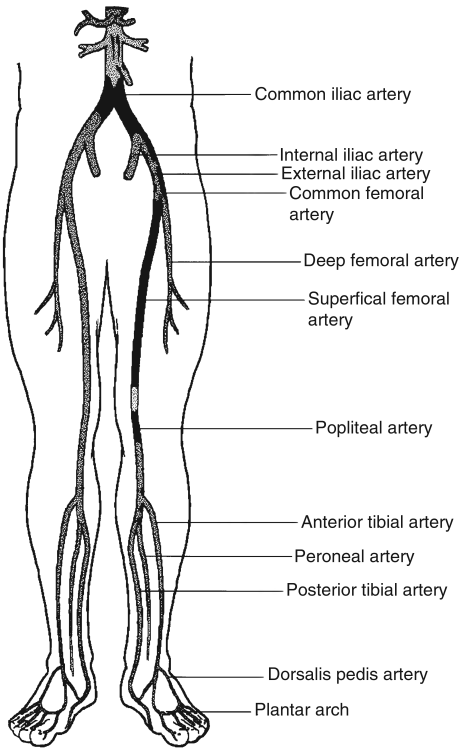


Figure 2.1.: Common anatomy of lower limb arteries (illustration from [AbuRahma, 2013])

The arterial supply of the lower limbs originates from the abdominal aorta, which divides into the common iliac arteries in each leg at the level of the L4/L5 vertebrae. Each common iliac artery branches into an internal iliac artery, which primarily supplies the pelvis, and an external iliac artery, which continues into the lower limb as the common femoral artery after passing beneath the inguinal ligament. The common femoral artery divides into the superficial femoral artery, which supplies the leg through the popliteal and tibial branches, and the deep femoral artery, which supplies the hip and thigh through the circumflex and perforating branches. Distally, the popliteal artery leads to the anterior tibial artery and the tibioperoneal trunk. The latter quickly divides into the posterior tibial artery, which supplies the posterior compartment and the plantar surface of the foot, and the fibular artery, which supplies the lateral leg. Together, these vessels form

an extensive network that ensures blood flow from the pelvis to the foot [Qazi, 2022; Netter, 2022].

In cases of chronic arterial obstruction, collateral pathways can develop to maintain distal blood flow, and their presence, number, and effectiveness vary between individuals as an adaptive response [Macchi, 1996; Hardman, 2011; Kruse, 2017].

2.2 Disease Characterization

2.2.1 Pathophysiology

Peripheral artery disease is defined as the partial or complete obstruction of arteries from the distal aorta to the foot, resulting in impaired blood flow to the lower extremities [Nordanstig, 2024], illustrated in Fig. 2.2.

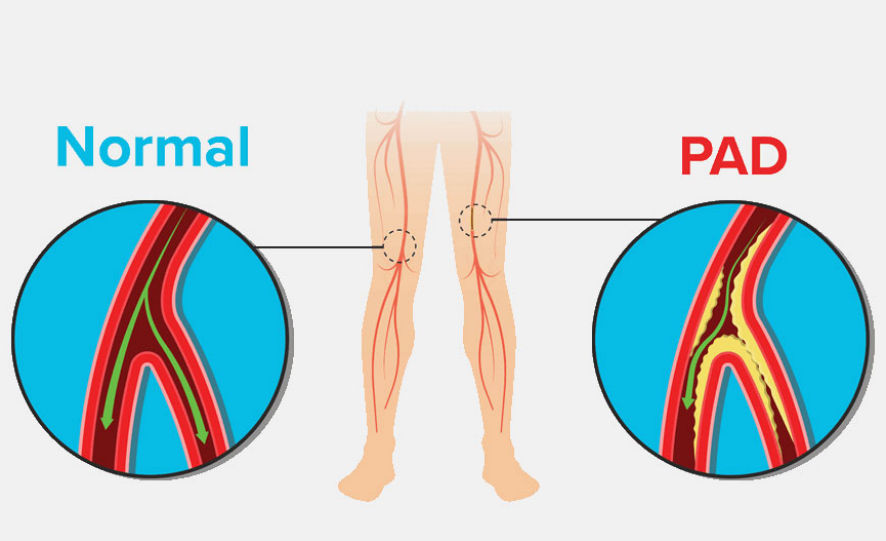


Figure 2.2.: Illustration of a normal versus a diseased artery caused by atherosclerosis (Source: University of Maryland Charles Regional Medical Center Blog, accessed October 7, 2025)

The obstruction is defined as a stenosis when the arterial lumen is narrowed and a thrombosis when the artery is completely occluded. Their manifestation is caused by atherosclerosis in 95% of cases [Frank, 2019; Hirsch, 2006; Gordon, 1972].

Atherosclerosis is a chronic inflammatory disease of the arteries driven by lipid accumulation and immune responses. It usually begins with the damage or dysfunction of the endothelium (the inner lining of blood vessels) that allows low-density lipoproteins and other lipids to enter and accumulate in the innermost layer (intima) of arteries. Inflammatory cells such as monocytes and macrophages are recruited, ingest lipids, and

become foam cells, forming fatty streaks. Over time, smooth muscle cells migrate from the media into the intima, proliferate, and secrete extracellular matrix molecules, which thickens the arterial wall and forms a fibrous cap over a lipid core. This thickening is the general cause of stenosis.

Occasionally, the fibrous cap becomes unstable and may rupture. When plaque ruptures, lipid core contents and calcium debris may be exposed to the blood in the lumen, triggering platelet aggregation, activation of coagulation cascades, and formation of a thrombus, which can partially or completely block the artery.

Calcification deposits sometimes build up in the plaque, making it stiffer. It occurs when inflammatory and metabolic changes within the plaque trigger vascular smooth muscle cells (VSMCs) to undergo a phenotypic switch, adopting osteoblast-like properties. These transformed VSMCs release matrix vesicles and bone-related proteins, which provide a scaffold for calcium phosphate crystal deposition. In parallel, cell death within the plaque releases apoptotic bodies that serve as nucleation sites for mineral deposition. Over time, these processes lead to patchy or extensive calcification within the intima. While some calcifications stabilize plaques by reinforcing the fibrous cap, others can increase mechanical stress and make rupture more likely. Thus, calcification represents both an adaptive response and a risk factor for plaque instability and thrombosis [Björkegren, 2022; Libby, 2021; Falk, 2006].

2.2.2 Clinical presentation

PAD has a broad spectrum of clinical manifestations. Many patients remain asymptomatic, while the most common symptomatic form is intermittent claudication, typically described as exertional leg pain that is relieved by rest [Criqui, 1996; Gardner, 2007]. However, symptoms are often atypical or masked by comorbidities such as diabetes, neuropathy, or limited mobility, which can lead to underdiagnosis [Santoro, 2018; Serhal, 2018]. In more advanced stages, PAD may manifest as chronic limb-threatening ischemia (CLTI) with rest pain, non-healing wounds, ulceration, or gangrene, particularly in patients with diabetes (“diabetic foot”). Acute presentations can also occur due to sudden arterial thrombosis or embolism, resulting in acute limb ischemia. This variability highlights the importance of careful clinical evaluation and risk factor assessment in patients at risk of PAD [Nordanstig, 2024].

2.2.3 Complications

If left untreated or in advanced stages, PAD can lead to serious complications. Local consequences include poor wound healing, tissue necrosis, and eventual limb loss re-

quiring amputation. Systemic risks are also significant, as PAD patients have a higher incidence of major cardiovascular events such as myocardial infarction and stroke. In addition, revascularization procedures carry their own risks, including bleeding, vessel perforation, infection, and restenosis. Together, these complications make PAD not only a limb-threatening but also a life-threatening disease [Nordanstig, 2024].

2.3 Epidemiology and risk factors

2.3.1 Prevalence and incidence worldwide

PAD affects hundreds of millions of people globally and its prevalence is rising [Kullo, 2016; Nordanstig, 2024; Song, 2019]. In 2015, an estimated 237 million adults aged 25 and older were affected worldwide, representing a 17% increase compared with 2010, with the largest relative increases observed in low and middle-income countries [Fowkes, 2013; Song, 2019]. Prevalence estimates vary due to differences in study populations, age, sex, ethnicity, and diagnostic methods, particularly the use of ankle-brachial index (ABI) thresholds. The prevalence of the disease increases with age, reaching more than 20% beyond 70 years. PAD is the local expression of a systemic disease whose prognosis is influenced by cardiac and cerebrovascular complications (mortality at the stage of intermittent claudication: 15% at 5 years; mortality at the stage of critical permanent ischemia: 25% at 1 year) [Nordanstig, 2024]. In general population studies, the prevalence of asymptomatic PAD often exceeds that of symptomatic disease: for example, in adults aged 50–90 years, 18% had PAD by ABI criteria, and more than 60% of these were asymptomatic [Cimminiello, 2010; Sigvant, 2007]. Risk factor prevalence, demographic changes, and regional differences in awareness and screening contribute to heterogeneity.

2.3.2 Key risk factors

Some risk factors increase the chances of suffering from PAD.

Lifestyle factors. Tobacco smoking is associated with increased risk of requiring revascularization, of chronic limb-threatening ischemia (CLTI), and amputation [Reitsma, 2017], and is the leading preventable cause of disease overall [Young, 2019]. Smoking cessation is therefore a primary behavioral intervention for PAD patients. Additionally, regular physical activity is recommended, as sedentary behavior increases PAD risk [Nordanstig, 2024].

Metabolic disease. Obesity and diabetes mellitus significantly increase the risk of PAD by promoting atherosclerosis and vascular dysfunction. Diabetes, in particular, is strongly associated with distal small-vessel disease, higher rates of foot ulcers, and progression to CLTI. Dyslipidaemia and hypertension are also major contributors, accelerating plaque formation and arterial stiffening [Nordanstig, 2024].

Intrinsic risk factors. Age is a major non-modifiable risk factor, with PAD prevalence increasing sharply after 50 years. Male sex is slightly more likely to develop symptomatic PAD in some cohorts, though differences are less pronounced for asymptomatic disease. Additional contributors include chronic kidney disease, systemic inflammation, and a history of cardiovascular disease [Nordanstig, 2024].

2.4 Diagnosis of PAD

2.4.1 Clinical assessment

The diagnosis of PAD begins with a thorough clinical assessment, focusing on patient-reported symptoms and physical examination. Notably, the Ankle-Brachial Index (ABI) is a simple, non-invasive diagnostic test for PAD that compares the systolic blood pressure at the ankle with that at the brachial artery. An ABI ≤ 0.90 is generally considered indicative of PAD. An ABI ≥ 1.40 suggests an arterial impressibility and high cardiovascular risk. In that case, the ankle pressure can not be used for the diagnosis and must be measured on the first toe instead. At the stage of intermittent claudication in the absence of calcifications, the ankle blood pressure is > 50 mm Hg and the ABI $<$0.90 [Høyer, 2013; Mills Sr, 2014].

The treadmill test is used for patients with suspected intermittent claudication and a normal ABI at rest. The ABI is measured before and after exercise, and the maximum walking distance is assessed during a graded treadmill test as the distance covered until the patient needed to stop due to claudication symptoms [Gardner, 1991; Birkett, 2021].

2.4.2 Imaging Modalities

Imaging is crucial for planning a revascularization procedure to determine the presence, location and extent of atherosclerotic lesions [Nordanstig, 2024]:

- **Duplex Ultrasound (DUS):** Combines B-mode imaging with Doppler to visualize vessel morphology and blood flow. Although it benefits from low costs compared

to other imaging modalities, it suffers from large inter-observer variability and reduced accuracy in obesity, excessive bowel gas and calcified lesions [Nordanstig, 2024; Collins, 2007].

- **Digital Subtraction Angiography (DSA):** Formerly the gold-standard technique providing high-resolution vascular images [Takahashi, 1984]. It has now been replaced by CTA and MRA, which are less invasive [Nordanstig, 2024].

- **Computed Tomography Angiography (CTA):** Uses X-rays and iodine-based contrast agent to generate high-resolution cross-sectional images of the vasculature. The sensitivity and specificity for the detection of arterial lesions are over 90% compared to the DSA reference standard [Nordanstig, 2024; Collins, 2007]. However, drawbacks include radiation exposure and potential misinterpretation of heavily calcified lesions.

- **Magnetic Resonance Angiography (MRA):** Uses magnetic fields and radio waves to create detailed images of blood vessels without ionizing radiation. The sensitivity and specificity for the detection of arterial lesions are over 90% compared to the DSA reference standard [Nordanstig, 2024; Collins, 2007]. However, this imaging modality does not provide information on calcification burden, and it is difficult to assess the lumen within metal stents. It also costs more than CTA and is contraindicated in patients unsuitable for MRI.

Therefore, CTA is particularly valuable in PAD and has become the gold standard due to its ability to provide detailed anatomical information on the arterial lumen, calcification plaques, and metallic stents. Tissue density differences are expressed in Hounsfield units (HU), allowing clear differentiation between vessels, plaques, and surrounding tissue. However, the artery lumen can show diverse HU ranges depending on the time between the injection of the contrast product and the image acquisition. Modern CTA systems offer high isotropic resolution, enabling visualization of small arterial branches and complex lesions.

2.5 Treatment strategies

2.5.1 Medical management

First-line management of PAD focuses on controlling risk factors and symptom relief. Lifestyle interventions such as smoking cessation and supervised exercise programs are foundational [Nordanstig, 2024]. Pharmacologic therapies include antiplatelet agents (e.g., aspirin, clopidogrel) to reduce thrombotic risk, statins to manage lipid levels, and

medications addressing comorbidities such as hypertension and diabetes [Nordanstig, 2024]. These measures aim to slow disease progression, alleviate symptoms, and prevent cardiovascular events.

2.5.2 Revascularization options

When medical therapy is insufficient, revascularization may be required (Fig. 2.3). Endovascular procedures are minimally invasive and commonly used for focal lesions. These include balloon angioplasty, where the vessel is dilated to restore flow, stent placement, which maintains lumen patency, and atherectomy, a technique in which a catheter-based device mechanically removes plaque from within the vessel. Endovascular approaches are particularly suited for short-segment or distal lesions, but may be limited by severe calcification or vessel tortuosity [Nordanstig, 2024].

Surgical bypass is considered for long-segment occlusions or when endovascular intervention is not feasible. It involves creating an alternative conduit for blood flow using either an autologous vein or prosthetic grafts, connecting proximal and distal healthy vessels. Endarterectomy, a surgical technique that removes plaque directly from the arterial lumen, may also be performed in selected cases, particularly in larger proximal vessels. The selection of bypass type, target site, and conduit material is dictated by lesion location, vessel quality, and patient comorbidities. Hybrid approaches, which combine endovascular and open surgical techniques, are increasingly employed for complex or multilevel diseases, offering the flexibility to treat different segments according to anatomical and pathological considerations [Nordanstig, 2024].

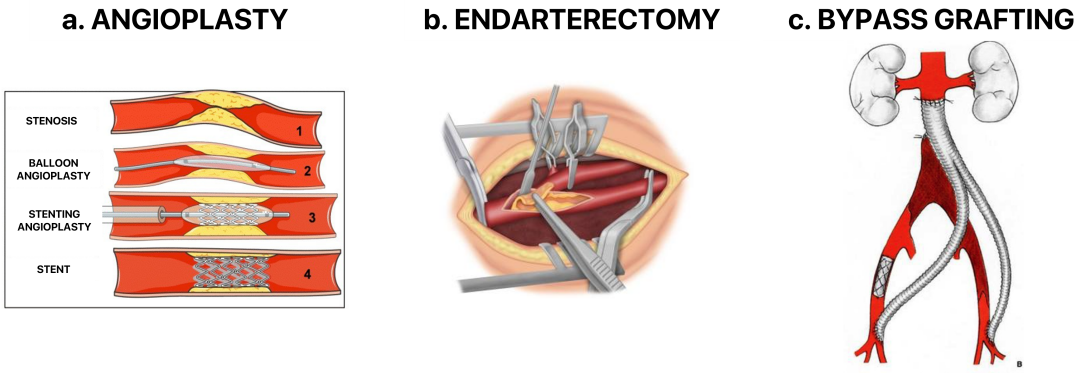


Figure 2.3.: Surgery strategies for revascularization. **a.** Balloon and stenting angioplasty, illustration from [Attias, 2018]. **b.** endarterectomy. **c.** Bypass grafting.

2.5.3 Main classification systems

The choice of treatment modality depends on multiple interrelated factors. Classification systems for PAD traditionally rely on clinical symptoms. The Fontain and Rutherford classifications are commonly used in clinical practice, but others additionally consider the severity and extent of vascular lesions from medical imaging. There is currently no single universally accepted classification system for PAD, but together, they can help decide the optimal treatment strategy [Nordanstig, 2024].

The Fontaine Classification: Is the oldest system [Fontaine, 1954]. It entirely relies on clinical symptoms and does not contain objective measures besides the maximum walking distance in stage II. It is used in clinical practice to guide decisions on conservative or invasive treatment [Nordanstig, 2024]:

- Stage I: Asymptomatic
- Stage IIa: Intermittent claudication after walking >200 meters
- Stage IIb: Intermittent claudication after walking <200 meters
- Stage III: Rest pain
- Stage IV: Ulcers or gangrene

The Rutherford Classification: Also considers patients' symptoms but also adds objective characteristics such as Doppler measurements, ABI, treadmill test and pulse volume recordings [Rutherford, 1997]:

- Grade 0: Asymptomatic
- Grade 1: Mild claudication
- Grade 2: Moderate claudication
- Grade 3: Severe claudication
- Grade 4: Rest pain
- Grade 5: Minor tissue loss
- Grade 6: Major tissue loss

Other classification systems rely on anatomical and clinical characteristics measured from imaging modalities [Lareyre, 2025; Nordanstig, 2024]. Lesion length and severity are primary considerations: short, focal lesions are usually amenable to endovascular therapy, whereas long occlusions may require bypass surgery. Lesion location, including proximal versus distal vessels and the presence of bifurcations, affects accessibility and intervention success. Vessel quality, particularly in heavily calcified or fragile arteries, dictates the suitability of stenting or bypass.

One of the most commonly used classification methods relying on imaging features is the Trans-Atlantic Inter-Society Consensus (TASC II) [Committee, 2015]. The arterial lesions are categorised into four categories based on the location and extent of lesions to guide decision-making of the revascularization technique. The Bollinger classification system was designed to describe the severity of the lesion based on the occlusion and atherosclerotic plaque patterns [Bollinger, 1981]. It enables the evaluation of atherosclerosis progression through follow-ups and can be used for assessment of angiograms and CT images [Chowdhury, 2017; Hardman, 2014]. In the case of CLTI, the global anatomic staging system (GLASS) score [Conte, 2019] aims to estimate the chance of success and patency of revascularization based on the extent and distribution of atherosclerotic lesions to improve the decision-making process. Finally, a classification system has been proposed for diabetic patients who present different lesions to categorize the severity of the disease into seven classes to facilitate the assessment of improvement in post-endovascular treatment [Graziani, 2007].

Additionally, vascular calcifications can also be assessed with specific scoring methods adapted from the Agatston score [Janowitz, 1991].

Patient-specific factors, such as age, comorbidities, and surgical risk, are also weighed to determine the optimal approach. Comprehensive imaging evaluation allows clinicians to integrate these considerations and plan individualized interventions.

2.5.4 Current challenges

PAD remains underdiagnosed and undertreated, contributing to delayed intervention and increased morbidity. While imaging-based assessment can improve the diagnosis and guide the revascularization strategy supported by classification systems, it is overall time-consuming and subject to inter-observer variability, particularly in complex cases with small, distal arteries, severe calcifications, or altered anatomy due to prior interventions. Additional challenges include occlusions that render vessels invisible on CTA due to the blockage of the blood flow and therefore of the contrast agent, high anatomical variability of collaterals, and confounding appearances of stents, bypass grafts, or bone on CTA due to similar HU ranges. Automatic segmentation and anatomical analysis

tools could accelerate diagnosis and intervention planning, reduce variability, and enable quantitative evaluation of lesion characteristics and vessel connectivity, providing critical support for patient-specific risk assessment and procedural guidance.

3 Deep Learning-based Segmentation for Peripheral Artery Disease

Contents

This chapter examines deep learning-based segmentation methods in the context of peripheral arterial disease. Given the clinical and anatomical background presented previously, the focus here is on image analysis techniques capable of accurately delineating the lower-limb arterial tree, calcified plaques, and endovascular stents from CTA data.

We begin with a concise methodological overview of convolutional neural network architectures and evaluation metrics relevant to medical image segmentation, intended to provide essential context for readers from diverse backgrounds. The chapter then presents an in-depth review of related work, organized into distinct thematic areas: segmentation approaches developed specifically for PAD, and methods targeting similar structures in other anatomical regions.

The chapter concludes with a synthesis highlighting current gaps and unresolved challenges, which motivate the contributions introduced in the following chapters.

3.1 Brief methodological background

In medical imaging, segmentation refers to the delineation of anatomical structures or pathological regions within an image, often at the pixel or voxel level. Traditional methods such as thresholding, region growing, or watershed algorithms relied heavily on handcrafted rules and low-level image features like intensity, edges, or texture

descriptors [Pham, 2000; Rogowska, 2009; Sharma, 2010]. While these approaches were effective in simple scenarios, they often failed in the presence of noise, anatomical variability, or imaging artifacts. The emergence of deep learning, and in particular convolutional neural networks (CNNs), has transformed this field by enabling models to automatically learn rich and hierarchical features directly from data, surpassing the performance and robustness of traditional methods [Shen, 2017].

3.1.1 CNN-based Segmentation

CNNs are a class of deep learning models designed to process images by learning spatially localized features through convolutional filters [LeCun, 1989; LeCun, 1998; LeCun, 2002]. CNNs automatically learn meaningful image representations directly from data, making them particularly powerful for medical image analysis [Shen, 2017].

In the context of segmentation, CNNs are typically trained in a fully supervised paradigm. The model is provided with pairs of input images and corresponding expert-annotated masks ("ground truth"), and learns to predict pixel or voxel-wise segmentation maps. Training proceeds iteratively over multiple epochs. During each epoch, batches of images are passed through the network, and predictions are compared to the ground-truth masks using a loss function. This loss quantifies the error in segmentation (e.g., mismatch in overlap or boundary distance). The network's parameters are then updated through gradient-based optimization, using backpropagation to propagate errors backward and adjust the convolutional filters so that predictions gradually improve over time, progressively refining their ability to delineate structures.

The core building block of CNNs is the convolutional layer. Learnable filters, called kernels, are systematically applied across the image, acting as localized detectors for characteristic patterns such as edges, textures, or more complex features at deeper layers. This scanning operation captures both local spatial correlations and hierarchical structures in the data (Fig. 3.1). These layers are typically followed by non-linear activations, which allow the network to capture complex relationships beyond simple linear combinations. Following convolutions, CNNs often include pooling layers, which downsample feature maps by summarizing local neighborhoods. The most common type is max pooling, which retains the maximum value in each region, effectively preserving the strongest activations while reducing spatial dimensions.

For medical image segmentation, CNNs are most often organized in an encoder-decoder architecture. The encoder progressively reduces the image resolution through successive convolutions and pooling operations, thereby extracting increasingly abstract and global features. Conversely, the decoder upsamples the encoded representation back to the original resolution, recovering spatial details required for precise delineation.

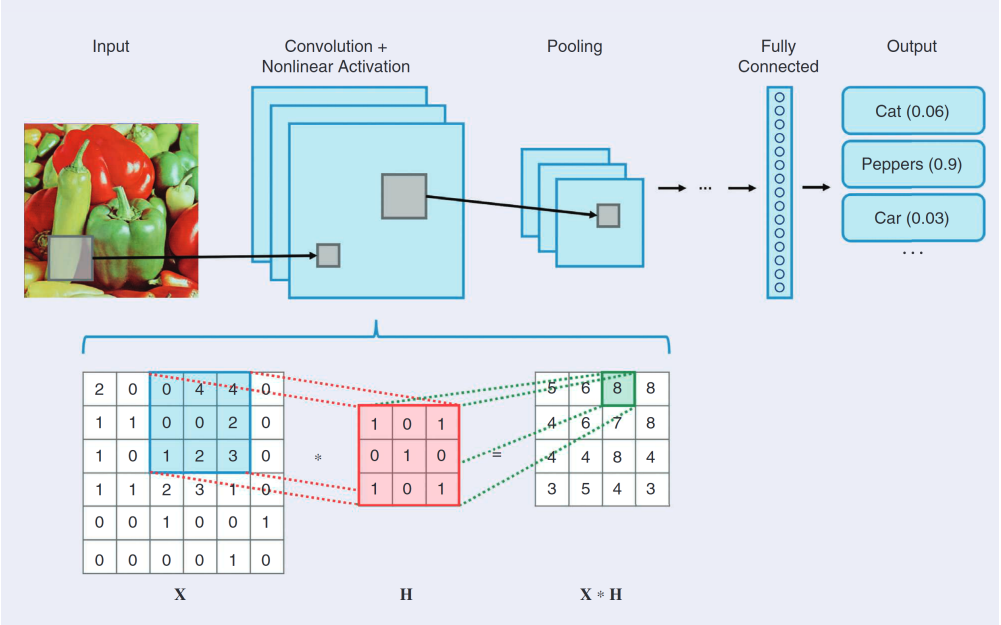


Figure 3.1.: Basic CNN architecture. The network applies convolutional layers with learnable filters that scan the input image to extract local features, interleaved with pooling layers for downsampling and non-linear activations to capture complex patterns. A learnable kernel (filter) H slides over the input image X, computing weighted sums at each position. (illustration from [Cheung, 2020]).

Among the most widely adopted architectures in biomedical imaging is the UNet [Ronneberger, 2015] (Fig. 3.2), which has become the standard backbone for medical image segmentation [Du, 2020; Siddique, 2020; Azad, 2024]. Its design, based on symmetric encoder–decoder pathways and dense skip connections, has proven effective across a wide range of anatomical targets. Skip connections between encoder and decoder stages play a crucial role in this design, as they directly transfer fine-scale information lost during downsampling. For volumetric imaging modalities such as CT or MRI, the 3D UNet [Çiçek, 2016] extends the same principle to capture contextual information in three dimensions. These architectures form the basis for many state-of-the-art segmentation frameworks, including the nnU-Net [Isensee, 2021; Isensee, 2024], which automatically adapts its configuration to the dataset at hand.

Therefore, CNNs are particularly well-suited to medical image segmentation because they can integrate both local texture information and larger anatomical context, making them ideal for capturing complex structures such as vascular trees.

3.1.2 Main loss functions for medical image segmentation

The loss function is a key component in the training of convolutional neural networks, independently of the chosen architecture. Losses quantify the discrepancy between the network's predictions and the expert-provided ground truth segmentations. This error

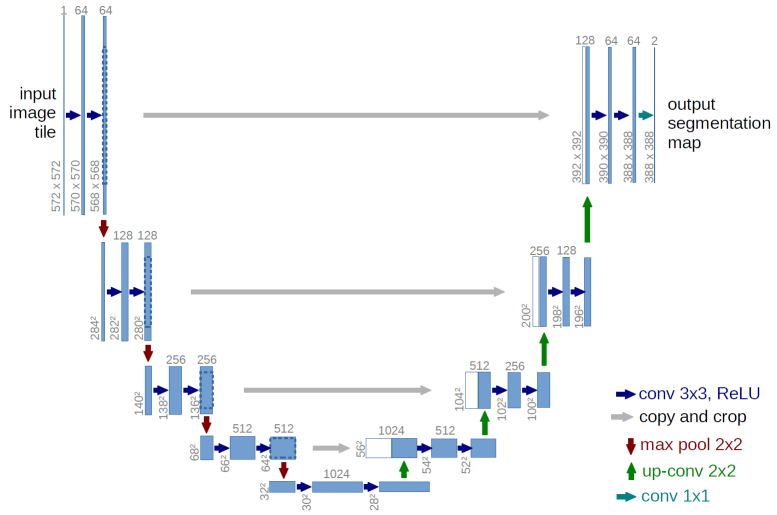


Figure 3.2.: Architecture of the UNet, showing the encoder-decoder structure with skip connections (illustration from [Ronneberger, 2015]).

signal is then backpropagated through the network to update its weights by gradient-based optimization.

A wide variety of loss functions have been developed in recent years to mitigate these challenges [LeCun, 2015; Ma, 2021]. While no single loss achieves universally optimal results, the most commonly used formulations are summarized below.

Cross-Entropy Loss (CE). The cross-entropy loss is the most standard formulation for classification problems, treating segmentation as voxel-wise classification. For a voxel i with true label $y_i \in \{0, 1, \ldots, C-1\}$ and predicted probability $\hat{p}_{i,c}$ for class c, the multi-class cross-entropy is:

$$\mathcal{L}_{CE} = -\frac{1}{N} \sum_{i=1}^{N} \sum_{c=1}^{C} y_{i,c} \log(\hat{p}_{i,c}),$$

where N is the number of voxels, and $y_{i,c}$ is a one-hot encoded label. While effective, CE is often sensitive to class imbalance, as the large background class can dominate the optimization. Particularly in medical image segmentation, there is usually a strong class imbalance between the background voxels and the target anatomical structures. For instance, small-caliber vessels or small lesions may occupy only a tiny fraction of the voxels in a 3D scan, which can bias the loss towards the background class.

Binary Cross-Entropy (BCE). For binary segmentation tasks, the CE can be reduced to the binary cross-entropy:

$$\mathcal{L}_{BCE} = -\frac{1}{N} \sum_{i=1}^{N} \Big(y_i \log(\hat{p}_i) + (1 - y_i) \log(1 - \hat{p}_i) \Big),$$

where $y_i \in \{0, 1\}$ is the ground truth and $\hat{p}_i \in [0, 1]$ the predicted probability. BCE is simple and widely used, but suffers from the same imbalance issue as CE.

Dice Loss. The Dice loss directly optimizes the overlap between the predicted segmentation and the ground truth mask. For predictions $\hat{p}_{i,c}$ and ground truth labels $y_{i,c}$, the soft Dice loss is defined as:

$$\mathcal{L}_{Dice} = 1 - \frac{2 \sum_{i=1}^{N} \sum_{c=1}^{C} y_{i,c} \hat{p}_{i,c}}{\sum_{i=1}^{N} \sum_{c=1}^{C} y_{i,c} + \sum_{i=1}^{N} \sum_{c=1}^{C} \hat{p}_{i,c}}.$$

Variants exist, for example, with squared terms in the denominator to further penalize large mismatches. This formulation is less sensitive to class imbalance by normalizing the overlap relative to the size of both prediction and ground truth.

Compound Losses. Compound losses are loss functions that combine two or more individual losses, usually through a weighted sum. For instance, to simultaneously optimize voxel-wise accuracy and region-level overlap, the CE and Dice losses are often combined:

$$\mathcal{L} = \alpha \, \mathcal{L}_{CE} + \beta \, \mathcal{L}_{Dice},$$

where α and β are weighting factors controlling the relative contribution of each loss. This hybrid approach leverages the complementary strengths of both losses, improving overall performance.

Beyond Classical Losses. In addition to the commonly adopted formulations described above, a wide variety of task-specific loss functions have been proposed in the literature. For instance, the Tversky loss [Salehi, 2017] and the focal loss [Abraham, 2019] introduce re-weighting strategies to emphasize hard-to-classify voxels or underrepresented classes. Other formulations target boundary precision, topological preservation, or uncertainty modeling for instance, depending on the application at hand.

3.1.3 Main evaluation metrics

Final predicted segmentations are evaluated by measuring their similarity or distance to the ground truth annotations. In medical image segmentation, a range of metrics are commonly used to quantify both volumetric accuracy and structural fidelity [Taha, 2015; Müller, 2022].

Overlap-based metrics. These metrics evaluate the agreement between the predicted segmentation P and the ground truth G in terms of shared voxels.

The Dice Similarity Coefficient quantifies the volumetric overlap between prediction and ground truth. It ranges from 0 (no overlap) to 1 (perfect overlap):

$$\text{Dice}\,(P, G) = \frac{2\,|P \cap G|}{|P| + |G|}$$

A related metric is the Jaccard Index or Intersection over Union (IoU):

$$\text{Jaccard}\,(P, G) = \frac{|P \cap G|}{|P \cup G|}$$

Precision and Recall assess over and under-segmentation tendencies:

$$\text{Precision} = \frac{TP}{TP + FP} \qquad \text{Recall} = \frac{TP}{TP + FN}$$

where TP, FP, and FN denote true positive, false positive, and false negative labeled voxels, respectively.

The F1 Score is the harmonic mean of precision and recall, providing a balanced measure:

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

True Positive Rate (TPR) and False Positive Rate (FPR) are formulated such that:

$$\text{TPR} = \frac{TP}{TP + FN}, \quad \text{FPR} = \frac{FP}{FP + TN}$$

Distance-based metrics. These metrics evaluate the spatial accuracy of boundaries.

The Hausdorff Distance (HD) measures the maximum distance between predicted and ground truth boundaries:

$$\text{HD}\,(P, G) = \max\left\{\sup_{p \in P} \inf_{g \in G} d(p, g), \sup_{g \in G} \inf_{p \in P} d(g, p)\right\}$$

where $d(p, g)$ is the Euclidean distance between points p and g. The 95th percentile HD (HD95) is commonly reported to reduce sensitivity to outliers.

Topology-based metrics. These metrics aim to assess the preservation of anatomical continuity and connectivity, which are crucial in vascular segmentation.

The centerline Dice (clDice) [Shit, 2021] score extends the traditional Dice metric by emphasizing the alignment of tubular structures along their centerlines. It ensures that not only the volume but also the continuity of structures is preserved.

$$\mathrm{clDice} = \frac{2 \cdot (|S_P \cap G| / S_P) \cdot (|S_G \cap P| / S_G)}{(|S_P \cap G| / S_P) + (|S_G \cap P| / S_G)}$$

where S_P and S_G are the skeletonized prediction and ground truth masks.

Betti numbers (β_0, β_1, β_2) quantify the number of connected components, loops, and enclosed cavities, respectively. The Euler characteristic is defined as:

$$\chi = \beta_0 - \beta_1 + \beta_2$$

Betti number errors can be reported as the mean absolute difference between predicted and ground truth values.

Detection-based metrics. These metrics assess the accuracy of identifying discrete structures. Precision and recall can be computed at the object level, evaluating whether individual structures are correctly detected. In practice, a predicted object is typically considered correct if it overlaps sufficiently with a ground-truth object, allowing object-level evaluation to focus on detection rather than exact voxel-level agreement.

The Free-response Receiver Operating Characteristic (FROC) [Bunch, 1977] curve plots the sensitivity (recall) of object detection against the average number of false positives per image or volume. It is particularly useful when multiple discrete structures exist per scan, and allows evaluation of performance across varying confidence thresholds.

The Mean Average Precision (mAP) [Buckley, 2017; Padilla, 2020] calculates the area under the precision-recall curve for each class and then averages across classes or targets:

$$\mathrm{mAP} = \frac{1}{N_{\text{objects}}} \sum_{i=1}^{N_{\text{objects}}} \mathrm{AP}_i$$

where AP_i is the average precision for object i, computed from the precision-recall curve.

Detection-based evaluation complements overlap metrics, particularly when segmenting small or sparsely distributed objects, where voxel-level measures may underestimate clinical relevance.

Together, these metrics provide a comprehensive assessment of segmentation quality, capturing both voxel-wise accuracy and higher-order anatomical fidelity.

3.2 Deep learning segmentation applications for PAD

In vascular imaging, CNN-based architectures have been widely adopted to tackle the challenges of segmenting small, thin, and highly connected structures, such as arteries and veins. In this section, we first review existing CNN-based approaches for the segmentation of vessels, then we examine the state-of-the-art for the segmentation of lower-limb arteries and PAD-relevant structures. The section concludes by highlighting key opportunities and remaining challenges in the automated segmentation of arteries, calcifications and stents in the context of PAD.

3.2.1 Vascular segmentation

Automated segmentation of blood vessels is a distinct and technically demanding subdomain of medical-image segmentation that has been extensively studied [Moccia, 2018]. Recently, the segmentation of vessels through CNNs has been explored for diverse vessel structures such as the aorta [Mohammadi, 2019; Lareyre, 2021; Imran, 2025], coronary arteries [Dong, 2022; Song, 2022; Wolterink, 2019], cerebral vasculature [Livne, 2019; Phellan, 2017; Hilbert, 2020] and other type of vessels [Tan, 2021; Nardelli, 2018; Taha, 2018; Xiao, 2018].

Vessel segmentation poses distinct challenges compared to the segmentation of other anatomical structures. Vessel voxels constitute only a very small fraction of the total image volume, leading to a severe class imbalance. This is particularly pronounced in distal branches, which often have very small radii, whereas deep learning models operate under the assumption that the object being segmented takes up a significant portion of the image [Deng, 2009; Shelhamer, 2017; Dang, 2022]. Additionally, vascular structures exhibit a wide range of scales and can extend over long anatomical distances, necessitating network architectures capable of capturing both fine local details and global context. Variability in contrast timing, imaging artifacts, and differences across imaging modalities further complicates model generalization. Another significant challenge is the labor-intensive nature of fully manual 3D annotation of entire vascular trees, which limits the availability of large, multi-center annotated datasets and impedes the training of data-intensive models. Collectively, these factors make vessel segmentation a particularly demanding task in medical image analysis.

To address these challenges, a variety of model architectures and methodological strategies have been proposed.

CNN encoder-decoder based models

UNet-based networks remain the backbone for vessel segmentation because they combine multilevel features with skip-connections that recover spatial detail [Livne, 2019; Chen, 2019; Pan, 2021; Xie, 2020].

The self-configuring nnUNet framework is widely adopted as a strong baseline for medical image segmentation [Isensee, 2021], because it automates preprocessing, patch sizing and hyperparameters to the dataset. Applications of nnU-Net span a wide range of vascular contexts, for example, in carotid artery and calcified plaque segmentation in CTA [Zhu, 2022] to general-purpose models covering coronary arteries, the aorta, and iliac vessels [Ouyang, 2024].

Building on the U-Net paradigm, several specialized architectures have been proposed to address the challenges inherent to vascular structures. For instance, BRAVE-NET [Hilbert, 2020] extends the 3D U-Net with multiscale feature extraction and deep supervision, enhancing sensitivity to small cerebral arteries. Similarly, 3D residual U-Nets [Yu, 2019] and models incorporating residual or dense blocks allow improved gradient flow and feature reuse, increasing segmentation accuracy in fine vascular structures.

Beyond traditional U-Net extensions, other encoder-decoder CNNs have been developed to further enhance vessel segmentation. SegResNet [Myronenko, 2018] integrates residual connections within its encoder to facilitate deeper network training and improved feature propagation, while DiNTS [He, 2021] leverages differentiable neural architecture search to automatically optimize U-Net-like topologies for a given dataset. DeepVesselNet [Tetteh, 2020] exemplifies a design that explicitly predicts centerlines and bifurcations in addition to segmentation masks, reducing topological errors. Its architecture avoids downsampling layers that can erase thin vessel structures and employs cross-hair 3D convolutions, which capture spatial information in all three dimensions efficiently, making it particularly effective for vascular segmentation.

Attention mechanisms

Attention modules enhance these architectures by emphasizing relevant vascular features and suppressing background noise across multiple scales. Attention U-Net [Oktay, 2018] introduces attention gates within skip connections, enabling the decoder to focus on pertinent encoder activations, thereby improving the detection of small vessels without significant computational overhead. Subsequent models have integrated advanced attention mechanisms: CS2-Net [Mou, 2021] employs both channel and spatial attention modules to capture hierarchical representations of curvilinear structures. GCA-Net [Ni, 2020] utilizes a global channel attention mechanism to capture long-range dependencies,

improving the segmentation of intracranial vessels in CTA images. SA-UNet [Guo, 2021] introduces cross-scale spatial attention in all skip connections, strengthening multi-scale feature fusion and enhancing segmentation performance on retinal vessel datasets. Overall, attention mechanisms enhance recall and focus on vessels, particularly when training data is limited, though they can increase computational cost, network complexity, and sensitivity to small datasets.

Transformers and hybrid CNN–Transformer models

Vision Transformers (ViT) [Dosovitskiy, 2020] bring a complementary strength for explicit modeling of long-range dependencies and global context. In vessel segmentation, these properties help connect discontinuous, low-contrast branches and relate distal anatomy to upstream trunks.

A representative hybrid model is TransUNet, which employs a CNN front-end to extract fine-grained local features and a Transformer encoder on tokenized patches to capture global context [Chen, 2021; Chen, 2024]. Its decoder then recovers fine localization through skip connections, making this design a practical compromise that has demonstrated improved performance across a range of medical segmentation tasks, including recent applications in vessel segmentation [Zhang, 2024; Lim, 2025].

While vanilla ViTs operate with global self-attention, their quadratic complexity makes them computationally prohibitive for high-resolution volumetric images. To address this, the Swin Transformer [Liu, 2021] introduces two key innovations: a window-based self-attention, which restricts computations to local non-overlapping windows, and shifted windows, which allow cross-window information exchange. Together with hierarchical patch merging, these properties make Swin Transformers structurally analogous to CNNs and well-suited for dense prediction tasks such as segmentation.

Building on this foundation, Swin-Unet [Cao, 2022], UNETR [Hatamizadeh, 2022], and Swin-UNETR [Hatamizadeh, 2021] adapt transformer backbones for medical segmentation. Swin-Unet is a transformer U-Net-like model, in which both encoder and decoder are composed of hierarchical Swin Transformer blocks (with shifted windows) and patch-expanding layers, without any convolutional backbone. By contrast, UNETR employs a vanilla patch-embedding transformer encoder (ViT-style) that directly operates on volumetric 3D inputs, with multi-scale features linked via skip-connections to a CNN decoder. Swin-UNETR adopts a hybrid design: a hierarchical Swin Transformer encoder, summarizing multi-scale features via shifted windows, feeds into a convolutional decoder through U-Net-style skip pathways. All three methods enhance the ability to capture long-range dependencies across anatomy, but Swin-Unet and Swin-UNETR are especially efficient for high-resolution or large-volume inputs due to their local-windowed attention.

Moreover, Swin-UNETR has been shown to benefit from self-supervised pretraining, which mitigates data scarcity in volumetric medical segmentation.

Finally, recent studies have applied transformer architectures directly to vascular segmentation [Maurya, 2022; Kesari, 2025; Li, 2024]. These directions remain promising, though progress is still constrained by the limited availability of annotated vascular datasets and by domain shifts across imaging modalities. Overall, transformers offer distinct advantages when long-range anatomical consistency is critical, but they remain computationally intensive and sensitive to data distribution.

Topology preservation and connectivity-aware losses

Preserving the topological integrity of vascular structures is a central challenge in vessel segmentation, as even small discontinuities or broken branches can have clinically significant implications. Consequently, a variety of losses and methods have been introduced to explicitly encourage connectivity and topology preservation.

A key contribution in this domain is the centerline Dice (clDice) loss [Shit, 2021], which optimizes the intersection between predicted masks and ground-truth skeletons. By relying on a differentiable soft-skeletonization algorithm, the clDice loss provides a training signal that encourages continuity along vessel centerlines and has been widely adopted for tubular structure segmentation. However, subsequent studies have questioned the strict topological reliability of the skeletonization step [Menten, 2023], motivating refinements. Extensions such as the centerline-boundary Dice (cbDice) [Shi, 2024] integrate centerline preservation with boundary-awareness to address vessel diameter imbalance, while the Skeleton Recall Loss [Kirchhoff, 2024] improves efficiency by reducing computational overhead. Together, these approaches attempt to balance connectivity preservation with geometric fidelity across a range of vessel sizes.

In parallel, topology-aware losses derived from persistent homology have been proposed. These methods compute persistence diagrams of predicted and ground-truth segmentations and penalize discrepancies in topological features such as connected components or loops. For example, Clough et al. [Clough, 2020] introduced a cost function that enforces agreement between the persistence barcodes of the predicted segmentation and the theoretical Betti numbers of the target object. More recently, Stucki et al. [Stucki, 2023] designed a loss that spatially matches topological features between predictions and ground truth, achieving a coherent correspondence of observed structures. While these approaches are powerful in enforcing global topological constraints, they are computationally demanding and require careful hyperparameter tuning to remain stable during training, which can limit their applicability, particularly for 3D volumetric datasets.

Beyond centerline and topology-specific losses, boundary-based formulations offer complementary perspectives. Losses such as the boundary loss [Kervadec, 2021], distance map loss [Caliva, 2019], and Hausdorff distance loss [Karimi, 2019] explicitly minimize distances between predicted and reference vessel surfaces. These distance-based losses, however, typically require auxiliary region-based losses such as the soft Dice loss to ensure stable training. Moreover, the differentiability of the distance transform function used in [Karimi, 2019] is not addressed.

Other methods explicitly incorporate anatomical priors to enhance vascular structure delineation. [Zhang, 2023] proposed TopoLab, a framework that integrates anatomical connections into the network design through intra and inter-segment feature interactions, along with an anatomy-aware connection classifier. Complementarily, [Zhou, 2025] recently introduced the Global-to-Local Connectivity Preservation (GLCP) framework, which jointly learns global segmentation, skeleton maps, and local discontinuity maps via an Interactive Multi-head Segmentation module. The addition of a Dual-Attention-based Refinement module further enhances segmentation quality, demonstrating superior accuracy and continuity in tubular structure. Finally, post-processing strategies have emerged to restore connectivity after segmentation [Qiu, 2023].

Trends and best practice

The segmentation of arterial systems has been the focus of several dedicated benchmarks, including the TOPCoW (Topology-Aware Anatomical Segmentation of the Circle of Willis) challenge [Yang, 2024] and the Aortic Segmentation Challenge [Imran, 2025]. These initiatives have stimulated the development of diverse segmentation strategies, ranging from classical CNN-based models to more recent transformer-based variants.

Overall, the nnU-Net framework has consistently emerged as the dominant baseline [Isensee, 2024]. Its self-configuring design enables automatic adaptation to new datasets and modalities, making it not only a universal starting point but also a strong upper bound that many domain-specific methods seek to surpass. The repeated success of nnU-Net across challenges underlines both its robustness and its role as the backbone upon which most competing methods build, often complemented with specialized loss functions, ensembling techniques and specific pre and post-processing pipelines.

These challenges additionally explore multi-class segmentation problems for semantic arterial labeling. Leading methods frequently extend baseline frameworks with targeted augmentations. For instance, winning solutions for TOPCoW 2024 employed a 3D nnU-Net backbone with topology-preserving loss functions.

3.2.2 Segmentation of the arteries of the lower limbs

While deep learning has been increasingly applied to PAD for tasks such as stenosis classification or severity grading [Dai, 2021; Goffart, 2025a; Ara, 2019; McBane, 2024], segmentation of lower-limb arteries remains a much less explored domain. Despite its clinical relevance, only a limited number of studies have specifically investigated the segmentation of PAD-relevant vascular structures in the lower limbs. This scarcity can be attributed to several challenges. First, lower limb arteries extend over a very long anatomical course, which increases the computational burden of automatic segmentation. For the same reason, manual annotation is highly labor-intensive, and consequently, no open-source reference dataset currently exists for this anatomy. Moreover, distal vessels such as the tibial and fibular arteries are very small in caliber, further complicating segmentation. In pathological cases of PAD, the vasculature demonstrates marked inter-patient variability due to the presence of collateral pathways, occlusions, calcifications, stents, and surgical bypasses. These factors increase the complexity of the task and hinder the development of robust and generalizable segmentation models.

The first attempts to address this problem relied on traditional model-driven techniques. [Chen, 2015], for example, proposed a semi-automated method for black-blood MRI in which manually initialized landmarks were refined using geometric operators and a 3D deformable model guided by intensity and gradient information to delineate the lumen and outer wall of femoral arteries. This approach achieved relatively low mean absolute distances (0.43 and 0.59 mm), but remained dependent on user input and was never extended to distal vessels.

More recently, the field has shifted towards CNN-based segmentation of CTA. [Bagheri Rajeoni, 2023] employed a U-Net with a pretrained ResNet-34 encoder to jointly segment the arterial lumen and calcifications, reporting a Dice score of 0.83 for arteries, yet their work was limited to iliac arteries down to the patella, therefore not including below-knee vessels. Similarly, [Guidi, 2022] developed an automatic segmentation framework for CTA focused on the aorta and iliac arteries, including calcification quantification, but did not attempt segmentation of peripheral vessels. [Zulfiqar, 2026] combined YOLOv8 for region-of-interest localization with a vascular segmentation network, achieving a Dice score of 0.91, yet their dataset contained mostly proximal arteries, with only 18% of cases including tibial vessels, and none of the samples represented pathological presentations. Other approaches have emphasized detection rather than full voxel-wise segmentation. [Anwer, 2025] used a region-based CNN (R-CNN) with a ResNet-101 backbone and a custom loss function to detect arteries and classify stenoses from the iliac arteries to the mid-knee level, reporting very high accuracy (98.79%, AUC = 0.95). However, their annotations were limited to bounding boxes and excluded below-knee vessels.

Beyond CTA, deep learning has also been explored for magnetic resonance angiography (MRA). [Ghodrati, 2022] proposed a 3D U-Net with local attention gates and a focal Tversky loss to segment arteries and veins from FE-MRA. By leveraging the temporal dynamics of time-resolved imaging, they separated arteries from veins and obtained promising results (F1 = 0.81, recall = 0.84). However, as with the CTA-based studies, the dataset was largely limited to proximal vessels above the knee and did not include cases with calcifications or occlusions.

Regarding calcifications, [Zou, 2025] applied an nnU-Net to calcification segmentation in the lower limbs, but performance was poor (Dice = 0.48), likely due to the small dataset of only 27 cases. Additionally, there is no information on the exact anatomical coverage. In coronary imaging, several methods have targeted plaque detection and characterization from intravascular optical coherence tomography (IVOCT) or CT. Approaches range from conventional encoder–decoder models such as SegNet refined with conditional random fields [Gharaibeh, 2019] to two-step pipelines combining detection and segmentation networks [Lee, 2020], as well as lightweight 3D CNNs for calcium scoring across multiple arteries [Santos, 2025].

Stent and graft segmentation constitutes another technically demanding task due to metallic artifacts and complex geometry. Early approaches relied on geometric graph-based methods for detecting stent grafts in the aorta [Klein, 2012], whereas more recent studies have leveraged 3D U-Nets with specialized patching strategies to improve lumen and stent delineation in the aorta [Sabrowsky-Hirsch, 2021].

Taken together, these works reveal important common limitations. Anatomical coverage remains restricted, with the majority of methods stopping at the femoral or popliteal level and rarely extending to tibial vessels, even though distal arteries are clinically relevant in PAD. Pathological robustness is also limited, as nearly all reported datasets consist of healthy or minimally diseased cases and do not always explicitly handle calcifications, occlusions, or metallic stents. Target structures are generally restricted to the lumen, with little attention given to calcifications or stents. Overall, despite recent progress, PAD-specific segmentation of lower-limb arteries remains in its infancy as current methods either stop at proximal levels, fail to include pathological cases, or only address partial structures. These gaps highlight the need for dedicated approaches capable of robustly segmenting the full vascular tree and PAD-relevant structures in real clinical data.

3.2.3 Synthesis: gaps, challenges and opportunities

Despite significant advances in vessel segmentation across various anatomical regions, few studies have specifically targeted lower-limb arteries. Existing works are generally limited to proximal arteries, rarely encompassing the full lower-limb arterial tree. Even

fewer approaches have addressed segmentation in pathological contexts characteristic of PAD, where occlusions, calcifications, and stents introduce additional complexity. While some efforts have attempted to segment calcifications or stents, these are often constrained either by limited anatomical coverage or by suboptimal performance. Notably, calcification segmentation appears to benefit from prior artery lumen segmentation, which likely provides crucial anatomical context. To date, no studies have specifically addressed stent segmentation or branch identification for lower-limb arteries in PAD.

Segmentation strategies developed for related vascular structures, however, illustrate both the versatility of deep learning approaches and the persistence of common challenges, including class imbalance, preservation of small or thin structures, and mitigation of imaging artifacts. Across diverse domains, nnU-Net has consistently demonstrated robustness, emerging as the preferred backbone in numerous segmentation challenges and serving as a foundation upon which task-specific refinements can be applied. This suggests that PAD-focused segmentation could similarly benefit from leveraging nnU-Net as a base, augmented with methods designed to explicitly preserve vascular topology and fine-scale structural details.

Recent methodological advances provide additional avenues for improvement, particularly in enforcing topological consistency and reducing boundary errors. For instance, refinements to skeleton-based losses, such as clDice, and to boundary-aware losses, including Hausdorff distance variants, could enhance the continuity of segmented vessels and better capture complex morphologies.

Collectively, these observations highlight clear opportunities for future work in applying and developing PAD-specific segmentation methods that comprehensively cover the entire lower-limb arterial tree in pathological conditions, integrate lumen, calcification, and stent information, and identify branch locations. Methodologically, there are additional opportunities to refine artery segmentation. For example, improvements could address the limitations of the skeletonization algorithm used in clDice, enhancing the topological preservation of segmented vessels. Similarly, refinements to Hausdorff distance–based losses could tackle both differentiability and numerical stability issues, enabling more reliable boundary-aware optimization.

Part II

Methodological Contributions

4

SoftMorph: Differentiable Probablistic Morphological Operators for Image Segmentation

Contents

Building on the gaps and challenges identified in the previous chapter, particularly the need for improved topological preservation in vascular segmentation, this chapter introduces SoftMorph, a novel framework for integrating accurate morphological operations directly into deep neural networks. The motivation originated from limitations observed in the original clDice loss formulation, where the skeletonization step could produce unreliable centerlines, ultimately affecting topological performance. This raised

a broader question: how can we embed robust mathematical morphological operations into convolutional networks in a way that remains fully differentiable and trainable?

SoftMorph addresses this challenge by reformulating any binary morphological operation into a differentiable probabilistic equivalent, enabling seamless integration either as a component of a loss function or as a final network layer. The approach preserves the logical structure of the original operation while allowing gradient-based optimization. Initially developed to provide a more consistent skeletonization process for clDice, the framework is also evaluated as a final segmentation layer across multiple segmentation tasks and imaging modalities, with a focus on the topological performance for tubular structures.

Beyond the specific applications explored here, such as skeleton-based losses and morphological post-processing, SoftMorph lays the groundwork for broader use in medical image analysis and potentially other computer vision tasks, offering a generalizable tool to incorporate morphological reasoning into deep learning pipelines.

The work presented in this chapter was accepted to MICCAI 2024 [Guzzi, 2024] and currently under review for the journal Medical Image Analysis.

The code is publicly available on GitHub[1].

4.1 Introduction

Mathematical morphology is a fundamental theory used to process images by extracting or enhancing information based on geometrical shapes, structures, and spatial relationships [Serra, 1994]. Its methodology revolves around two fundamental operations: dilation expands the foreground objects in an image, increasing their size and connectivity, while erosion shrinks the foreground, eliminating small objects. Dilation and erosion can be combined or iterated to achieve more complex operations. These operations can enhance the essential shape and characteristics of an image, and extract shape features such as edges, holes, cracks, and corners [Shih, 2017]. Morphological operations operate on a Structuring Element (SE) that defines the pattern of the image feature to be processed, encompassing a large variety of shapes and sizes. Originally defined on binary images, morphological operations have subsequently been extended to gray-scale images. An important extension involves *fuzzy morphological operators* [Sinha, 1992; Bloch, 1995], which adapt the operations to continuous values, mainly by replacing erosion and dilation with minimum and maximum operations. Therefore, morphological operations are particularly useful for many image analysis and processing tasks such as

[1] https://github.com/lisaGUZZI/SoftMorph2.

image segmentation [Nomura, 2005; Xia, 2006; Said, 2006; Yucheng, 2009], feature extraction [Angulo, 2008; Feehs, 1987; Krishnapuram, 1992; Angulo, 2003; Zhao Yu-qian, 2005], noise reduction [Asano, 2004; Schonfeld, 1991; Peters, 1995; Stringhini, 2019; yong, 2010], or image enhancement [Mukhopadhyay, 2000; Román, 2021; Kimori, 2011; Amutha, 2011].

For instance, morphological operations were applied in radar imaging for image restoration by effectively removing noise while retaining critical information [yong, 2010]. Similarly, new operations were designed for edge detection in noisy medical imaging [Zhao Yu-qian, 2005]. In medical applications, skeletonization is the operation that extracts the centerline of a segmented structure such as vessels or the heart. It is often used to assess the organ's topology, connectivity and trajectory [Lidayová, 2017].

Morphological operations can enhance the performance of convolutional neural networks (CNNs) by extracting meaningful structural information to analyze images. Researchers have started integrating these operations into loss functions to improve specific characteristics of segmentation tasks [Jurdi, 2021; Shit, 2021]. Moreover, some operations such as dilation and erosion could be integrated as layers within CNN architectures, allowing to handle different patterns and tasks than traditional convolutional layers, hereby refining complex image analysis [Bloch, 2021; Nogueira, 2021]. However, traditional binary morphological operations are based on Boolean expressions, and some are implemented using minimum and maximum operations, ill-suited for the gradient-based optimization methods used in deep learning models. Indeed, that optimization process requires differentiable functions that are smooth and continuous to compute gradients effectively during the training phase [LeCun, 2015; Kumar, 2024]. Binary decisions and min/max operations lead to discontinuous or undefined gradients, making them inherently non-differentiable.

To address this, fuzzy morphological neurons were designed to align operators with gradient-based optimization, by approximating the *soft minimum and maximum functions* [Oh, 1998; Nakashizuka, 2009]. Standard min/max operations can also be seen as piecewise differentiable, directly replacing erosion and dilation with *min and max-pooling layers*, and combining them to achieve more complex operations [Franchi, 2020; Mondal, 2019; Jurdi, 2021; Shit, 2021]. Although these methods can successfully incorporate morphological operators in neural networks, they do not generalize beyond a restricted subset. Additionally, the use of min and max-pooling only, can lead to discontinuous skeletons and homotopy inaccuracies [Menten, 2023; Bloch, 2017]. Others have focused on *Learning Morphological Operations* by training CNNs to replicate specific morphological operations [Nguyen, 2021; Panichev, 2019]. These models can backpropagate gradients and be integrated into deep learning pipelines. However, they correspond to approximative morphological operators trained on a specific domain and may require fine-tuning. Lastly, one may consider *Convolution-like approaches* such as successive

convolutional layers with pre-defined kernels to replicate a particular operation [Menten, 2023], but this method can be hard to apply to some complex operations. To the best of our knowledge, there is no existing approach to integrate any morphological operation with any SE into a CNN in a smooth and differentiable manner.

To address these limitations, we introduce SoftMorph, a family of operators based on various fuzzy logics, to extend binary morphological operations on probability maps that can be seamlessly integrated into neural networks either as a loss function or as a final morphological layer. It provides a generalized approach for translating any binary morphological operator with any SE into differentiable and probabilistic equivalents. This chapter presents the following contributions:

- *We define probabilistic morphological operators* as the expectation of the binary operator over the probability of generating each possible binary configuration, expressed as a multi-linear polynomial deduced from the truth table.

- *We propose a family of quasi-probabilistic operators* for intractable truth tables of binary morphological operators. We use various fuzzy logic operators to convert the Boolean expression defining the morphological operation into a soft, differentiable expression. These quasi-probabilistic operators approximate the probabilistic operator while maintaining the complexity of the original binary operator. Unlike previous works on fuzzy morphological operators that applied to erosion and dilation only, our proposed expressions are applicable to any Boolean expression, hence easily adaptable to new custom operations.

- *We validate the SoftMorph operators* in a binary context to ensure consistency with the corresponding binary counterparts, and assess their backpropagation capability to ensure their smooth integration into deep learning pipelines.

- *We demonstrate applications of this framework* to various 2D and 3D segmentation tasks. We integrate the SoftMorph operators as final layers of a neural network's architecture and in loss function applications.

Overall, these contributions make the SoftMorph framework a powerful tool for integrating morphological operations into deep learning models. In section 4.2, we review related works to make soft, differentiable morphological operations and their integration into CNNs. Section 4.3 details the methodology of SoftMorph. We validate the soft morphological operators in a binary setting in section 4.4.3 and in section 4.4.4 we evaluate the backpropagation capability of the proposed operator representations. Finally, we demonstrate segmentation applications of the framework across multiple datasets in section 4.4.5.

4.2 Related Works

4.2.1 Continuous fuzzy morphological operators

Mathematical morphology has been extended from binary to gray-scale image processing partly through fuzzy set theory, as introduced by [Sinha, 1992]. In this context, gray-scale images are treated as fuzzy sets, where each pixel value is interpreted as a degree of membership to a set, and morphological operators are redefined within that fuzzy space. This approach enables more nuanced image analysis and has led to various representations, employing fuzzy conjunction and disjunction operators. The most notable are the minimum and maximum operations to represent erosion and dilation respectively [Bloch, 1995].

4.2.2 Differentiability of morphological operators

Subsequent works have addressed the differentiability of morphological operators through different techniques to integrate them into deep learning pipelines.

Counter harmonic mean

The counter harmonic mean (CHM) has been studied to approximate the erosion and dilation operations in [Angulo, 2010]. This approach involves raising the pixel values of an image to a certain power, combining them with weights, and normalizing the result. Thus, morphological dilation and erosion correspond to the limit cases when the power tends to $+\infty$ (dilation) and $-\infty$ (erosion). The differentiable nature of the CHM formulation facilitates gradient-based optimization and its application in CNNs [Masci, 2013; Mellouli, 2017; Mellouli, 2019]. A morphological neuron, the so-called Pconv layer, implements the CHM filter to learn the appropriate power value and weight parameters to optimize the operations. However, this method is subject to exploding gradients due to the power function [Franchi, 2020].

Log-Sum-Exp

The morphological operations have also been approximated using differential approximations of the minimum and maximum operations with the Log-Sum-Exp function. In [Nakashizuka, 2009] and [Shih, 2019], it has been tailored to optimize the SE of the morphological operations. In [Shen, 2022], it is furthermore improved to learn

non-flat SEs by applying bias variables to correct the rounding errors caused by the soft approximation of min/max.

Non-smooth minimum and maximum operators

In theory, minimum and maximum operations are not fully differentiable. While they are differentiable almost everywhere, they lack a well-defined gradient at points where multiple arguments of the operations are equal. In practice, these operations are treated as piece-wise differentiable. That means that the gradient is propagated only towards the maximum or minimum element during backpropagation. Consequently, the gradient is non-zero only at the element where the maximum or minimum is attained. When multiple arguments share the same maximum or minimum value, different strategies are employed depending on the context. In typical minimum/maximum operations, the gradient is distributed equally between those arguments. In commonly used max-pooling layers, the gradient is assigned exclusively to the first occurrence of the maximum value within each pooling window. Therefore, some researchers argue that these non-smooth operators can be used in deep learning pipelines to replicate the erosion and dilation operators, while still ensuring the optimization of models. For example, in [Jurdi, 2021], the combination of min and max-pooling layers allows to obtain the borders of the foreground. Similarly, the pooling layers are iterated to extract the skeleton in [Shit, 2021]. It is obtained by iteratively getting the difference between the erosion of the image and the opening of that erosion. However, this method leads to inconsistencies, creating disconnected skeletons and topological errors. As the standard pooling layers are fixed with a square SE, [Franchi, 2020] and [Mondal, 2019] proposes to replace them with learned morphological pooling layers that can optimize the SE. In [Franchi, 2020], non-flat SEs are optimized to learn the exact morphological operators.

Learning-based emulation of morphological operations

Morphological operations trained from data by neural networks inherently support gradient-based optimization. It can provide differentiable solutions for complex operations such as skeletonization, trained in CNNs for example in [Nguyen, 2021; Panichev, 2019]. Other works optimize the morphological operation and the SE learned from training data, by alternating dilation and erosion-like convolution layers [Aouad, 2022]. However, training from data is prone to domain shift and can produce topological errors [Menten, 2023].

Convolution-like approaches

A specific operation can be emulated by applying several convolution layers with predefined kernels to detect expected configurations and patterns in an image. For example, [Menten, 2023] detects simple points, corresponding to pixels that can be removed from the foreground to obtain a skeleton, only using convolutional and matrix operations followed by non-linear functions. They defined specific kernels to recognize 2D and 3D configurations and their rotational equivalent to check the presence of simple points. That process requires 57 convolutions to match the Boolean rules for simple point detection. The reparametrization trick is employed as the detection is based on binary criteria. Depending on the operation to replicate, this method can be computationally expensive due to the high number of convolutions required. Designing specific kernels to detect various configurations and their rotational equivalents can be complex and time-consuming, potentially limiting its adaptability.

4.2.3 Morphological operations in deep learning applications

Through the aforementioned methods to approximate differentiable morphological operations defined on continuous values, several applications of these operators have been explored in neural networks [Mondal, 2022; Hirata, 2021; Bloch, 2021].

Deep morphological neural networks

There has been a growing interest in replacing traditional convolutional layers of CNNs with morphological operations. Deep Morphological Neural Networks (DMNNs) leverage the inherent non-linearity of morphological operations to substitute the linear convolution and non-linear activation functions typically used in CNNs. These networks define morphological layers that approximate the erosion or dilation operations and optimize the SE based on target data. In some works, the appropriate sequence of erosion and dilation within the network architecture is also learned [Shen, 2022; Mondal, 2020; Masci, 2013]. DMNNs can alternate morphological layers with standard convolutional layers, and fully connected layers for classification tasks [Shen, 2022; Mellouli, 2019; Mellouli, 2017]. Other main applications of these networks include image restoration [Franchi, 2020; Mondal, 2019; Masci, 2013] and edge detection [Franchi, 2020]. Besides, [Mellouli, 2019] consider these networks more interpretable than usual CNNs because the learned sequence of morphological operations can be explicitly recovered and analyzed.

Loss functions with morphological operators

Some morphological operations can extract specific image features to be used in the loss function of a CNN to improve particular characteristics of the model. The clDice loss function in [Shit, 2021] requires the extraction of the skeleton to compute its intersection with the foreground volume, to maximize the topological preservation in the segmentation of tubular structures. In [Jurdi, 2021], they extract the boundaries of the segmentation to minimize the perimeter difference of the prediction with the ground truth volume. The segmentation of small and thin structures is enhanced in [Pihlak, 2021] by integrating the white top-hat operation to detect small structures. Compared to existing methods, our approach does not involve optimizing the SE or learning the morphological operations. Instead, we focus on adapting any known binary operator into a differentiable and probabilistic form, that one might want to use in a CNN to optimize an image analysis problem. It provides a solution to precisely replicate the binary operator, that can be integrated either in the loss function of a CNN or as a post-processing layer within the network. One key innovation of our approach is defining the differentiable operators on probability maps, enhancing their applicability. Additionally, we ensure to define quasi-probabilistic operators that match the computational complexity of the binary operators. Previous methods either approximate those operations, resulting in a mismatch with the exact binary operator, or are challenging to apply to complex operations. Overall, our method offers a more precise and flexible solution to incorporate morphological operations into deep learning frameworks.

4.3 Methodology

4.3.1 Definition of binary morphological operators

We define an image X consisting of N voxels $\{X_n\}$, $n = 1 \dots N$ and its corresponding binary segmentation image Y with $Y_n \in \{0, 1\}$.

A binary morphological operator $F()$ is applied to the binary image Y resulting in a final binary segmentation Z such that $Z = F(Y) \in \{0, 1\}^N$.

This operator $F()$ takes as input k binary variables selected from the neighborhood function $\mathcal{N}(i, n)$, providing the index of the ith neighbor of the voxel n in Y as $Y_{\mathcal{N}(1,n)}, \dots, Y_{\mathcal{N}(k,n)}$. It outputs a binary variable $Z_n = F(Y_{\mathcal{N}(1,n)}, \dots, Y_{\mathcal{N}(k,n)}) \in \{0, 1\}$. The neighborhood function corresponds to the SE of F. It defines the domain of the geometrical features processed in the morphological operation. For instance, in a 2D image, typical neighbor-

hoods are defined as $k = 4 + 1$ (Fig 4.1a) or $k = 8 + 1$ (Fig 4.1b). Similarly, in 3D they are defined as $k = 6 + 1, 18 + 1$ or $26 + 1$.

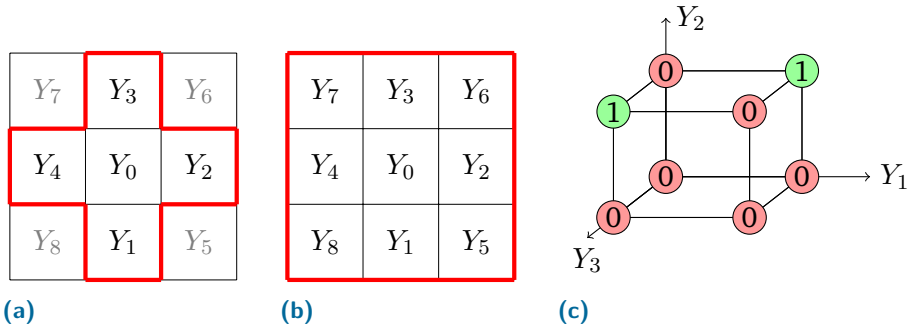


Figure 4.1.: Definition of the neighborhood variables in a 2D image. (a) 4+1 neighborhood. (b) 8+1 neighborhood. (c) Representation of the truth table of F_{Ex} on a hypercube. Each dimension corresponds to a binary input variable of the morphological operator.

The binary operator $F()$ is a Boolean function that has 2^k different possible input values and outputs a binary variable. The exhaustive list of those values $F(a) \in \{0, 1\}$, for $a \in \{0, 1\}^k$ is called the *truth table* of $F()$ and can be provided exhaustively for small values of k. Besides, it can be shown that any Boolean function can be written as a propositional formula involving the k binary variables with the logical operators AND ($\wedge$), OR ($\vee$) and NOT ($\neg$). In fact, two operators are sufficient in Boolean logic, since the AND and OR operators can be expressed with the remaining two according to De Morgan's theorem [Copi, 2016]. Similarly, the XOR ($\oplus$) operator can rewritten as: $(A \oplus B) = (A \vee B) \wedge \neg(A \wedge B)$.

Based on the notation defined in Fig.4.1, the dilation operator acting on a $4 + 1$-neighborhood of a 2D image can be written as $f_{\text{Dil}} = Y_0 \vee Y_1 \vee Y_2 \vee Y_3 \vee Y_4$ whereas the erosion operator is $f_{\text{Ero}} = Y_0 \wedge Y_1 \wedge Y_2 \wedge Y_3 \wedge Y_4$.

As an example, the operator f_{Ex} will be reused throughout this section to illustrate key concepts. The operator f_{Ex} is defined on a 1x3 grid such that $Z_n = f_{\text{Ex}}(Y_{n-1}, Y_n, Y_{n+1}) = F_{\text{Ex}}(Y_1, Y_2, Y_3)$. It is computed with the Boolean expression:

$$F_{\text{Ex}} = Y_2 \wedge ((Y_1 \wedge \neg Y_3) \vee (\neg Y_1 \wedge Y_3))$$

Consequently, with $k = 2 + 1$, there are 8 possible input configurations represented on the 3-hypercube of Fig. 4.1c.

4.3.2 Definition of probabilistic morphological operators

We want to extend these morphological operators to the output of segmentation algorithms. Typically, the output of a neural network consists of probabilities y_n, which

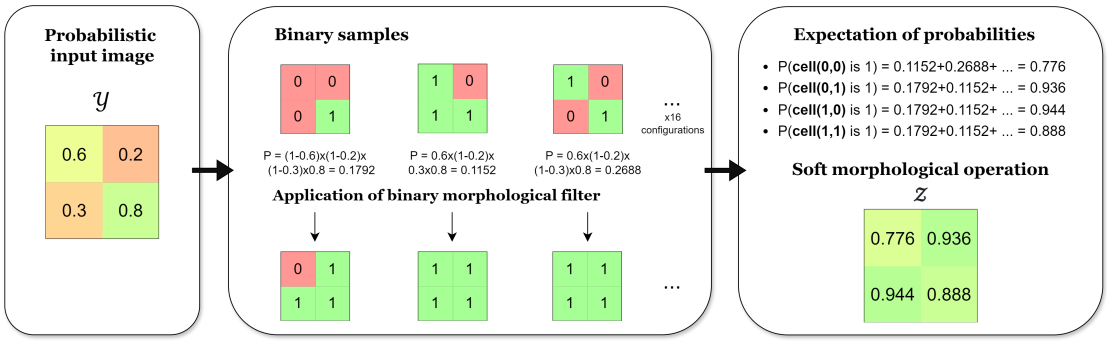


Figure 4.2.: Given a 2x2 input grid with probabilistic values, the probability of each possible binary configuration is calculated. By applying a specific morphological operator to each binary configuration, we compute the expectation over all configurations to estimate the result of the soft operator on the probabilistic input.

correspond to the posterior probability $y_n = p(Y_n = 1|X) \in [0, 1]$ of the binary variables $Y_n \in \{0, 1\}$.

We seek to formalize the definition of a probabilistic morphological operator $\mathcal{F}^\star()$ applied on the probabilistic segmentation image $\mathcal{Y} = \{y_n\} \in [0, 1]^N$ and generate a new probabilistic image $\mathcal{Z} = \{z_n\} \in [0, 1]^N = \mathcal{F}^\star(\mathcal{Y})$. This operator should generalize the given binary operator $Z = F(Y)$ such that both give the same result when the input probabilistic image is binary $Z = F(Y) = \mathcal{F}^\star(Y)$. More precisely, we aim to apply the deterministic morphological operator $F()$ on a binary image Y only known through its posterior probability $\mathcal{Y} = p(Y|X)$. Therefore, we estimate the posterior $z_n = p(Z_n = 1|X)$ of the final segmentation Z knowing that it results from the application of the morphological operation $Z = F(Y)$.

Lemma 1 *The posterior probability z_n can be obtained through the law of total probability as the expectation of the filtered binary segmentation $F(a)$:*

$$\begin{aligned} z_n = p(Z_n = 1|X) &= \sum_{Y_1=0}^{1} \dots \sum_{Y_N=0}^{1} p(Z_n = 1|Y)\, p(Y|X) \\ &= \sum_{a \in \{0,1\}^N} F(a)\, p(a|X) = \mathbb{E}_{a \sim p(Y|X)} F(a) \end{aligned}$$

Lemma 1, as represented in Fig.4.2, defines implicitly the relationship $z_n = \mathcal{F}^\star(\mathcal{Y})$ of the soft morphological operator.

4.3.3 Soft operators using Multi-linear polynomials

It is furthermore required to make the operator differentiable, i.e., to estimate the derivatives $\frac{\partial \mathcal{F}^\star(\mathcal{Y})}{\partial y_m}$. To provide a closed-form expression of a soft operator defined in

Lemma 1, we propose to adopt a polynomial representation of the Boolean function $F(a)$, $a \in \{0,1\}^k$. Indeed we can associate with any Boolean function $F(a)$ a multilinear polynomial $\mathcal{F}^\star(x)$, $x = (x_1, \ldots, x_k)^T \in \mathbb{R}^k$ defined as:

$$\mathcal{F}^\star(x) = \sum_{a \in \{0,1\}^k} F(a) \prod_{i=1}^{k} x_i^{a_i} (1 - x_i)^{1-a_i} \tag{4.1}$$

It is easy to see that $F(a) = \mathcal{F}^\star(a)$, i.e., that the polynomial $\mathcal{F}^\star(x)$ coincides by construction with the Boolean function on the hypercube $\{0,1\}^k$. Each monomial $F(a) \prod_{i:a_i=1} x_i \prod_{i:a_i=0} (1 - x_i)$ is equal to 0 if $x \neq a$ and equal to $F(a)$ otherwise. Besides, it is of degree k and linear with each variable x_j, making the polynomial multilinear. This property leads to the following result:

Theorem *The expectation of a Boolean function $F(a) \in \{0,1\}$, $a \in \{0,1\}^k$ over a set of k independent variables with $a \sim$ Bernoulli(p), $p \in [0,1]^k$ is $\mathcal{F}^\star(p)$*

Proof *It is easy to show that $\mathbb{E}_{a_i \sim \text{Bernoulli}(p_i)}(\alpha + \beta a_i) = \alpha + \beta p_i$ using the linearity of expectation. Thus, we have :*

$$\begin{aligned} &\mathbb{E}_{a_1}(F(a) \prod_{j=1}^{k} a_j^{a_j} (1 - a_j)^{1-a_j}) \\ &= F(a) p_1^{a_1} (1 - p_1)^{1-a_1} \prod_{i=2}^{k} a_i^{a_i} (1 - a_i)^{1-a_i}). \end{aligned}$$

By taking the expectation over each variable a_i, we get :

$$\mathbb{E}_a(F(a)) = \sum_{a \in \{0,1\}^k} F(a) \prod_{i=1}^{k} p_i^{a_i} (1 - p_i)^{1-a_i} = \mathcal{F}^\star(p).$$

□

Therefore, assuming that the marginal posteriors $y_n = p(Y_n|X)$ are independently distributed (which is the case when dealing with the output of segmentation neural networks or mean field approximations), we define the soft morphological operator associated with the binary operator $F()$ as the polynomial value $\mathcal{F}^\star(y_{\mathcal{N}(1,n)}, \ldots, y_{\mathcal{N}(k,n)}) \in [0,1]$.

With our example binary operator F_{Ex}, we get the multi-linear polynomial representation of the soft operator :

$$\begin{aligned} \mathcal{F}^\star_{\text{Ex}}(y_1, y_2, y_3) &= \sum_{a \in \{0,1\}^3, F(a) \neq 0} \prod_{i=1}^{3} y_i^{a_i} (1 - y_i)^{1-a_i} \\ &= y_1 y_2 (1 - y_3) + (1 - y_1) y_2 y_3 \\ &= y_1 y_2 + y_2 y_3 - 2 y_1 y_2 y_3 \end{aligned}$$

This function defined over the 3D cube can be seen as the trilinear interpolation of the binary truth table as illustrated in Fig. 4.3a. More generally, the probabilistic morphological operator defined in Theorem 1, can be interpreted as the multilinear interpolation of the binary operator F over the hypercube. Multilinear interpolation is among the most basic interpolation methods and is symmetric with respect to all variables, which makes it a natural extension of the binary operator F to the probabilistic context.

4.3.4 Limitations of Multi-linear polynomial representation

The construction of the multilinear polynomial $\mathcal{F}^\star()$ (Eq.equation 4.1) requires the summation over non-zero elements $F(a)$ of the truth table of size 2^k. For non-trivial truth tables, writing such polynomials requires the use of symbolic computation software such as SymPy or Maple. But the complexity of such polynomials grows exponentially with the number k of variables. For example, if $k = 26 + 1$ in a 3D images, there are $2^{27} = 134\ 217\ 728$ possible input configurations. In practice, when $k > 10$ for non-trivial operators, the number of monomials often becomes prohibitively large.

4.3.5 Quasi-probabilistic operators using fuzzy logic

The computation of $\mathcal{F}^\star$ is based on the exhaustive list of positive binary configurations which can become intractable to produce. We are looking for alternative computation methods that have the same complexity as the binary operator $F(a)$, $a \in \{0,1\}^k$. To this end, we notice that the probabilistic version of the AND operator is the product of the probabilities $\text{AND}(X_1, X_2) \longrightarrow x_1 x_2$ while the OR operator is transformed into $\text{OR}(X_1, X_2) \longrightarrow x_1 + x_2 - x_1 x_2 = 1 - (1 - x_1)(1 - x_2)$ and the NOT operator as $\text{NOT}(X_1) \longrightarrow 1 - x_1$. For any Boolean operator $F()$ represented by a proposition formula, involving the AND, OR and NOT operators, we propose to create a soft *quasi-probabilistic* operator $\mathcal{F}^\bullet()$ by substituting the logical operators AND, OR, and NOT by their probabilistic versions. This soft operator is a polynomial expression which can be of a degree greater than k, and is not necessarily multilinear. However, its computational complexity is the same as the one of the propositional form of the operator since we have substituted a simple logical expression with another simple polynomial one. The *quasi-probabilistic* polynomial corresponds to a compact and factorized form whereas the multi-linear polynomial corresponds to its polynomial expansion combined with the application of the idempotence rule $x^i = x, \forall i > 0$ on all probability variables. The quasi-probabilistic polynomial can be interpreted as an alternative non-linear interpolation method of the values over the hypercube whereas the probabilistic polynomial is a (multi) linear one.

While AND, OR and NOT operators were substituted by their probabilistic versions, one could think of other ways to replace the 3 logical operators with algebraic expressions of continuous values in the range $[0, 1]$. This has been the focus of the Fuzzy logic [Novák, 1999] and Fuzzy set theories. They introduce Triangular Norms (T-norms), and Triangular Conorms (S-norms) as substitutes for the AND and OR operators with the complementation operator $1 - x$ as the substitute for the NOT operator. T and S-norms are defined on the unit square $[0, 1] \times [0, 1]$ with values on the unique segment $[0, 1]$, and follow the commutativity, monotonicity, associativity and element identity properties. Besides, an S-norm $S()$ is dual of a T-norm $T()$ under the action on the complementation operator, $S(x_1, x_2) = 1 - T(1 - x_1, 1 - x_2)$, which can be seen as the generalization of the De Morgan's rules.

There exist many different T/S-norms proposed in the literature, which admit a partial pointwise ordering as follows:

$$T_1 \leq T_2 \quad \text{if} \quad T_1(a, b) \leq T_2(a, b) \quad \text{for all} \quad a, b \in [0, 1].$$

The expressions of the main T and S-norms are listed in Table 4.1 and their graphs on the unit square are displayed in Fig. 4.4. The *Product* T-norm corresponds to the probabilistic AND operators introduced previously and it is easy to see on Fig. 4.4 that the Drastic T-norm is the smallest whereas the Minmax logic is the largest. A number of those T-norms have additive generators [Dombi, 1982], which means that there exists a function $f : [0, 1] \longrightarrow \mathbb{R}^+$ such that $T(x, y) = f^{-1}(f(x) + f(y))$. This is the case for the Product logic (with $f(x) = -\log(x)$) and this allows to easily factorize multiple applications of the T and S-norms. For the Product rule, we have for instance : AND$(Y_0, Y_1 \ldots, Y_l) \equiv y_0 y_1 \ldots y_l$, OR$(y_0, y_1 \ldots, y_l) \equiv 1 - (1 - y_0)(1 - y_1) \ldots (1 - y_l)$

With each T-norm, we can substitute the AND, OR, and NOT operations of the binary function $F()$ to obtain an approximation $\mathcal{F}^\bullet()$ of the associated probabilistic operators $\mathcal{F}^\star()$. Both functions coincide on the vertices of the unit hypercube, but they correspond to different interpolation functions inside the hypercube. Thus, we call *SoftMorph* the family of soft operators derived from a binary morphological operator $F(Y_{\mathcal{N}(1,n)}, \ldots, Y_{\mathcal{N}(k,n)})$ which can be either *probabilistic operators* as multilinear polynomials $\mathcal{F}^\star(y_{\mathcal{N}(1,n)}, \ldots, y_{\mathcal{N}(k,n)})$ or *quasi-probabilistic operators* $\mathcal{F}^\bullet(y_{\mathcal{N}(1,n)}, \ldots, y_{\mathcal{N}(k,n)})$ as derived from a T / S-norm. The Einstein and Product logic generate smooth and differentiable expressions whereas the other T-norms are only piecewise differentiable, in particular due to the min and max functions. As an example, the *SoftMorph* operator associated with the binary function F_{Ex} using Product logic writes as $\mathcal{F}^\bullet{}_{\mathrm{Ex}}(y_1, y_2, y_3) = y_2(1 - (1 - y_1(1 - y_3))(1 - y_3(1 - y_1)))$. This is a factorized polynomial which differs from its probabilistic version and its expansion involves monomials of degree 5 (instead of 3 for the multilinear case) $y_1 y_2 + y_2 y_3 - 3 y_1 y_2 y_3 + y_1^2 y_2 y_3 + y_1 y_2 y_3^2 - y_1^2 y_2 y_3^2$. Both polynomials have the same values on the unit hypercube, but differ elsewhere as seen in

Table 4.1.: T-norms and S-norms formula for different fuzzy logics

Logic	Operation	Formula
Boolean	*AND*	$a \wedge b$
	OR	$a \vee b$
Drastic	*T-norm*	$\min(a,b)\ if\ \max(a,b) = 1, \quad else = 0$
	S-norm	$max(a,b)\ if\ min(a,b) = 0, \quad else = 1$
Bounded	*T-norm*	$max(0, a+b-1)$
	S-norm	$min(1, a+b)$
Einstein	*T-norm*	$(ab)/(2-(a+b-ab))$
	S-norm	$(a+b)/(1+ab)$
Product	*T-norm*	$a * b$
	S-norm	$a+b-ab$
Hamacher	*T-norm*	$0\ if\ a=b=0,$ $else\ (ab)/(a+b-ab)$
	S-norm	$1\ if\ a=b=1,$ $else\ (a+b-2ab)/(1-ab)$
Min-Max	*T-norm*	$min(a,b)$
	S-norm	$max(a,b)$

Fig. 4.3. The largest difference between the probabilistic and quasi-probabilistic functions on the unit cube is only around 0.06, showing that $\mathcal{F}^{\bullet}{}_{\mathrm{Ex}}$ is a good approximation of $\mathcal{F}^{\star}_{\mathrm{Ex}}$. In this simple case, there is no benefit to use $\mathcal{F}^{\bullet}{}_{\mathrm{Ex}}$ instead of $\mathcal{F}^{\star}_{\mathrm{Ex}}$, but for more complex functions, one must resort to quasi-probabilistic *SoftMorph* functions.

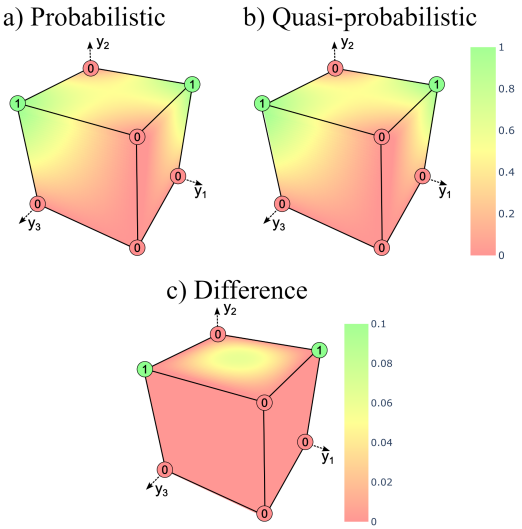


Figure 4.3.: Field plots of the probabilistic (a) and Product-based quasi-probabilistic (b) morphological operators of F_{Ex} over the hypercube. (c) Field plot of the difference between the two operators. The maximum difference is reached at the center of the faces or of the volume.

Computational Complexity of Probabilistic and Quasi-Probabilistic Operators

The probabilistic operator $\mathcal{F}^{\star}()$ is defined as a multilinear polynomial computed from the full truth table of a Boolean function with k binary inputs. Given its expression in Eq. 4.1, the expansion of the formula has a worst-case time and space complexity of $\mathcal{O}(k \cdot 2^k)$.

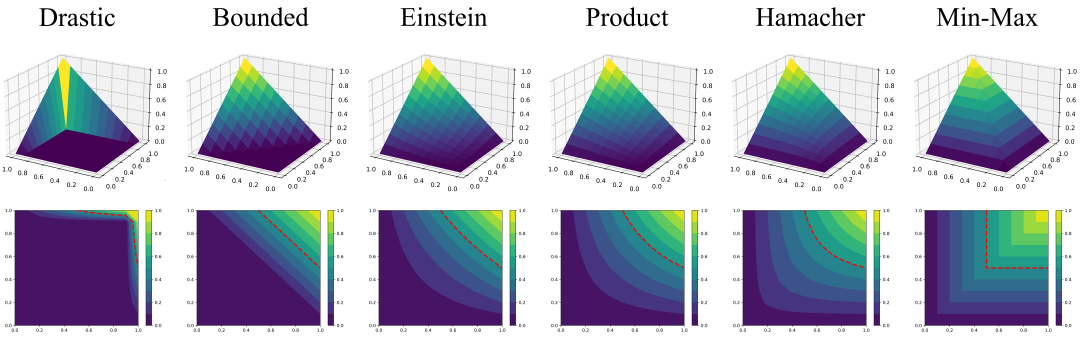


Figure 4.4.: Graphs of the Drastic, Bounded, Einstein, Product, Hamacher and minmax T-norms on the $[0,1]\times[0,1]$ unit square. The red dashed contours correspond to the isocontour at 0.5.

This arises because the full truth table of size 2^k must be traversed, and each entry yields a unique monomial term of up to degree k. For the quasi-probabilistic operators $\mathcal{F}^{\bullet}()$, logical operators in the Boolean expression are directly replaced by their fuzzy logic counterparts. Let m denote the number of logical operations in the Boolean expression defining F. Then the soft operator $\mathcal{F}^{\bullet}()$ has $\mathcal{O}(m)$.

Unlike $\mathcal{F}^{\star}()$, which must be fully expanded in advance, $\mathcal{F}^{\bullet}()$ is obtained by a symbolic substitution of logic operators with fuzzy arithmetic, leading to the same computational complexity as the binary boolean expression F. A direct runtime comparison between $\mathcal{F}^{\star}()$ and $\mathcal{F}^{\bullet}()$ is generally not meaningful since the cost of evaluating $\mathcal{F}^{\star}()$ depends on the size of the truth table (2^k), whereas $\mathcal{F}^{\bullet}()$ depends solely on the size of the logic expression (m). In summary, $\mathcal{F}^{\star}()$ is feasible and exact for small k and when the full truth table is available while $\mathcal{F}^{\bullet}()$ is tractable for any k and is recommended for large structuring elements (Table 4.2). Although $\mathcal{F}^{\star}()$ and $\mathcal{F}^{\bullet}()$ may have similar evaluation times for a fixed input once constructed, the construction of $\mathcal{F}^{\star}()$ is exponential due to the truth table size, while the construction of $\mathcal{F}^{\bullet}()$ is linear in the number of logical operations in the original Boolean formula. This makes $\mathcal{F}^{\bullet}()$ preferable for scalability and deployment.

Table 4.2.: Comparison of operator complexity.

Operator	Construction Time	Variable
Binary F	$\mathcal{O}(m)$	Depends on the logic expression
Probabilistic $\mathcal{F}^{\star}()$	$\mathcal{O}(k \cdot 2^k)$	Depends on the structuring element
Quasi-prob. $\mathcal{F}^{\bullet}()$	$\mathcal{O}(m)$	Depends on the logic expression

4.3.6 Relation to prior work

The dilation and erosion operators have been defined on any grayscale images using the notion of supremum and infimum[Shih, 2017]. On finite sets, dilation and erosion are

obtained by considering the maximum, and minimum values within a SE. They have been further generalized as fuzzy morphological operators [Bloch, 1995] with fuzzy (i.e. within the $[0, 1]$ range) SE and based on fuzzy logic (T / S-norms). However, when restricted to binary SEs, the fuzzy erosion and dilation operators are also equivalent to the ones defined in mathematical morphology (independently of the selected fuzzy logic), i.e. taking the maximum/minimum values in the neighborhood of a pixel.

It is important to note that those dilation and erosion operators are specific cases of the quasi-probabilistic morphological operators when adopting the Min-Max logic to these particular operations. Therefore, the *SoftMorph* operators are novel *soft* formulations of binary morphological operators that supersede the existing erosion and dilation operators. Those multi-linear and Product-based operators are smooth and fully differentiable whereas most other quasi-probabilistic ones require to adopt smooth min-max approximations to be differentiable. Furthermore, probabilistic morphological operators and their approximations are "averaged morphological operators" as the expectation of a binary operator. This implies that the *SoftMorph* closing and opening operations are not idempotent (producing the same output irrespective of the number of times it is applied) unlike the classical closing and opening using the Min-Max logic.

4.3.7 Morphological operators of interest

In this chapter, we generate the *SoftMorph* operators on 5 main morphological operations: erosion, dilation, closing, opening, and skeletonization. Depending on the SE and the image dimension, the erosion and dilation F_{Dil} Boolean expressions are written as $F_{\text{Ero}} = \bigwedge_{i=1}^{k} Y_k$ and $F_{\text{Dil}} = \bigvee_{i=1}^{k} Y_k$. The closing operation is a dilation followed by an erosion whereas the opening is an erosion followed by a dilation.

Skeletonization is a more complex morphological operation for which many algorithms have been proposed [Lam, 1992]. This operation involves repeated thinning operations, applied iteratively in all directions (North, South, East, West in 2D, with Up, Down added in 3D) until the final skeleton is obtained. We have translated the morphological thinning Boolean expression on 2D images proposed by Wagner *et al.* [Wagner, 2020] and the 3D thinning algorithm proposed by Palàgy [Palàgyi, 1998], as described in A.1.1 and A.1.2. In binary skeletonization, the algorithm ends when the binary structure is stable upon the application of each thinning sub-iteration. However, the SoftMorph are in general not idempotent (except with the Min-Max logic) and therefore a stopping criterion must be defined. We propose to stop the soft thinning process when the change between two thinning operations is less than 2% for all pixels in the initial foreground object (Algorithm 1).

Algorithm 1 Skeletonization operation

```
procedure SKELETON(img)
    obj ← sum(img)
    img_temp ← img
    change ← true
    while change do
        change ← false
        for orientation in {North, South, East, West, (Up, Down)} do
            img ← apply_thinning_formula(img, orientation)
        end for
        change ← TEST_CHANGE(img_temp, img, obj)
        img_temp ← img
    end while
    return img
end procedure

procedure TEST_CHANGE(img_temp, img, obj)
    change ← sum(img - img_temp) / obj
    return change > 0.2
end procedure
```

4.4 Experiments

4.4.1 Experimental setup

Datasets

Experiments are conducted on six 2D and two 3D datasets. In 2D, The DRIVE dataset [Staal, 2004] corresponds to retinal blood vessels. The Massachusetts Road [Mnih, 2013] dataset comprises satellite images of road networks. Labeled images are extracted from the Open Images Dataset V7 [Krasin, 2017] for classes Sea turtle, Starfish and Croissant. The Butterfly dataset [Wang, 2009] is composed of butterfly images. In 3D, We use the Vessap dataset featuring synthetic brain vessels [Paetzold, 2019] and the Liver task from the Medical Segmentation Decathlon [Antonelli, 2022].

Evaluation metrics

To evaluate the experimental performance on the final segmentations, we use the Dice similarity coefficient to measure the overlap with the ground truth. For datasets containing tubular structures we additionally use the clDice metric [Shit, 2021] to assess the topological preservation. Given the importance of topological accuracy in morphological operations, we compute the mean absolute error of topological invariants: The Betti

numbers $\beta 0$, $\beta 1$, $\beta 2$, representing respectively number of objects, holes and cavities, and Euler's characteristic ($\beta 0 - \beta 1 + \beta 2$). We use a connectivity of 26 in 3D and of 8 in 2D in the foreground.

Implementation

Experiments are implemented with Python 3.11.4, Pytorch 2.0.1 and 3 Nvidia A40 PCIe GPUs.

4.4.2 Skeletonization threshold

In the binary case, skeletonization terminates when there are no changes in pixel values between two consecutive thinning iterations. As discussed in Section 4.3.7, SoftMorph operators are generally not idempotent, meaning the skeletonization operation does not converge to a stable result without a clearly defined stopping criterion in a probabilistic setting. To address this, we aim to define a convergence criterion based on the relative change in pixel values between successive iterations. Specifically, we quantify the absolute pixel-wise difference between two consecutive thinning steps, and normalize it by the total sum of pixel values in the original image:

$$\Delta_t = \frac{\sum_{i,j} |I_t(i,j) - I_{t-1}(i,j)|}{\sum_{i,j} I_0(i,j)}$$

where $I_t(i,j)$ denotes the pixel value at position (i,j) after the t-th thinning iteration, and I_0 is the original input image. We monitor Δ_t over 20 iterations to evaluate the convergence behavior of each SoftMorph operator as illustrated in Fig. 4.5.

The analysis is first conducted on the whole DRIVE dataset. To accommodate probabilistic inputs, each binary mask is transformed into a continuous-valued representation by computing the signed distance transform, followed by min-max normalization to constrain values to the $[0,1]$ interval. To validate generality, we replicate the experiment on a synthetically generated dataset comprising 50 random probabilistic images of varying sizes (64×64 to 256×256). The images represent a combination of six morphologically distinct blob types: circular, elliptical, irregular star-like, branched, connected multi-component structures, and multiple separate blobs. Pixel intensities within object boundaries are randomly sampled from a uniform distribution over $[0,1]$ (Example images are shown in A.2).

Results from both datasets reveal consistent convergence behavior across operators. The Min-Max and Hamacher operators show rapid early changes followed by fast convergence. The multi-linear and Product operators follow a similar pattern, but with a slower decay.

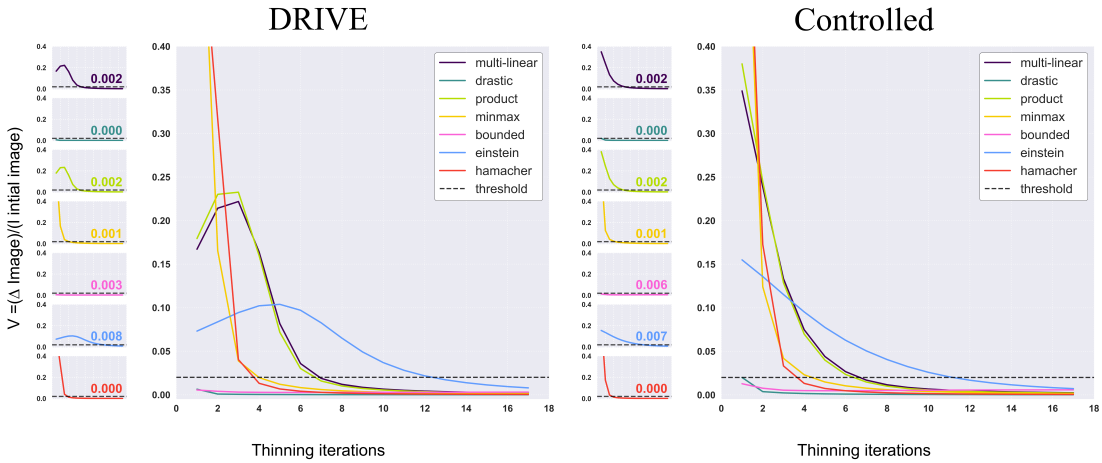


Figure 4.5.: Percentage of pixel-wise change between thinning operations Δ_t, evaluated on the DRIVE dataset and a controlled synthetic dataset.

In contrast, the Einstein operator shows a more gradual reduction in change magnitude across iterations. The Drastic and Bounded operators produce minimal per-iteration change, likely due to their restrictive formulations.

Empirically, all operators converge below 2% pixel-wise change within the first 15 iterations. Based on this observation, we set the stopping criterion at this value for all subsequent experiments.

4.4.3 Validation of SoftMorph on binary images

We evaluate the reliability of our designed probabilistic morphological operators on binary images, as summarized in Table 4.3. The primary objective is to ensure that our probabilistic operators replicate accurately the binary operators. The reference corresponds to the non-differentiable morphological operations for erosion, dilation and skeletonization, from the widely used scikit-image package [Van der Walt, 2014]. We do not assess the opening and closing operations as they correspond to iterative applications of erosion and dilation. Additionally, we compare the performance of our operators against other existing differentiable morphological operators on binary images. For the erosion and dilation operations, we test the max and min-pooling layers with a kernel of 3x3 to simulate a $k = 8 + 1$ SE. For skeletonization, we compare our method with the *soft-skeleton* approach from the clDice paper [Shit, 2021] (corresponding to a combination of min and max-pooling layers) and the method by Menten et al. [Menten, 2023] (corresponding to the convolutional layers with specific kernels to detect simple points). In 2D, we also assess a neural network model trained for skeletonization from [Nguyen, 2021]. Each method is evaluated on 15 randomly selected images in 2D from the DRIVE dataset and in 3D with the VesSap dataset.

Our method precisely replicates the erosion and dilation operations of the reference, whereas the min and max-pooling layers demonstrate significant topological errors when the SE of the operations is defined on the $k = 4+1$ neighborhood. Especially, the soft min-pooling approach exhibits an average $\beta 0$ absolute error of 214.10. These discrepancies do not arise when the reference SE is set to $k = 8 + 1$. This is because these pooling layers are defined with a 3x3 kernel that corresponds to the $k = 8 + 1$ SE. Because they are limited to square SEs, they can not replicate the $k = 4 + 1$ SE as defined in 4.1a.

The skeletons produced by our method and from Menten et al.are both topologically accurate. However, the soft-skeleton from clDice and the trained model show high topological errors. It is important to note that the Dice score is not an optimal metric to assess skeletonization performances compared with the reference. Multiple valid skeletons can be derived from the same initial volume, making the definition of the centerline ambiguous, for example when the width of the object is set on an even number of pixels. Therefore, the Dice scores from our method and Menten et al. indicate similarities with the reference with values of 0.65 in 2D and 0.71 in 3D.

Although our probabilistic operators are considerably slower than other methods, this trade-off ensures topological correctness. These results demonstrate that our probabilistic morphological operators are accurately designed and can replicate the exact traditional binary morphological operators, whereas most other differentiable methods lack topological reliability or the adaptability to scale to various morphological operations.

Table 4.3.: Validation of differentiable morphological operators in a binary setting to verify the replication of the binary operators.

			$\beta 0 \downarrow$	$\beta 1 \downarrow$	**Euler** $\downarrow$	**Dice** $\uparrow$
DRIVE 2D dataset	Dilation ($k = 4 + 1$)	*Ours*	0	0	0	1
		maxpooling 3x3	0.10 $\pm$0.31	6.65 $\pm$6.64	6.75 $\pm$6.60	0.95 $\pm$0.002
		maxpooling 2x2	0.10 $\pm$0.31	4.05 $\pm$3.10	4.05 $\pm$3.05	0.92 $\pm$0.003
	($k = 8 + 1$)	*Ours*	0	0	0	1
		maxpooling 3x3	0 $\pm$0	0 $\pm$0	0 $\pm$0	1 $\pm$1
		maxpooling 2x2	0.20 $\pm$0.41	7.40 $\pm$8.54	7.50 $\pm$8.59	0.90 $\pm$0.0031
	Erosion ($k = 4 + 1$)	*Ours*	0	0	0	1
		minpooling 3x3	214.10 $\pm$98.27	7.60 $\pm$3.69	211.80 $\pm$92.64	0.85 $\pm$0.02
		minpooling 2x2	204.15 $\pm$55.66	5.50 $\pm$2.42	209.65 $\pm$55.69	0.82 $\pm$0.01
	($k = 8 + 1$)	*Ours*	0	0	0	1
		minpooling 3x3	0	0	0	1
		minpooling 2x2	86.75 $\pm$93.79	13.10 $\pm$5.07	87.25 $\pm$102.54	0.74 $\pm$0.02
	Skeleton	*Ours*	0	0	0	0.65 $\pm$0.02
		Neural Network	206.15 $\pm$19.66	22.30 $\pm$14.27	226.85 $\pm$30.40	0.77 $\pm$0.02
		Soft-skeleton	1414.20 $\pm$191.19	66.50 $\pm$15.63	1480.70 $\pm$203.26	0.65 $\pm$0.01
		Menten et al.	0	0	0	0.65 $\pm$0.01
VesSap 3D dataset	Skeleton	*Ours*	0	0	0	0.72 $\pm$0.002
		Soft-skeleton	8362.20 $\pm$360.33	8.80 $\pm$4.30	8371.00 $\pm$358.99	0.64 $\pm$0.003
		Menten et al.	0	0	0	0.71 $\pm$0.001

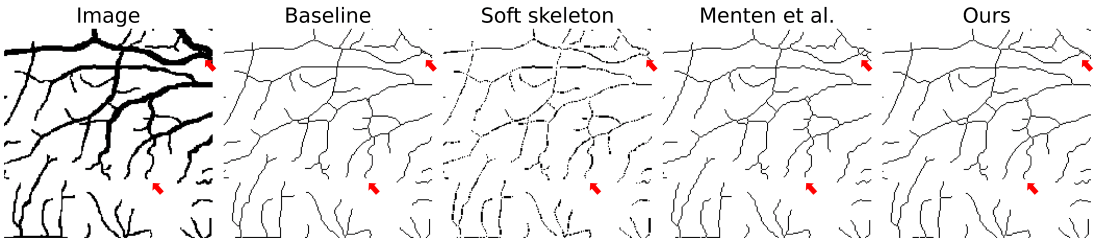


Figure 4.6.: Comparison of 2D differentiable skeletonization methods in a binary setting. Red arrows highlight significant differences.

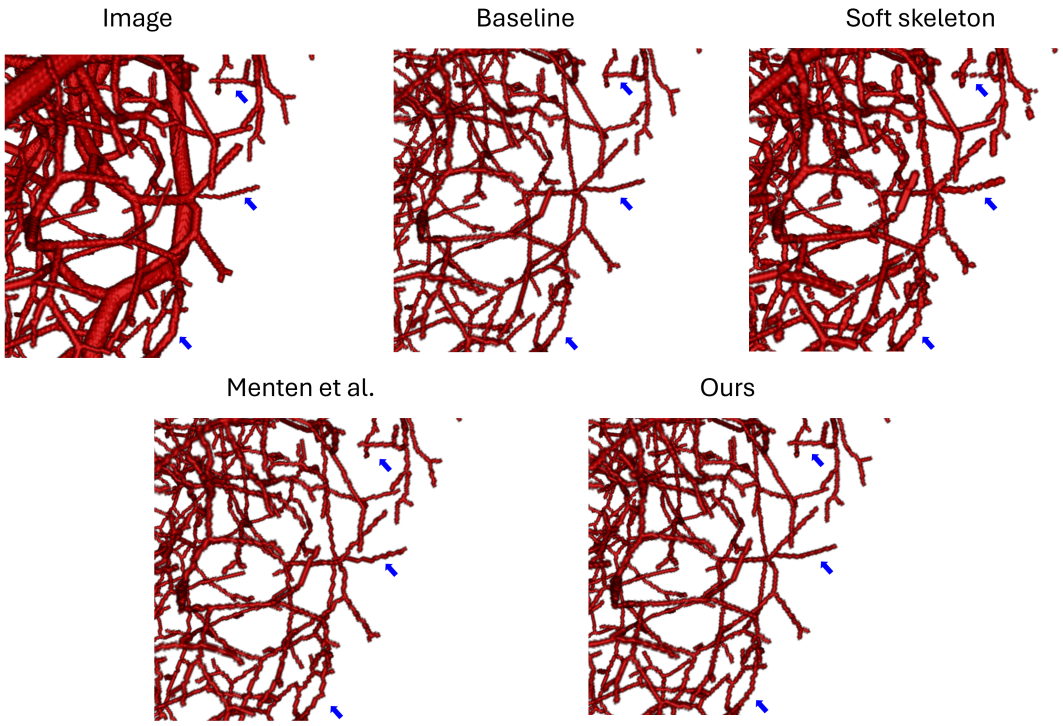


Figure 4.7.: Comparison of 3D differentiable skeletonization methods in a binary setting. Blue arrows highlight significant differences.

4.4.4 Backpropagation capability of SoftMorph

We quantify the backpropagation capabilities of the probabilistic and quasi-probabilistic representations to evaluate their impact on gradient computation for optimization within neural networks. This experiment is similar to the one described in [Menten, 2023]. We initialize a tensor with random values that we pass through a morphological operator. The soft-Dice is used to compute the loss between the operation's output and a ground truth image that is also passed through the same operator. The propagation of gradients enables the adjustment of the tensor's values until the operation's output converges towards that of the ground truth image. The experiment is performed for dilation, erosion (with the SE $k = 4 + 1$) and skeletonization operations, converging with a learning rate set to 1 over 20 epochs. The operators compared are the probabilistic multi-linear polynomial operator, and the family of proposed quasi-probabilistic operators based on fuzzy logic. We record the loss values at each epoch, as shown on Fig 4.8 to compare the convergence speed and performances of the different operators.

For dilation, the Drastic and Bounded logic-based operators fail to facilitate gradient backpropagation across all epochs. The other operators converge around the 15th epoch, with the Product and Einstein operators reaching the best final loss values of 0.01 and 0 respectively. For erosion, we observe the same trend for the Drastic and Bounded logic operators, whereas all other operators converge to the exact ground truth operation output values after around 8 epochs only. In the skeletonization operation, the Drastic and Bounded logic operators achieve the best convergence after 10 epochs, fully learning the operation's output. The Hamacher and Product operators follow, reaching a loss of approximately 0.10, with the Einstein and multi-linear operators trailing at 0.20, 0.24. The Min-Max operator obtains a Dice loss of 0.47 after 20 epochs, although the Eintein, multi-linear and Min-max do not appear to have fully converged within this time frame.

Overall, the Product-based operator emerges as the most stable operator representation, allowing an efficient convergence across all operations. The Hamacher and Einstein operators also perform effectively but are less consistent than the Product operator. Finally, the Drastic and Bounded logic exhibit unique behavior: their very restricted formulations seem to hinder the backpropagation in simpler operations like erosion and dilation, yet allowing effective gradient backpropagation in the more complex skeletonization operation, compensating for their inherent sparsity and rigidity. Therefore, most of these operators effectively support gradient backpropagation for the optimization of CNNs. The Product-based logic is recommended for such use due to its consistent performance.

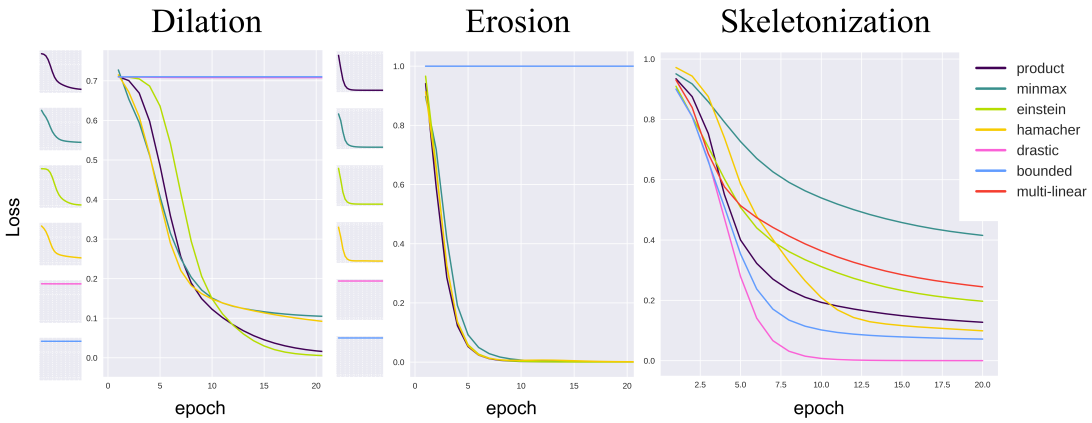


Figure 4.8.: Results of the backpropagation capability between the SoftMorph probabilistic and quasi-probabilistic operators for dilation, erosion and skeletonization.

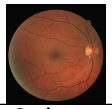
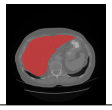
4.4.5 Applications

Final morphological layer for segmentation

We test the integration of the probabilistic morphological operators into a CNN's architecture, specifically in the U-net model [Ronneberger, 2015] for semantic segmentation. In our implementation, we add the Product-based quasi-probabilistic morphological operator as the final layer of the network, following a sigmoid activation function. The operator is applied twice to enhance its effect on the network. We use a simple U-Net implementation in 2D using the soft-dice loss, a batch size of 16, and a learning rate of $1\mathrm{e}{-3}$ over 250 epochs. For the 3D Liver dataset, we use the nnUNet model [Isensee, 2021]. The training is conducted using the soft-dice and cross-entropy loss function, with an automatically determined batch size of 2 over 1000 epochs. The dataset is split into 80% for cross-validation (65% training, 15% validation) and 20% for testing. The operators tested include the erosion, dilation, opening and closing operators with a SE of $k = 4 + 1$ in 2D and $k = 6 + 1$ in 3D as the final network layers. We compare these configurations against the baseline model without any morphological operator. The averaged results on the test set are presented in Table 4.4. The clDice metric is only measured for the tubular structures (DRIVE, Croissant and Massachusetts datasets).

Our results indicate a tendency for topological improvements across all datasets with the application of morphological operators. The $\beta 0$ absolute error is significantly lower with the inclusion of the morphological operators while maintaining a high Dice score on the DRIVE dataset. Euler's number absolute error is also significantly reduced with all but the Erosion operator, and the $\beta 1$ with the Dilation. The clDice is also maintained in the tubular structures. Examples are shown in Appendix A.2. However, no definitive pattern emerges to predict which operator can yield the best performances based on the data characteristics. We conclude that integrating morphological operators as final layers of

Table 4.4.: Double morphological operators as final U-Net layers. Bold indicates improved performance over baseline; blue cells (*) mark statistically significant improvements using the Wilcoxon rank test.

DATASET	Final layer	$\beta 0 \downarrow$	$\beta 1 \downarrow$	$\beta 2 \downarrow$	Euler $\downarrow$	Dice $\uparrow$	clDice $\uparrow$
DRIVE (2D)	*Normal*	36.00 ±7.26	33.80 ±15.75	-	69.90 ±16.37	0.84 ±0.02	0.89 ±0.03
	*Opening*	**24.10*** ±6.29	**30.95** ±16.12	-	**55.05*** ±16.41	0.84 ±0.02	0.89 ±0.03
	Closing	**25.45*** ±8.49	**31.45** ±13.80	-	**56.90*** ±11.40	0.84 ±0.02	0.89 ±0.03
	Erosion	**28.20*** ±8.58	33.90 ±15.50	-	**62.10** ±20.03	0.84 ±0.02	0.89 ±0.03
	Dilation	**27.65*** ±8.31	**21.30*** ±14.60	-	**48.55*** ±16.37	0.82 ±0.01	0.89 ±0.03
Croissant	*Normal*	3.30 ±2.52	3.55 ±5.42	-	4.55 ±5.56	0.57 ±0.32	0.56 ±0.32
	*Opening*	3.30 ±3.06	5.15 ±10.82	-	6.55 ±9.48	**0.60** ±0.33	**0.59** ±0.33
	Closing	**1.80*** ±2.19	**3.40** ±6.06	-	**3.90** ±7.30	**0.60** ±0.34	**0.58** ±0.34
	Erosion	7.15 ±4.82	4.45 ±6.16	-	5.60 ±5.59	0.57 ±0.29	0.53 ±0.27
	Dilation	**2.85** ±2.54	4.25 ±5.38	-	4.80 ±5.45	**0.58** ±0.32	**0.57** ±0.32
Massachusetts	*Normal*	31.00 ±29.18	30.71 ±31.66	-	61.47 ±49.27	0.57 ±0.12	0.64 ±0.13
	*Opening*	**26.71** ±19.58	**29.65** ±31.84	-	**56.35** ±43.30	0.54 ±0.10	**0.67** ±0.14
	Closing	**23.29** ±17.55	**29.82** ±31.43	-	**53.00** ±41.32	**0.58** ±0.13	**0.65** ±0.15
	Erosion	**20.41** ±18.04	31.24 ±32.17	-	**50.94** ±42.01	0.56 ±0.11	0.64 ±0.13
	Dilation	**19.29** ±15.70	**20.71** ±26.85	-	**37.29*** ±29.24	0.54 ±0.11	**0.69** ±0.15
Butterfly	*Normal*	1.45 ±1.50	1.25 ±1.68	-	1.80 ±1.47	0.87 ±0.25	-
	*Opening*	**1.20** ±1.32	**0.70** ±1.22	-	**1.40** ±1.64	0.86±0.0.27	-
	Closing	1.45 ±1.90	1.60 ±2.54	-	2.45±3.55	0.86 ±0.23	-
	Erosion	1.90 ±1.80	2.55 ±3.39	-	2.75 ±2.65	**0.88** ±0.17	-
	Dilation	2.55 ±2.82	3.30 ±4.66	-	2.65 ±2.89	0.86 ±0.20	-
Sea turtle	*Normal*	6.00 ±4.83	2.60 ±2.37	-	4.80 ±3.49	0.54 ±0.28	-
	*Opening*	**5.65** ±4.49	3.20 ±3.02	-	**4.55** ±3.65	**0.56** ±0.28	-
	Closing	**3.20** ±2.38	**1.90** ±2.59	-	**3.70** ±3.51	0.41 ±0.30	-
	Erosion	9.15 ±9.30	2.70 ±3.21	-	8.25 ±6.88	**0.57** ±0.26	-
	Dilation	**5.70** ±5.02	3.90 ±3.73	-	**4.50** ±2.56	0.54 ±0.28	-
Starfish	*Normal*	4.1 ±6.89	2.9 ±3.70	-	5.3 ±6.21	0.60 ±0.31	-
	*Opening*	5.35 ±4.46	**2.6** ±3.39	-	5.45 ±5.16	0.55 ±0.32	-
	Closing	4.75 ±6.09	2.95 ±3.44	-	5.40 ±4.51	0.54 ±0.30	-
	Erosion	5.00 ±4.93	**2.15** ±2.92	-	5.65 ±5.48	0.55 ±0.35	-
	Dilation	4.30 ±4.79	3.25 ±4.73	-	**4.05** ±4.61	0.53 ±0.37	-
Liver (3D)	*Normal*	2.81 ±3.84	3.19 ±4.29	2.48 ±3.57	5.15 ±4.37	0.94 ±0.07	-
	*Opening*	3.00 ±3.16	**3.07** ±4.57	**2.07** ±3.21	**4.15** ±3.16	0.92 ±0.07	-
	Closing	**1.85** ±2.14	**3.07** ±2.80	**1.33** ±2.32	**4.04** ±4.47	0.92 ±0.07	-
	Erosion	**1.63** ±1.88	3.70 ±4.58	**1.63** ±3.19	**3.85** ±3.37	0.92 ±0.07	-
	Dilation	**1.19** ±1.55	**2.59**±3.86	**2.04** ±3.24	**3.00** ±2.13	0.92 ±0.07	-

a U-net architecture for segmentation tasks can considerably improve the topological accuracy. Currently, determining the most beneficial morphological operation for a specific dataset is based only on a trial-and-error approach.

clDice loss function

We test the integration of the probabilistic morphological operators into a loss function. The loss corresponds to the clDice loss [Shit, 2021] designed to improve the topological preservation of tubular structure segmentation. This loss is combined with the SoftDice loss, weighted by a parameter α. We compare the segmentation performance on the 2D DRIVE dataset with the SoftDice alone, and the combination of clDice and SoftDice with $\alpha = 0.5$ and $\alpha = 0.7$. In clDice, it is required to extract the skeleton from both the prediction and the ground truth. To do so, we test the soft-skeleton from clDice [Shit, 2021], the method from Menten et al., and our Product-based quasi-probabilistic skeleton operator. We also evaluate the effects of dilation and erosion as final layers (with the $k = 4 + 1$ SE) of the network in conjunction with the clDice loss computed with our

Table 4.5.: 2D U-Net segmentation performance on DRIVE dataset. Blue cells (*) denote statistically significant differences with soft-skeleton and Menten et al. methods having the same α value using the Wilcoxon rank test.

LOSS \METRICS			$\beta 0 \downarrow$	$\beta 1 \downarrow$	Euler $\downarrow$	Dice $\uparrow$	clDice $\uparrow$
BCE			48.55 ±13.52	40.75 ±15.16	89.30 ±19.08	0.81 ±0.03	0.85 ±0.04
SoftDice (baseline)			45.35 ±6.58	35.15 ±16.32	80.50 ±17.99	0.82 ±0.02	0.86 ±0.03
clDice	α=0.7	*Soft-Skeleton*	17.90 ±4.63	**27.2** ±14.54	44.60 ±14.92	0.79 ±0.01	0.88 ±0.02
		Menten et al.	18.15 ±5.14	38.40 ±14.20	56.55 ±14.37	0.76 ±0.04	0.85 ±0.03
		Ours	**14.70** ±4.37	27.95 ±15.28	**42.65** ±14.68	**0.81**±0.01	**0.89** ±0.02
	α=0.5	*Soft-Skeleton*	24.70 ±4.69	**32.05** ±14.57	56.75 ±15.41	0.82 ±0.02	0.88 ±0.03
		Menten et al.	25.20 ±6.34	34.35 ±15.10	59.55 ±15.75	0.78 ±0.04	0.87 ±0.03
		Ours	**16.50*** ±5.72	32.90 ±15.04	**49.40** ±16.83	**0.83** ±0.02	**0.89*** ±0.03
clDice + final layer	Ours α=0.7	Dilation	11.30* ±4.74	**30.35** ±15.16	**41.65** ±16.94	0.80 ±0.02	**0.90*** ±0.03
		Erosion	**10.15*** ±4.33	32.75 ±14.49	42.90 ±15.87	**0.82*** ±0.01	0.89 ±0.03
	Ours α=0.5	Dilation	**11.10*** ±4.71	**30.55** ±15.67	**41.55*** ±16.08	0.81 ±0.02	**0.90*** ±0.03
		Erosion	**11.10*** ±4.23	36.85 ±15.09	47.95 ±15.82	**0.83** ±0.02	0.89 ±0.03

Product operator. We use a U-Net model trained with a batch size of 16 and a learning rate of $1e{-}4$ over 500 epochs.

Our results in Table 4.5 show that topological performances are improved with the clDice loss function compared to BCE and SoftDice alone. The application of clDice with our skeletonization method yields the best topological performances in terms of $\beta 0$ and Euler numbers, while also maintaining or even slightly improving the Dice and clDice scores compared to other methods. The addition of the final morphological layer further enhances these results by significantly reducing the number of topological errors.

In conclusion, our skeletonization method demonstrates the best overall performance compared to the other skeletons. The soft-skeleton method is prone to topological errors[Menten, 2023], which likely contributes to its lower performances. The lower performances with the skeleton from Menten et al. may be due to the reparametrization trick used in that method. This trick involves a sensitive set of parameters that can be hard to tune for the specific characteristics of the task and dataset. Here it is applied with the default parameters proposed by Menten et al. In this experiment, training and inference times were comparable across skeletonization methods, despite our operator being slower in standalone validation [Guzzi, 2024], suggesting its computational overhead is minimal within CNNs.

Therefore, these findings suggest that SoftMorph operators can effectively enhance topological performance in segmentation tasks by extracting accurate morphological features that can be integrated into loss functions or used as the final layer of a neural network. It can achieve improved results without requiring any parameter tuning.

4.4.6 Operator comparison

Last Layer

We compare in more detail the influence of the different fuzzy logic operators and SE when integrated as the final layers of the segmentation network following the previous experiments (Table 4.6). Specifically, we apply a double dilation layer as last layers of the Unet to segment the DRIVE dataset with each SoftMorph operators. The training is conducted over 250 epochs, with a learning rate $1\mathrm{e}{-3}$, a batch size of 16, the softDice loss, and the same train/test split as in prior experiments. We consider the Min-Max operator with two SE configurations: $k = 4 + 1$ and $k = 8 + 1$. For comparison, we also evaluate MaxPooling layers with kernel sizes 2×2 and 3×3.

Among all tested operators, the Min-Max consistently yields the best segmentation performance, maintaining high Dice and clDice scores while achieving low Betti number and Euler characteristic errors. In contrast, the Drastic, Bounded, and Hamacher operators result in poor performance, with Dice and clDice scores below 0.4, indicating insufficient gradient flow for effective optimization. The Product and Einstein operators significantly outperform the baseline in terms of topological accuracy, suggesting improved structural preservation. For the Min-Max operator, the $k = 4 + 1$ SE outperforms the $k = 8 + 1$ configuration, highlighting the impact of SE size on the learning dynamics. Surprisingly, the MaxPooling layers perform worse than Min-Max operator, even though the mathematical formulation is the same. These differences is easily explainable for the square 2x2 kernel as it does not exactly exactly replicate the $k = 4 + 1$ SE. However, the 3x3 kernel yields the exact $k = 8 + 1$ SE, and in this case, this discrepancy may result from implementation differences in PyTorch. Indeed, as stated in section 4.2.2, standard MaxPooling layers propagate gradient exclusively through the first occurrence of the maximum value within each pooling window, whereas the minimum/maximum operations distribute gradients evenly across all equal elements. Additionally, the Maxpooling layers underperform Einstein and Product-based operators overall, particularly in $\beta 1$ and Euler characteristic errors.

Loss function

We evaluate the impact of the different SoftMorph operators within the clDice loss function for the skeletonization operation, using the same training setup as the previous experiment (Table 4.6).

All operators except Hamacher lead to a statistically significant reduction in topological errors, including connected components ($\beta 0$) and Euler characteristics. The Min-Max

Table 4.6.: 2D U-Net segmentation performance on DRIVE dataset comparing SoftMorph operators, structuring element, and Pooling layers. Bold indicates improved performance compared to the baseline; Blue cells (⋆) denote statistically significant improvements using the Wilcoxon rank test.

Loss	Layer	Logic	$\beta 0 \downarrow$	$\beta 1 \downarrow$	Euler $\downarrow$	Dice $\uparrow$	clDice $\uparrow$
SoftDice	**None (Baseline)**		36.00 ±7.26	33.80 ±15.75	69.90 ±16.37	0.84 ±0.02	0.89 ±0.03
	Dilation	*Product*	**27.65**⋆ ±8.31	**21.30**⋆ ±14.60	**48.55**⋆ ±16.37	0.82 ±0.01	0.89 ±0.03
	$k = 4 + 1$	*Einstein*	**15.75**⋆ ±3.93	**28.40** ±15.21	**44.15**⋆ ±16.14	0.83 ±0.02	0.89 ±0.04
		Min-Max	**16.40**⋆ ±4.60	**17.95**⋆±14.40	**17.75**⋆ ±12.27	0.82 ±0.02	0.88 ±0.03
		Drastic	**0.75**⋆ ±0.79	65.50 ±15.63	65.55 ±15.45	0.32 ±0.03	0.01 ±0.04
		Bounded	**0.95**⋆ ±1.10	67.50 ±15.63	66.55 ±15.45	0.22 ±0.02	0.04 ±0.18
		Hamacher	**0.95**⋆ ±1.10	67.50 ±15.63	66.55 ±15.45	0.22 ±0.22	0.04 ±0.18
		MaxPool $2x2$	**23.50**⋆ ±6.11	35.00 ±15.34	**58.50**⋆ ±15.02	0.83 ±0.01	0.89 ±0.03
	$k = 8 + 1$	*Min-Max*	42.90 ±8.37	36.65 ±15.19	79.55 ±18.00	0.79 ±0.01	0.87 ±0.04
		MaxPool $3x3$	61.90 ±16.10	38.20 ±16.83	100.10 ±21.23	0.78 ±0.01	0.87 ±0.04
clDice	**None**	*Product*	**21.90**⋆ ±5.81	**30.60** ±15.86	**52.50**⋆ ±16.92	0.83 ±0.02	**0.90** ±0.03
$\alpha = 0.5$		*Einstein*	**15.80**⋆ ±3.72	**33.25** ±17.17	**49.05**⋆ ±17.73	0.83 ±0.02	**0.90** ±0.03
		Min-Max	**13.25**⋆ ±3.74	**30.45** ±13.65	**43.70**⋆ ±14.30	0.83 ±0.01	0.89 ±0.03
		Drastic	**23.00**⋆ ±6.12	34.65 ±14.68	**57.65**⋆ ±14.27	0.83 ±0.02	0.89 ±0.03
		Bounded	**15.45**⋆ ±6.05	35.55 ±14.32	**51.00**⋆ ±15.86	0.83 ±0.02	**0.90** ±0.03
		Hamacher	**0.95**⋆ ±1.10	64.35 ±15.51	63.30 ±15.33	0.27 ±0.03	0.22 ±0.07

operator achieves the best overall topological performance, with the lowest error in Euler number (43.70) and $\beta 0$ (13.25). Contrary to the previous dilation experiment, the Bounded and Drastic operators are functional in this experiment, yielding significant improvements compared to the baseline. These results coincide with the backpropagation experiment of section 4.4.4. The Hamacher operator, consistent with earlier observations, failed to facilitate effective learning. It resulted in a poor Dice score (0.27) and clDice (0.22), indicating low segmentation accuracy.

4.5 Discussion

In this work, we have demonstrated that SoftMorph operators can be effectively integrated into deep learning pipelines, either as last layers of a segmentation network or directly within loss functions requiring a morphological operation. Their integration showed topological improvements across multiple segmentation tasks compared to the baseline and other existing differentiable morphological operators. Our experiments reveal key considerations to adopt the SoftMorph framework:

Regarding the choice of morphological operation to apply as the final layer in a segmentation network, dilation delivered the best performance across the evaluated tasks. However, this difference is marginal in some applications where other operations showed comparable results. These differences suggest that the optimal morphological operation is task-dependent and should be selected empirically.

Secondly, the size and shape of the SE play a crucial role. In our evaluations, the $k = 4+1$ SE in 2D and $k = 6+1$ SE in 3D provided the most consistent improvements in topological accuracy. Nevertheless, these results may not generalize universally, and tuning the SE

based on the characteristics of the target structure may have an important impact on performance.

Lastly, all SoftMorph operators can be used interchangeably in terms of interface and integration. However, their gradient propagation characteristics vary substantially. Specifically, the Drastic and Bounded operators were not able to backpropagate gradients for simple morphological operations like dilation and erosion, but were very effective for more complex operations such as skeletonization. Accordingly, they enabled significant segmentation improvements when integrated for the skeletonization operation of the clDice loss compared to the dilation and erosion layers applications. Conversely, the Hamacher operator theoretically supports gradient flow but exhibits poor empirical performance in our segmentation experiments. The Min-Max, Product, and Einstein operators consistently provided both strong topological improvements and reliable gradient propagation. Notably, the Min-Max operator yielded the best results overall and should be considered a default choice when converting traditional morphological operations into differentiable forms. While conventional min-pooling and max-pooling layers can replicate erosion and dilation, respectively, they lack the flexibility to generalize to more complex morphological operations. Furthermore, their performance was consistently inferior to the Product, Einstein, and Min-Max SoftMorph operators, highlighting the practical advantages of our approach.

The transition from a probabilistic to a fuzzy interpretation with the quasi-probabilistic operators corresponds to a change in semantics rather than an exact equivalence. In the probabilistic formulation, input values represent Bernoulli probabilities and morphological operators are defined as expectations over all binary configurations. In contrast, fuzzy logic interprets values as degrees of membership and combines them deterministically using T/S-norms. In this work, fuzzy operators are not introduced as probabilistic models, but as computationally tractable surrogates that preserve key properties of the probabilistic operators, with an exact agreement on binary inputs, bounded outputs, and smooth interpolation for intermediate values. Among these, the product T-norm is of particular interest as it coincides with probabilistic conjunction under independence assumptions. The evaluation of alternative fuzzy logics, including min/max, bounded, and drastic norms, was conducted to assess the robustness of the proposed framework beyond the product logic.

For new applications, we recommend starting with the Product, Min-Max, or Einstein operators, and empirically testing different SEs and operations tailored to the specific task. Meanwhile, Drastic and Bounded operators, while limited in simple operations, can be highly effective for more structurally complex transformations like skeletonization. One can use SoftMorph to design new morphological operations based on a logic expression, specific for an image analysis task, and that needs to be integrated into deep learning segmentation networks.

4.6 Conclusion

We presented SoftMorph, a family of differentiable probabilistic and quasi-probabilistic morphological operators for deep learning frameworks. SoftMorph successfully bridges the gap to translate any morphological operation defined on any SE in its soft counterpart while maintaining the computational complexity of the original binary operator. We have demonstrated that probabilistic operators can be defined as the expectation of the binary operator and represented as a multi-linear polynomial. We have also shown that the factorized form of the original binary operator can be approximated as a quasi-probabilistic operator using fuzzy logic. Besides, these operators replicate the exact output of binary operators on binary images while enabling gradient-based optimization and handling probabilistic maps. Integrating some basic morphological operations can improve the topological performances of segmentation networks when inserted as the final layer or within loss functions. Possible improvements include the optimization of the Boolean 2D and 3D skeletonization algorithms, the definition of morphological operations with non-flat (fuzzy) SEs, and learning new morphological operations and their associated optimal SE. This work focuses on binary morphological operators, as they are naturally expressed through Boolean formulas and thus admit a probabilistic extension via expectation over binary configurations. Grayscale morphological operators, by contrast, are defined directly on complete lattices of real-valued functions and already possess continuous formulations. Extending the proposed probabilistic framework to such operators would therefore require a different theoretical construction and is left for future work. Overall, this work opens avenues for defining new task-specific morphological operations. Their application could be extended to other image analysis applications, neural network architectures, and new loss functions requiring the extraction of morphological features.

5 Regional Hausdorff Distance Losses for Medical Image Segmentation

Contents

Building on the SoftMorph framework introduced in the previous chapter, which enabled the integration of differentiable morphological operations into convolutional neural networks, we next explored its potential for defining more reliable distance-based loss functions. In particular, we focused on the Hausdorff Distance, a metric that captures spatial boundary errors and is especially relevant for small or thin structures where overlap-based measures such as the Dice coefficient can be less informative.

Using SoftMorph, we implemented a fully differentiable erosion-based distance transform that can be applied directly to probability maps. This allowed us to compute positive, signed, and unsigned distance maps without relying on binary inputs, enabling stable optimization within neural network training. From the positive distance transform, we derived three novel formulations of regional Hausdorff Distance losses: the maximum, modified, and averaged, encouraging precise boundary alignment without the need for auxiliary losses.

We validated these loss functions on multiple public medical image segmentation datasets, demonstrating that they substantially reduce distance-related metrics while maintaining,

and in some cases improving, performance on standard overlap metrics. This effect was particularly notable for small lesions such as small tumors, highlighting the relevance of this approach to the segmentation challenges addressed in this thesis for the small structures with hard-to-define boundaries such as calcifications or stents.

The work presented in this chapter was accepted at the 16th International Workshop on Machine Learning in Medical Imaging (MLMI 2025), held in conjunction with MICCAI 2025.

To foster reproducibility, our code is publicly available on GitHub[1]

5.1 Introduction

The Hausdorff Distance (HD) metric quantifies the distance between two point sets [Huttenlocher, 1993]. In medical image analysis, it is widely used to assess the segmentation performance of deep learning models [Crum, 2006; Taha, 2015]. Unlike the Dice Similarity Coefficient (DSC), which only measures the volumetric overlap between a predicted and the ground truth segmentation, the HD metric offers a more sensitive assessment of boundary discrepancies by explicitly capturing the maximal distance between misaligned parts. This makes it particularly relevant for assessing the quality of the delineation of small structures such as tumors or lesions by highlighting the worst-case mismatch between two objects [Rizzetto, 2020], or the boundary alignment of anatomical structures like the heart or the liver [Moradi, 2019]. Despite its importance, few attempts have been reported to directly optimize the HD metric during the training of deep learning-based image segmentation.

Some studies have designed losses to optimize the boundary accuracy without explicitly reproducing the HD [Kervadec, 2021; Caliva, 2019; Yang, 2019; EL Jurdi, 2021]. The weighted Hausdorff Distance loss from [Ribera, 2019] approximates the average HD for isolated points but is designed only for points, not image masks. Only the work from [Karimi, 2019] aims to reduce the HD by approximating an HD loss, relying on the Euclidean Distance Transform (EDT) of the segmentation boundaries. They also propose an alternative loss through morphological operations, which underperformed compared to the EDT-based formulation. While this loss improved performance regarding HD, it must be combined with Dice loss to prevent any instabilities, and its formulation does not exactly replicate the HD formula. Additionally, the EDT-based loss lacks differentiability properties. Indeed, Distance Transform Maps (DTM) provide an alternative representation of a binary shape where each voxel's intensity corresponds to its distance from the nearest foreground boundary voxel [Rosenfeld, 1968]. This representation enables

[1] https://github.com/lisaGUZZI/HD-Loss

explicit distance computations between point sets. However, conventional distance transform algorithms are inherently non-differentiable [Zhang, 2022]. The differentiability of the approach in [Karimi, 2019] is not explicitly addressed, raising questions about its suitability for gradient-based optimization.

Several studies have tried approximating soft distance transforms (DT) by emulating the operation through CNN-based learning [Bai, 2017; Bui, 2019; Navarro, 2019]. For instance, in [Bai, 2017], a CNN-based approach was introduced to learn a watershed transform for instance-based segmentation. Other methods have directly regressed the DTM as a CNN output to be used as a loss function or as a regularization term [Wang, 2020; Dangi, 2019; Xue, 2020]. However, these approaches are prone to domain shift since the learned DT is dataset-dependent, and some require additional geometry-aware refinements to improve generalization. Recent works have proposed Convolutional-based differentiable Distance Transform (CDT) operations [Zhu, 2024; Zhang, 2022; Pham, 2021]. One major limitation of CDT is that the kernel size must be as large as the image diagonal to process sparse binary images correctly. If the kernel size is too small, background pixels far from the foreground will mistakenly receive a zero-distance value. This leads to increased computational complexity due to the large kernel size and potential numerical instability when the kernel's exponential term approaches zero, particularly in large images with only a few foreground pixels. A cascaded CDT approach was proposed by [Pham, 2021] to mitigate these issues, but it requires soft binarization of segmentation outputs, hence not directly applicable to raw probability maps.

In this chapter, we propose a novel family of regional HD-based loss functions that rely on a soft and differentiable DT. As a first contribution, we introduce a novel morphological erosion-based differentiable distance transform that can be applied directly on probability maps. This method allows to compute the accurate signed, unsigned, and positive distance maps in a differentiable manner, ensuring compatibility with deep learning models. In a second contribution, we derive three well-established variations of the HD loss leveraging the differentiable distance transform. Specifically, we provide a smooth formulation of the Hausdorff, Modified Hausdorff, and Symmetric Averaged Hausdorff Distances. These losses achieve state-of-the-art performance without auxiliary losses.

We analyze the effect of key hyperparameters and validate our method through controlled validation experiments. Then, we evaluate the applicability of our losses across various public 2D and 3D medical segmentation datasets, proving the applicability of the soft distance transforms and the effectiveness of the proposed loss functions to minimize the distance of the predicted masks to the ground truth while maintaining an equivalent Dice score.

5.2 Method

5.2.1 Definition of distance transforms

We define a ground truth binary segmentation X consisting of N voxels such that $X = \{X_n\}_{n=1}^{N}, \quad X_n \in \{0, 1\}$. The boundary ∂X of the foreground object in the image is defined as $\partial X = \{q \mid X_q = 1, \ \exists r \text{ such that } X_r = 0, \ dt(q, r) \leq 1\}$ where $dt(q, r)$ is a chosen distance metric between two voxels, typically the Euclidean or Manhattan distance.

The distance transform map $\mathcal{D}(X)$ assigns to each voxel X_n its minimum distance to the closest foreground boundary voxel q such that:

$$\mathcal{D}(X_n) = \min_{q \in \partial X} dt(X_n, q) \tag{5.1}$$

Eq. 5.1 actually corresponds to the unsigned distance transform $\mathcal{D}^u$ where the distance is computed for each voxel, both inside and outside the boundary ∂X. Alternatively, one can also compute the signed distance transform $\mathcal{D}^s$ such that the distance is negative inside the object and positive outside:

$$\mathcal{D}^s(X_n) = \begin{cases} \min_{q \in \partial X} dt(X_n, q), & \text{if } X_n = 0 \quad \text{(outside the object)} \\ -\min_{q \in \partial X} dt(X_n, q), & \text{if } X_n = 1 \quad \text{(inside the object)} \end{cases} \tag{5.2}$$

The positive distance transform $\mathcal{D}^+$ assigns a distance of 0 to all foreground voxels and computes the distance only for background voxels:

$$\mathcal{D}^+(X_n) = \begin{cases} \min_{q \in \partial X} dt(X_n, q), & \text{if } X_n = 0 \\ 0, & \text{if } X_n = 1 \end{cases} \tag{5.3}$$

The relationships between these distance transforms are therefore given by: $\mathcal{D}^u(X) = \mathcal{D}^+(X) + \mathcal{D}^+(1 - X)$ and $\mathcal{D}^s(X) = \mathcal{D}^+(X) - \mathcal{D}^+(1 - X)$. Also, $\mathcal{D}^u(X) = |\mathcal{D}^s(X)|$ and $\mathcal{D}^+(X) = \max(\mathcal{D}^s(X), 0)$.

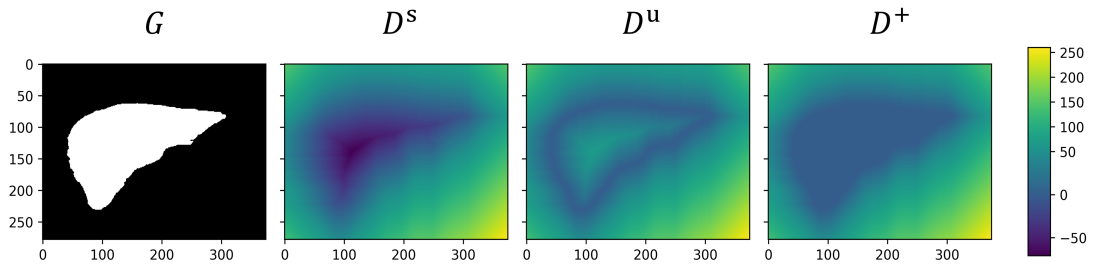


Figure 5.1.: Illustration of distance transforms $\mathcal{D}^s$, $\mathcal{D}^u$, $\mathcal{D}^+$ on a binary image X.

5.2.2 Distance transforms with morphological operations

The positive distance transform $\mathcal{D}^+$ can be computed with morphological operations, specifically using successive erosions, when morphological distances are considered. The maximum distance corresponds to the number of erosions required to completely remove the background. The erosion filter $\ominus$ is defined by an isotropic structuring element k (typically $k = 4 + 1$ or $k = 8 + 1$ in 2D, and $k = 6 + 1$, $k = 18 + 1$ or $k = 26 + 1$ in 3D) which defines the unit ball of the considered metric. The erosion of shape X with the k structuring element is written as $X \ominus k$, and we note $X \ominus^i k$ its ith eroded shape.

We propose to compute the positive distance transform $\mathcal{D}^+$ as:

$$\mathcal{D}^+(X) = \sum_{i=1}^{I} [(1 - X) \ominus^i k] \tag{5.4}$$

where I represents the number of iterations required to compute the complete distance map, or in other words, the number of successive erosions to perform to remove the background. Therefore, each background voxel receives a value equal to the number of erosions required to remove the shape complement $1 - X$. The other distance transforms $\mathcal{D}^u$ and $\mathcal{D}^s$ can then be derived from their respective relationships to $\mathcal{D}^+$.

Algorithm 2 Positive Distance Transform with Morphological Erosions

```
procedure PositiveDistance(img, I, k)
    D+ ← 0                                   ▷ Initialize distance map
    img_temp ← 1 − img                       ▷ Work on background complement
    for i ← 1 to I do
        D+ ← D+ + img_temp
        img_temp ← Erosion(img_temp, k)
        if sum(img_temp) = 0 then
            break                            ▷ Stop if background fully eroded
        end if
    end for
    return D+
end procedure
```

5.2.3 Soft distance transforms

Neural networks typically output a segmentation $\mathcal{Y}$ of an image X corresponding to the probability map $\mathcal{Y} = \{y_n\} \in [0,1]^N$ with $y_n = p(X_n = 1|X) \in [0,1]$. We want to formalize the $\mathcal{D}^u$, $\mathcal{D}^s$ and $\mathcal{D}^+$ distance transforms to their probabilistic equivalents d^u, d^s and d^+ which can be applied on the probabilistic segmentation $\mathcal{Y}$. The probabilistic distance transform d should generalize to the discrete case $\mathcal{D}(Y) = d(Y)$ such that both give the same output when Y is binary.

Given the soft and differentiable morphological filters introduced in the previous chapter and corresponding publication [Guzzi, 2024], we propose to compute the soft distance transform by replacing the binary erosion filter $\ominus^i k$ in Eq. 5.4 with a soft erosion filter $\ominus_s^i k$ applied on the probability segmentation $\mathcal{Y}$. The soft erosion $\ominus_s$ is expressed as the multilinear polynomial $(\mathcal{Y} \ominus_s k)_n = \prod_{j \in \mathcal{N}^k(n)} \mathcal{Y}_j$, where $\mathcal{N}^k(n)$ are the neighboring pixels of voxel n according to the selected structuring element. This soft erosion operator makes the soft distance transforms applicable on probability maps and fully differentiable.

5.2.4 Regional Hausdorff Distance Losses

The Hausdorff Distance (HD) quantifies the maximum discrepancy between two point sets. Given the point sets A and B, the HD is defined as:

$$HD = \max(\max_{a \in A} dt(a, B), \max_{b \in B} dt(b, A))$$

where the directed distance of a point a to the set B is given by $dt(a, B) = \min_{b \in B} dt(a, b)$. This formulation of the HD satisfies the triangle inequality property. However, its sensitivity to outliers has led to the introduction of alternative HD formulations.
The Modified Hausdorff Distance (MHD) [Dubuisson, 1994] averages the minimal distance instead of taking the maximum:

$$MHD = \max\left(\frac{1}{|A|}\sum_{a \in A} dt(a, B), \quad \frac{1}{|B|}\sum_{b \in B} dt(b, A)\right) \tag{5.5}$$

The Symmetric Averaged Hausdorff Distance (MHD_{sym}) [Dubuisson, 1994] takes the mean of the two directed distances:

$$MHD_{sym} = \frac{1}{2}\left(\frac{1}{|A|}\sum_{a \in A} dt(a, B) + \frac{1}{|B|}\sum_{b \in B} dt(b, A)\right) \tag{5.6}$$

The MHD is more robust in matching objects based on edge points than the HD [Dubuis-

son, 1994], while MHD_{sym} provides a smooth alternative with a slightly reduced discriminatory power.

To optimize the HD into neural network-based segmentation, we propose differentiable approximations of the HD formulas based on the positive distance transform $\mathcal{D}^+(X)$ and its probabilistic counterpart $d^+(\mathcal{Y})$.

Given a binary ground truth segmentation X and a probabilistic prediction segmentation $\mathcal{Y}$, we define:

The Hausdorff Loss (L_H), approximating the classical HD:

$$L_H = \text{smax}(\text{smax}(\mathcal{D}^+(X)_n \circ \mathcal{Y}_n), \text{smax}(d^+(\mathcal{Y})_n \circ X_n)) \tag{5.7}$$

The Averaged Hausdorff Loss (L_{AH}):

$$L_{AH} = \text{smax}\left(\frac{1}{|Y|+\epsilon}\sum_n \mathcal{D}^+(X)_n \circ \mathcal{Y}_n, \quad \frac{1}{|X|+\epsilon}\sum_n d^+(\mathcal{Y})_n \circ X_n\right) \tag{5.8}$$

The symmetric Averaged Hausdorff Loss ($L_{AH_{sym}}$):

$$L_{AH_{sym}} = \frac{1}{2}\left(\frac{1}{|Y|+\epsilon}\sum_n \mathcal{D}^+(X)_n \circ \mathcal{Y}_n + \frac{1}{|X|+\epsilon}\sum_n d^+(\mathcal{Y})_n \circ X_n\right) \tag{5.9}$$

where $\text{smax}(x)$ is a smooth approximation of the maximum operator implemented using the LogSumExp function (with a α scaling parameter), $\circ$ represents the Hadamard product and ϵ a small positive constant. The directed distance $\mathcal{D}^+(X)_n \circ \mathcal{Y}_n$ represents the false positive points of $\mathcal{Y}$ and their distance to the ground truth X whereas $d^+(\mathcal{Y})_n \circ X_n$ is the directed distance representing the distance of the false negatives to X.

5.2.5 Computational complexity.

Considering that the complexity of the soft dice loss is $\mathcal{O}(b \cdot c \cdot V)$, the proposed losses have a complexity of $\mathcal{O}(I \cdot b \cdot c \cdot V)$ where I is the number of soft erosion iterations, and b,c,V denote batch size, number of channels, and number of voxels, respectively. To reduce computational overhead, we propose constraining I to a small fixed value (e.g., 2–5), resulting in a saturated positive distance transform. The truncation attenuates the influence of distant errors without removing them, maintaining meaningful gradient

signals and additionally improving numerical stability during optimization, ensuring that overly large distances do not dominate the loss.

5.3 Results

5.3.1 Experimental Setup

Datasets. We evaluated our method on five public datasets: the DRIVE dataset of 2D retinal blood vessels [Staal, 2004], the 3D CT Liver, Spleen, Pancreas and Hepatic vessels datasets from the Medical Segmentation Decathlon [Antonelli, 2022], and the ACDC dataset [Bernard, 2018] comprising multi-slice 2D cine MRI images of the heart. 3D datasets were considered as a stack of 2D slices when used in the 2D applications.

Evaluation Metrics. We evaluated segmentation performance using the Dice coefficient, conventional HD, HD95, modified HD (MHD) to verify that Hausdorff distance losses regress distance metrics, and clDice [Shit, 2021] to assess the global shape and object-level alignment.

Implementation Details. Experiments were conducted with Python 3.12.8 and Pytorch 2.5.1 on 3 Nvidia A40 PCIe GPUs.

5.3.2 Validation of Distance transforms

Our differentiable distance transforms, based on morphological operations, is sensitive to the connectivity choice. To evaluate the validity of the generated distance maps, we analyzed the variance of the gradient norm of the signed distance map computed with $\mathcal{D}^s$ under different connectivity settings. Indeed, the ideal configuration should exhibit minimal variance, as the norm of the gradient should be equal to 1 everywhere: $\|\nabla \mathcal{D}^s(X)\| = 1$. We considered signed distance maps produced with a connectivity of 4 and 8 in 2D and of 6, 18, and 26 in 3D. We additionally compared these results with i) the *distance_transform_edt* function from scipy which provides the exact distance but is not differentiable, and ii) the convolutional distance transforms from [Zhu, 2024] and [Pham, 2021]. In Table 5.1, we reported the variance averaged over a database of 50 images randomly selected from the DRIVE, Liver, Spleen, and ACDC datasets, both in 2D and 3D. In 2D, our method achieved a lower norm of the gradient variance around the value 1 than SciPy with both 8 and 4-connectivity. In 3D, the Scipy package got the smallest variances, and the 26-connectivity led to the smallest variance to 1. In the remainder, a connectivity of 4 is selected in 2D and of 26 in 3D.

Table 5.1.: Variance of the norm of the gradient of different SDT.

Modality	DT		Variance to 1 ↓	Variance to mean ↓
2D	Scipy		0.96	0.03
	CDT [Zhu, 2024]		2.80	2.16
	cascaded CDT [Pham, 2021]		0.24	0.21
	Ours	*4 connec*	0.14	**0.02**
		8 connec	**0.12**	0.12
3D	Scipy		**0.97**	**0.03**
	CDT [Zhu, 2024]		4.68	4.18
	Ours	*6 connec*	74.64	72.63
		18 connec	21.87	21.62
		26 connec	19.89	19.69

5.3.3 Evaluation on Public Datasets

2D Applications.

We used our Hausdorff loss functions to train a 2D U-Net [Ronneberger, 2015] for the segmentation of the DRIVE and Spleen datasets (Table 5.2). Datasets were split into 75% training, 15% validation, and 15% testing sets, and the network was trained with a batch size of 16 through 250 epochs and a learning rate of 1e-3 for the DRIVE dataset and of 1e-4 for the Spleen. We set the number of erosion to compute the distance transform to 5 iterations for the L_H loss and to 2 for L_{AH} and $L_{AH_{sym}}$. The α scale factor of the LogSumExp function was empirically set to $\alpha = 150$. We compared our method with models trained with the same UNET and the Dice loss as a baseline and the HD loss proposed in [Karimi, 2019].

All of our loss functions improved the HD metrics (HD, HD95, MHD) compared to the baseline (Dice loss) and the additional CE+Dice loss while maintaining equivalent Dice and clDice scores on both datasets. The Hausdorff loss from [Karimi, 2019] did not show any improvement compared to the baseline. Best results were obtained with L_{AH} for the Spleen dataset with a HD of 5.51 compared to 10.95 for the baseline, while achieving the same Dice score of 0.93. For the Drive dataset it is the $L_{AH_{sym}}$ loss that yielded best results with a HD of 14.80 compared to 16.77 for the baseline while keeping the Dice score at 0.82. Additionally, while the losses significantly improved the HD metrics on the Spleen dataset, the results were less significant regarding the retinal dataset. We believe this is because the DRIVE dataset contains images of retinal blood vessels that are dense and spread through the entire image while the Spleen contains masks that are smaller and centered in the image, meaning that the distance of false positive points could be higher and the losses correct these segmentation discrepancies more effectively.

Table 5.2.: 2D U-Net segmentation on the Spleen and DRIVE datasets.

		Dice ↑	HD ↓	HD95 ↓	MHD ↓	clDice ↑
Spleen	*Dice (baseline)*	**0.93**	10.95	10.10	5.93	**0.96**
	CE + Dice	0.93	25.11	24.34	19.60	0.96
	HD [Karimi, 2019]	0.88	15.08	13.97	11.56	0.92
	LH	0.90	6.82	5.62	3.51	0.94
	LAH	**0.93**	**5.51**	**4.50**	**0.87**	**0.96**
	LAH sym	0.92	9.66	8.88	5.32	0.95
Drive	*Dice (baseline)*	**0.82**	16.77	2.20	0.50	**0.84**
	CE + Dice	0.80	16.36	2.80	0.50	0.82
	HD [Karimi, 2019]	0.78	17.80	3.04	0.67	0.81
	LH	0.80	15.58	2.10	0.49	0.82
	LAH	0.80	15.53	1.96	0.48	0.84
	LAH sym	**0.82**	**14.80**	**1.85**	**0.45**	**0.84**

When we compared the number of erosion iterations to compute the distance map (Table 5.3 and 5.4), increasing the number of erosion for the L_H loss decreased the HD metrics in both datasets. However, the averaged losses L_{AH} and $L_{AH_{sym}}$ did not exhibit the same behavior as the performance sometimes decreases as the number of iterations increases. This could be because the LAH criterion is a regularized version of LH, which does not verify the triangular inequality and can entail extra local minima. Indeed, the normalization term by the size of the predicted mask $\mathcal{Y}$ can lead to a local minimum for slightly over-segmented predicted masks. Increasing the number of iterations to compute the distance transform might increase the chances of attaining those local minima.

Table 5.3.: 2D U-Net segmentation on the DRIVE dataset comparing the number of erosion iterations to compute the positive distance map.

		Dice ↑	HD ↓	HD95 ↓	MHD ↓	clDice ↑
Soft Dice (baseline)		0.82	16.77	2.20	0.50	0.84
LH	*1 iter*	0.81	16.59	2.38	0.56	0.83
	2 iter	0.81	15.91	2.13	0.48	0.83
	5 iter	0.80	15.58	2.10	0.49	0.82
LAH	*1 iter*	0.82	16.09	2.05	0.51	0.84
	2 iter	0.80	15.53	1.96	0.48	0.84
	5 iter	0.80	16.14	1.88	0.46	0.83
LAH sym	*1 iter*	0.82	15.68	1.96	0.48	0.84
	2 iter	0.82	**14.80**	**1.85**	**0.45**	0.84
	5 iter	0.81	16.31	**1.85**	0.47	0.84

Table 5.4.: 2D U-Net segmentation on the Spleen dataset comparing the number of erosion iterations to compute the positive distance map.

		Dice ↑	HD ↓	HD95 ↓	MHD ↓	clDice ↑
Soft Dice (baseline)		0.93	10.95	10.10	5.93	0.96
LH	*1 iter*	0.92	11.02	10.06	7.52	0.95
	2 iter	0.92	10.75	8.68	4.25	0.95
	5 iter	0.90	6.82	5.62	3.51	0.94
LAH	*1 iter*	0.92	8.76	6.74	5.08	0.96
	2 iter	0.93	5.51	4.50	**0.87**	0.96
	5 iter	0.92	6.84	5.87	2.12	0.95
LAH sym	*1 iter*	0.92	**5.37**	**4.40**	2.78	0.95
	2 iter	0.92	9.66	8.88	5.32	0.95
	5 iter	0.93	12.10	11.29	7.61	0.95

3D Applications.

We used our Hausdorff loss functions to train a 3D nnUNet [Isensee, 2021] for the segmentation of the Pancreas (Table 5.5) and Hepatic vessels (Table 5.6) datasets in 3d full resolution. Each dataset has two labels: *Pancreas* and *Cancer* for the Pancreas dataset and *Vessels* and *Tumor* for the Hepatic vessels. Datasets were split into 65% training, 15% validation, and 20% testing sets, and the network was trained with a batch size of 12 through 250 epochs and an initial learning rate of 1e-2. We set the number of erosion to compute the distance transform to 2 iterations in every case. The α scale factor of the LogSumExp function was still set to $\alpha = 150$. We compared our methods with an nnUNet trained with the Dice loss as baseline. We also investigated the integration of the CE loss with every Hausdorff loss and the baseline (CE + Dice).

The three Hausdorff losses consistently improved the distance metrics across all labels, with the largest gains observed for the cancer and tumor classes. An exception was the MHD with L_H, which showed only marginal improvement in the Pancreas dataset and no improvement for the Hepatic Vessels. This behavior can be explained by the fact that the other two losses are directly aligned with the MHD formulation, making them more effective for optimizing this metric. On the Pancreas dataset, clDice was improved for both labels. The Dice score was also improved using L_{AH} compared to the baseline. The cancer and tumor labels exhibited the most substantial improvements in both datasets, including in Dice. This effect is likely due to their structure, as they typically consist of multiple small, disconnected components. Such structures are difficult to capture using overlap-based losses like Dice alone, whereas distance-based losses are better suited to recover the precise locations of these small objects. This advantage is also evident in the qualitative examples (Figure 5.2 and Figure 5.3).

Table 5.5.: 3D nnU-Net segmentation on the Pancreas dataset.

		Dice ↑	HD ↓	HD95 ↓	MHD ↓	clDice ↑
Pancreas	*Dice (baseline)*	0.77	25.14	13.37	524.49	0.72
	LH	0.77	25.22	12.22	519.86	0.73
	LAH	0.80	24.65	10.43	282.03	0.75
	LAH sym	**0.82**	**16.94**	**7.10**	**221.80**	**0.79**
	CE + Dice (baseline)	**0.81**	17.63	7.08	162.20	0.78
	CE + LH	0.80	**16.96**	6.74	261.21	0.77
	CE + LAH	0.80	18.69	8.16	314.91	0.77
	CE + LAH sym	**0.81**	17.10	**6.09**	**161.46**	**0.79**
Cancer	*Dice (baseline)*	0.43	42.28	29.95	1018.01	0.48
	LH	0.42	23.75	20.79	948.13	0.52
	LAH	0.48	23.29	20.73	867.23	0.53
	LAH sym	**0.54**	**22.09**	**19.23**	**283.82**	**0.60**
	CE + Dice (baseline)	0.46	20.12	17.66	355.61	0.48
	CE + LH	0.48	21.41	15.99	398.72	0.54
	CE + LAH	0.49	**16.51**	14.00	523.89	**0.57**
	CE + LAH sym	**0.52**	17.00	**13.92**	**306.56**	0.56

When combined with the CE loss, the improvements over the baseline became less pronounced overall, although distance metrics and Dice score, particularly for the cancer and tumor labels, still benefited from the HD losses.

5.4 Conclusion

We introduced regional Hausdorff Distance (HD) losses to regress the Hausdorff, modified Hausdorff, and average Hausdorff Distances in CNN-based medical image segmentation. The introduced method relies on a smooth formulation of the distance transform that can be applied to probability maps. Finally, we have shown that these losses reduce the HD while preserving the Dice score without requiring any auxiliary loss. Furthermore, we show that only two iterated erosions are sufficient to compute the regional HD losses and improve the performance in terms of HD. In future work, we aim to investigate more accurate distance maps in 3D, such as Chamfer distances, and the adoption of other smooth maximum functions, rather than the LogSumExp function. Furthermore, alternative normalization terms for L_{AH} and $L_{AH_{sym}}$ could be explored to mitigate the issue of local minima that arises with an increasing number of erosion iterations. This work holds the potential to improve the segmentation of medical structures where the HD is of higher relevance, offering a new paradigm for optimizing segmentation models beyond Dice-centric metrics.

Table 5.6.: 3D nnU-Net segmentation performances on the Hepatic Vessels dataset combining the losses with the Cross Entropy (CE) Loss.

		Dice ↑	HD ↓	HD95 ↓	MHD ↓	clDice ↑
Vessels	*Dice (baseline)*	0.61	104.16	22.56	765.84	0.69
	LH	0.58	109.76	17.07	809.26	0.65
	LAH	0.61	73.46	16.04	**698.74**	0.69
	LAH sym	**0.63**	**68.80**	**15.22**	707.71	**0.71**
	CE + Dice (baseline)	**0.63**	64.76	15.67	681.14	**0.70**
	CE + LH	0.58	82.50	16.47	864.52	0.65
	CE + LAH	0.62	71.82	15.93	**660.60**	0.69
	CE + LAH sym	0.62	**59.49**	**15.39**	687.80	**0.70**
Tumor	*Dice (baseline)*	0.55	109.36	71.33	992.69	0.60
	LH	0.60	**66.74**	46.50	980.03	0.70
	LAH	**0.64**	72.48	**44.01**	751.71	**0.71**
	LAH sym	0.63	79.08	45.71	**688.56**	0.70
	CE + Dice (baseline)	0.66	70.70	42.80	499.33	0.72
	CE + LH	0.62	**61.71**	39.94	837.60	0.73
	CE + LAH	**0.67**	62.06	**33.11**	600.73	**0.76**
	CE + LAH sym	0.66	67.22	34.88	**529.09**	0.74

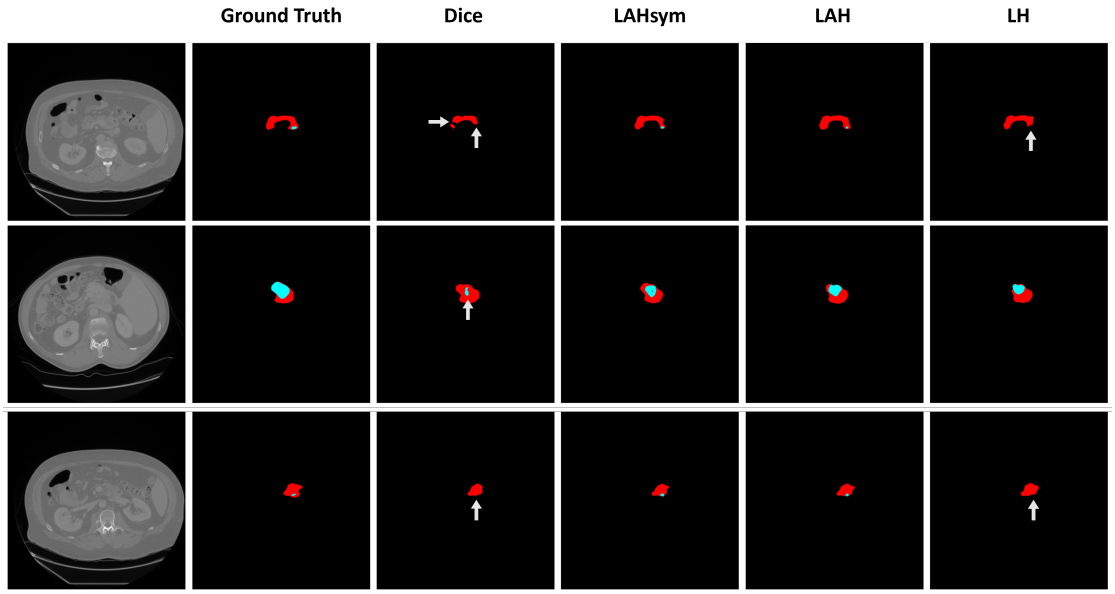


Figure 5.2.: Qualitative example based on the axial view of segmentation results on the Pancreas Dataset with the Dice and Hausdorff distance loss functions. The Pancreas label is in red, and the cancer label is in cyan. The white arrows show missing labels.

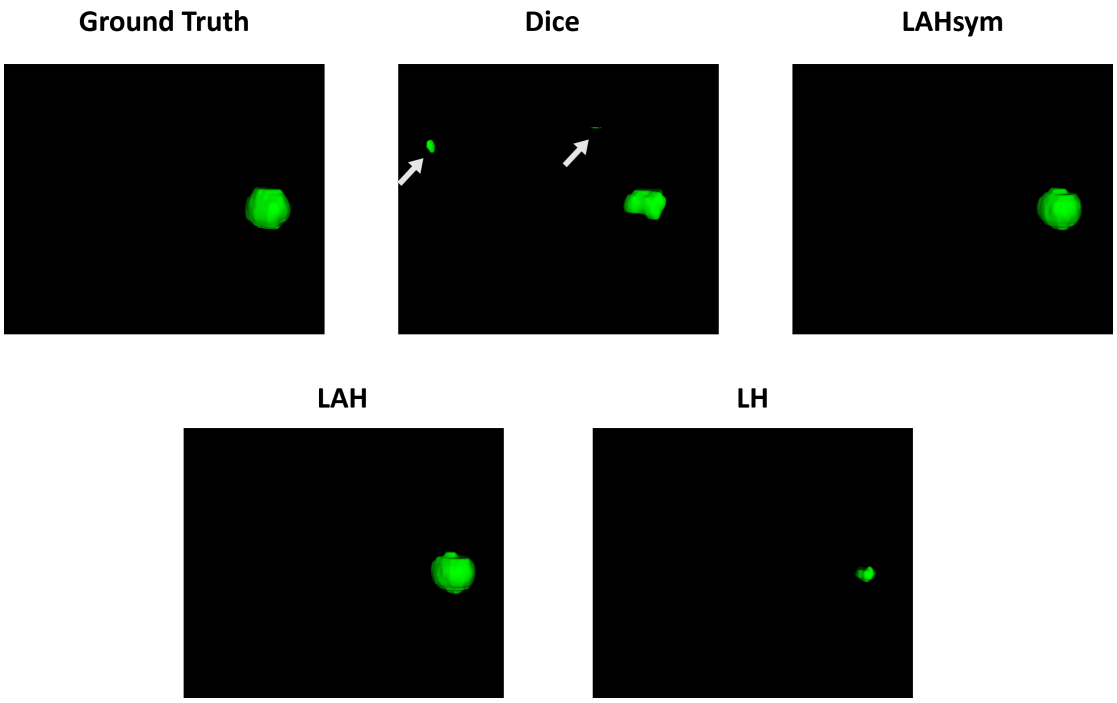


Figure 5.3.: Qualitative example based on 3d rendering of the cancer label segmentation results on the Pancreas Dataset with the Dice and Hausdorff distance loss functions. The cancer label is in green. The white arrows show mislabelings.

Part III

Clinical Application

6 PADSET: A Private CTA Dataset for Peripheral Artery Disease

Contents

While the previous chapters introduced the methodological contributions of this thesis, their full potential could only be assessed in a realistic and clinically relevant setting. For Peripheral Artery Disease, however, no open-source CTA dataset exists that offers the level of detail and annotation required for the type of analyses presented here. This motivated the creation of PADSET, a dedicated in-house dataset curated specifically for this work.

In this chapter, we present PADSET, consisting of lower-limb CTA scans acquired at the Hospital of Nice, accompanied by high-quality annotations of the arterial lumen, calcification plaques, stents, and individual arterial branches. The manual segmentation of lower-limb arteries in CTA is particularly challenging due to their small size in distal regions, the frequent presence of collateral arteries, and the large number of slices per scan, sometimes exceeding more than 2000 slices. These difficulties required a carefully designed annotation protocol and close collaboration with clinical experts to ensure both anatomical accuracy and consistency across the dataset.

6.1 Motivations

Lower-limb CTA is the clinical gold standard for assessing PAD due to its non-invasive nature, high spatial resolution, and rapid acquisition time. Despite this, no publicly or privately available dataset currently provides comprehensive coverage of the entire lower-limb arterial tree with sufficient data in either healthy individuals or PAD patients. Table 6.1 summarizes existing datasets. Several are limited in scope, focusing only on the aorto-iliac region [Imran, 2025; Radl, 2022], which excludes the femoral, popliteal, and below-the-knee arteries that are frequently affected in PAD. Other resources rely on alternative imaging modalities such as Ferumoxytol-enhanced MRA (FE-MRA) [Ghodrati, 2022]. This technique corresponds to an MRI-based vascular imaging using ferumoxytol as a contrast agent. While FE-MRA provides high-quality vascular images, it is insensitive to calcium, prone to artifacts in the presence of metallic implants, and is less commonly used in clinical practice compared to CTA. As a result, it is unsuitable for studying calcification plaques or stents. A few recent private CTA datasets do cover the lower limbs, but they remain insufficient for robust research and, most critically, they are not publicly accessible. For instance, the dataset in [Zulfiqar, 2026] contains only 36 volumes without cases of occluded arteries, while the dataset in [Anwer, 2025] includes 80 volumes, of which only about 10 have detailed annotations. Moreover, annotations are restricted to bounding boxes [Anwer, 2025], and anatomical coverage is either partial or not clearly specified. Hence, there is a lack of a relevant dataset designed explicitly for the research presented in this thesis.

To address this gap, we curated a dedicated lower-limb CTA dataset that encompasses a wide spectrum representation of PAD lesions, including calcifications and stents, with annotations of artery lumen, calcification plaques, and stents. We additionally generated annotations of each arterial branch in the lower limb arterial tree. This dataset enables the development and evaluation of segmentation methods for PAD on clinically realistic data, capturing the full complexity of lower-limb arterial anatomy and pathology.

6.2 Presentation of the dataset

A dataset of 196 lower-limb CTA anonymized scans from patients with PAD was acquired from the Hospital of Nice. The study was conducted in accordance with the World Medical Association Declaration of Helsinki and French ethical regulations, and was approved by the University Hospital of Nice's review board (register n°294). Twenty-five scans were excluded due to suboptimal image quality, including issues related to anatomical coverage, resolution, or non-visible contrast agent in the arteries. The remaining 171 scans were converted from DICOM to NIfTI format to stack 2D axial slices in a single 3D volume, resulting in 3D images with a mean image dimension of $512 \times 512 \times 1657$

Table 6.1.: Summary of available datasets for the segmentation of lower-limb arteries.

Dataset	Modality	Size	Anatomy	Limitations
AVT (Public) [Radl, 2022]	CTA	56 scans from KiTS [Heller, 2019; Heller, 2021], RIDER [Zhao, 2015], Dongyang Hospital	Aorta and major branches (arch vessels, thoracic/abdominal aorta, bilateral iliac arteries)	Stop at iliac arteries, no PAD patients
AortaSeg24 MICCAI Challenge (Public) [Imran, 2025]	CTA	100 scans from multiple institutions	23 classes covering the aorta and main branches. Includes iliac branches (common, external, internal)	Stop at iliac arteries, no PAD patients
Private (UCLA) [Ghodrati, 2022]	FE-MRA (Ferumoxytol MRA)	45 images	Segmentation of all blood vessels in calves, including arterial and venous trees	Poor signal for calcifications and stents, low number of samples, no indication on the anatomical coverage
Gennimatas datasets [Zulfiqar, 2026] (private)	CTA	36 scans from Gennimatas Hospital, Athens, Greece	Labels for the common, external and internal iliac arteries, common femoral, popliteal and anterior and posterior tibial arteries from PAD patients	Small data amount and no coverage of superficial femoral arteries and fibular artery. Absence of partially or fully occluded arteries
CTA dataset [Anwer, 2025] (Private)	CTA	67,850 slices from 80 subjects.	Aortoiliac, femoropopliteal and proximal tibial arteries from patients with suspected PAD. Bounding-box annotations on a subset of 13,429 slices.	Low amount of data (on average 848 slices per patient, with only one-fifth of slices annotated.), no fibular arteries and no delineation annotations, released recently (2025)

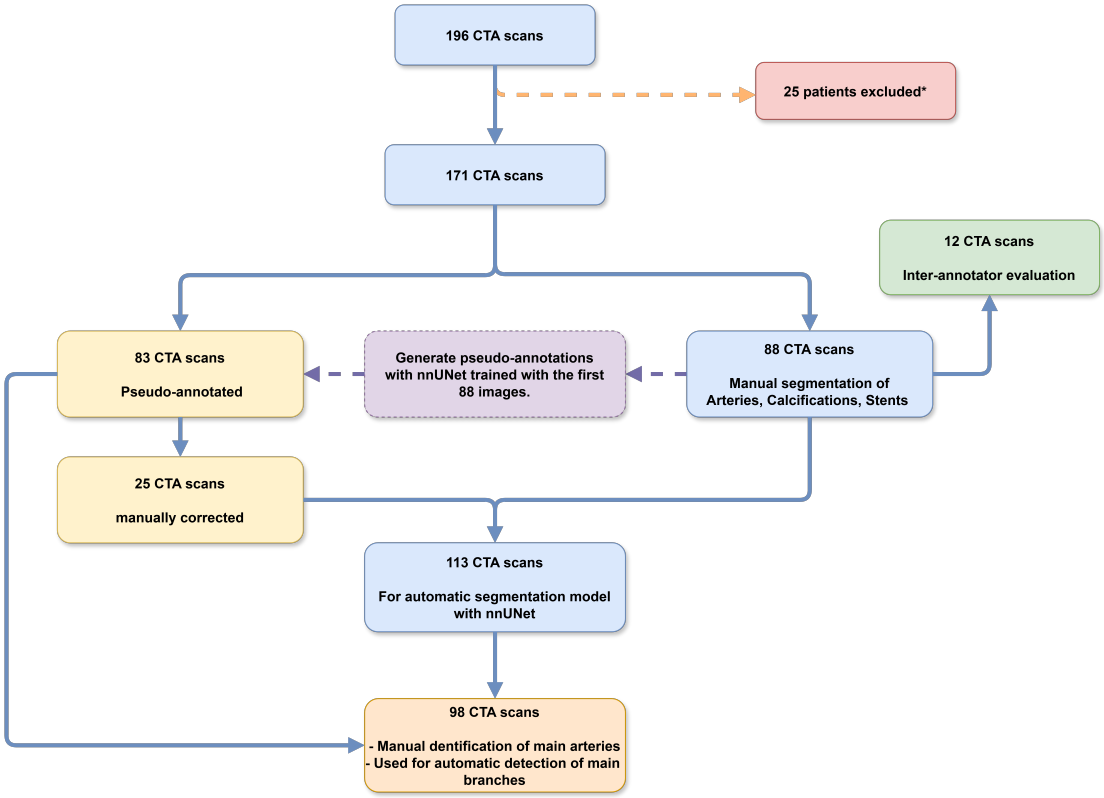


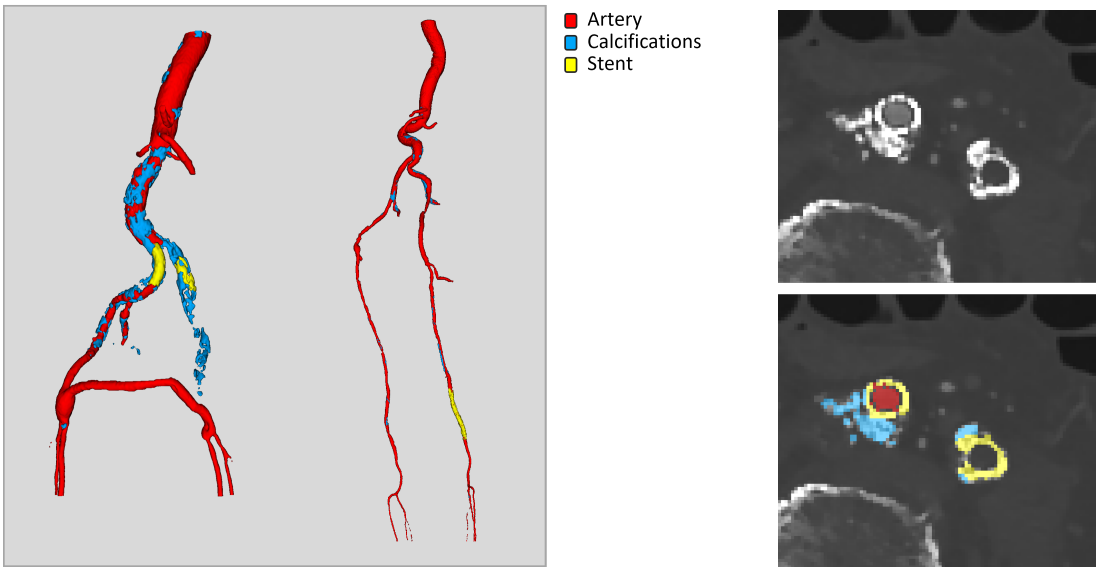
Figure 6.1.: Flowchart of dataset distribution across tasks. Blue: initial CTA scans acquired and retained for artery, calcification, and stent segmentation. **Red:** excluded patients. **Green:** scans selected for inter-annotator evaluation. **Yellow:** scans pseudo-annotated with nnUNet for artery, calcifications, and stents labels, with a subset manually corrected to increase the number of annotated images. **Orange:** subset of pseudo-annotated and manually annotated artery masks used to generate secondary annotations of individual arterial branches.

slices ($\pm$ 595) and a mean voxel spacing of 0.81 $\times$ 0.81 $\times$ 0.77 mm ($\pm$ 0.11 $\times$ 0.11 $\times$ 0.83 mm).

6.3 Annotation of Arteries, Calcifications, and Stents

6.3.1 Manual annotation protocol

From the 171 selected CTA scans, 88 were randomly selected for manual segmentation (Fig. 6.1). After some training guided by an expert vascular physician to recognize PAD-specific structures, we performed the initial delineations of arteries, calcification plaques, and stents (Fig. 6.2). To assess inter-annotator variability, a subset of 12 manually annotated scans was also independently annotated by the physician.



(a) (b)

Figure 6.2.: (a) A 3D representation of Artery, Calcification, and Stent segmentation. (b) Axial CTA representation with and without segmentation annotations.

Manual segmentation was performed using ITK-SNAP v4.0 [Yushkevich, 2016] with an adaptive brush tool based on the watershed algorithm. The arterial tree was annotated in the axial plane, beginning at the abdominal aorta and extending to the fibular and tibial arteries. Collateral arteries and secondary branches were segmented up to their first bifurcation or until they became indistinguishable.

Manually segmenting each image took on average 3 to 4 hours due to the number of slices in the dataset (on average, 1657 axial slices per scan), the number of classes and the difficulty in delivering precise annotations.

6.3.2 Inter-annotator evaluation

The inter-annotator agreement was assessed on the set of 12 independently labeled scans using the same tools by ourselves and the expert (the vascular physician) reported in the first row of Table 6.2.

The agreement for the global arterial tree achieved a Cohen's kappa agreement of $\kappa = 0.91$, indicating an almost perfect concordance [Landis, 1977]. When looking at the artery sub-regions as illustrated in Fig. 6.3, agreement was lower for distal branches in the below-knee region, reaching fair agreement ($\kappa = 0.40$), compared to the Aorto-Iliac ($\kappa = 0.94$) and Femoral-Popliteal ($\kappa = 0.88$) regions, reflecting greater variability in segmenting small-caliber vessels. This reduced agreement is attributed firstly to mislabelled annotations from our part due to inexperience, the sensitivity of the Dice

metric to minor boundary differences in small structures and in the difficulties of manually segmenting very long and thin vessels. Qualitative assessment of differences per region is shown in Fig. 6.4.

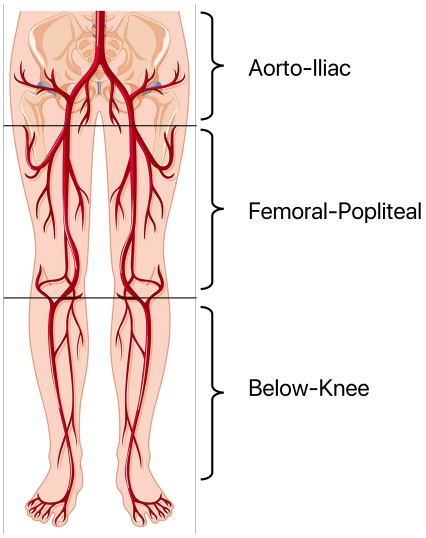


Figure 6.3.: Representation of the Aorto-Iliac, Femoral-Popliteal and Below-Knee artery regions.

For calcifications and stents, inter-annotator agreement was substantially lower, with a moderate agreement of $\kappa = 0.42$ for the calcifications and a fair agreement of $\kappa = 0.21$ for the Stents. This is mainly due to the small size of these structures and the inherent difficulty in visually identifying their exact boundaries in CTA as shown in Fig. 6.5, but also due to mislabeling between stents and calcifications due to lack of expertise on our part as shown in Fig. 6.4. It highlights the subjectivity and variability of manual annotation for such small targets and the importance of expert knowledge.

Table 6.2.: Inter-annotator evaluation of artery, calcification and stent masks.

Kappa κ	Global Artery	Aorto-Iliac	Femoral-Popliteal	Below knee	Calcifications	Stents
Initial	0.91	0.94	0.88	0.40	0.42	0.21
Corrected	0.91	0.94	0.88	0.40	0.43	0.40

6.3.3 Semi-automated labeling of calcifications and stents

Due to the difficulty of manually delineating the exact boundaries of calcifications and stents, we developed a semi-automatic tool to reduce human operator bias in detecting boundaries, and therefore improve annotation consistency.

A positive distance transform was first applied to the previously annotated artery mask to extract voxels located within a 10-voxel radius around it. Within this region, Gaussian mixture clustering was performed on voxel intensities using Scikit-learn [Pedregosa, 2011], with full covariance matrices and K-means initialization, EM optimization up to a maximum of 100 iterations with convergence tolerance of $1e^{-3}$ and default regularization

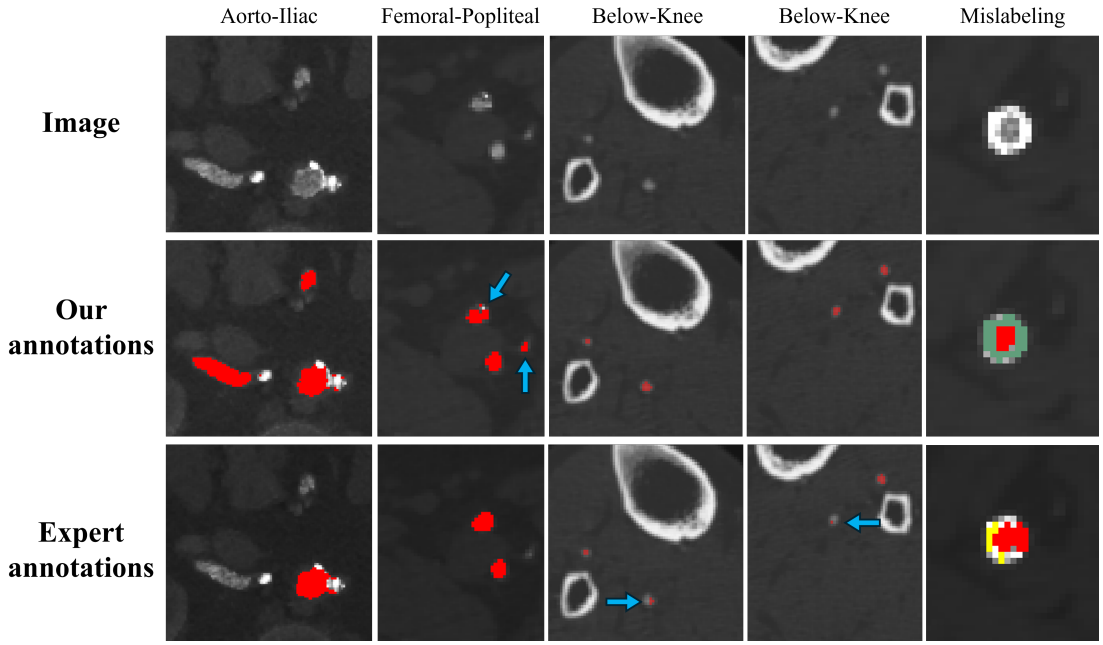


Figure 6.4.: Representation of manual annotation differences between our annotations and the expert clinician's annotations. Arteries are in **red**, Calcifications in **green**, and stents in **yellow**. Subtle differences are highlighted with **blue arrows**. From left to right: differences in the aorto-iliac; femoral-popliteal; below-knee regions; mislabeling of stent annotation.

of $1e^{-6}$. We set the number of clusters to 3, which generally corresponded to anatomical structures (including the artery lumen), the background, and high-density structures such as calcifications and stents. Respectively, the mean intensities of each cluster correspond to Hounsfield Units (HU) of around 100, -600 and 1100. Since calcifications and stents typically appear with saturated intensities in CTA (HU values above 700), the cluster with the highest mean intensity was initially assigned to the generic class "calcifications".

However, because stents and bones may overlap in intensity with calcifications, manual corrections were applied post-clustering to distinguish between these structures accurately, manually replacing the calcification components that should be labeled to the stent class, and removing components that actually correspond to bone structures.

This semi-automatic approach improves consistency and reliability in segmenting calcifications and stents compared to purely manual annotation, as shown in Fig. 6.5.

6.3.4 Final ground truth

Final ground truth segmentations of the artery, calcification, and stent masks were established through a systematic review by the vascular physician, in which our initial annotations were revised to rectify mislabeled objects and ensure clinical accuracy. For

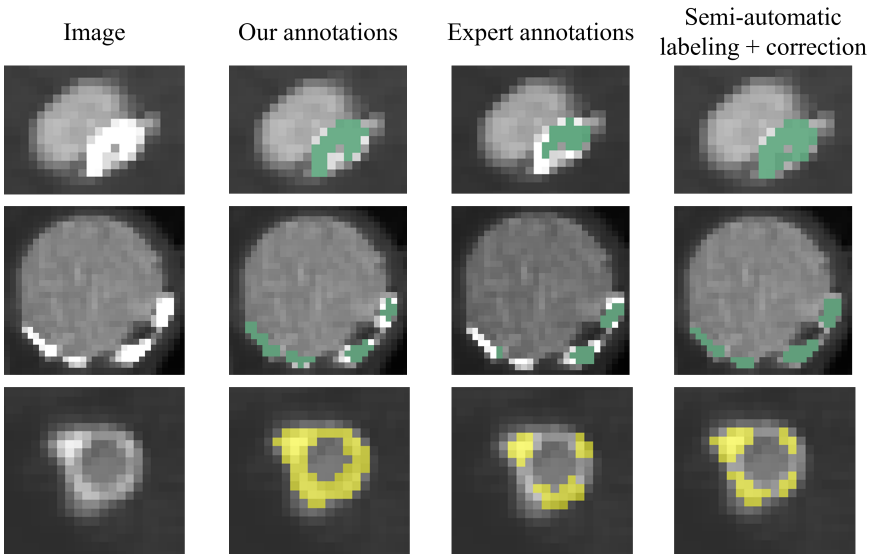


Figure 6.5.: Representation of annotation differences between our annotations, the expert clinician's annotations and the semi-automatic tool with correction for calcifications and stents. Calcifications in **green**, and stents in **yellow**.

arteries, this refinement did not alter the overall Cohen's kappa coefficient between the expert annotations and our corrected segmentations (Table 6.2). Agreement for calcifications showed only a marginal improvement ($\kappa = 0.43$ vs. 0.42), whereas agreement for stents improved more substantially, though it remained low overall ($\kappa = 0.40$ vs. 0.21). These findings underscore that even when there is consensus on the anatomical structures to be delineated, manual annotations remain prone to substantial inter-annotator variability, particularly for complex structures such as small arteries, calcifications, and stents. This challenge has been consistently reported in the literature [Falcetta, 2025; Renard, 2020; Joskowicz, 2019].

Each segmentation underwent further refinement by removing small connected components (< 5 voxels). For each segmentation label, Tukey's fences method [Tukey, 1977] was applied to identify voxel intensity outliers, defined as values lying outside the interval spanning 1.5 times the interquartile range beyond the first and third quartiles.

After the initial manual segmentations, a preliminary nnUNet [Isensee, 2021] model was trained on the 88 labeled images and applied to 25 additional unlabeled scans to alleviate the manual segmentation burden. These pseudo-labels were manually corrected to increase the sample size to 113 annotated scans (Fig. 6.1).

The correction process after the clinical verification took on average 30 minutes up to one hour per image, and the correction of the pseudo-annotated scans with the nnUNet took around one hour per image. Based on time estimations, the full annotation process required approximately 461 hours. Assuming an average of 7 hours of effective

annotation per day, this corresponds to an annotation effort of roughly 66 full working days.

6.4 Annotation of Arterial Segments

6.4.1 Landmark annotation protocol

A second segmentation protocol was implemented to label individual arterial segments on 98 scans, specifically to differentiate main artery branches, secondary branches and bypass arteries (Fig. 6.6). The clinical expert provided proximal and distal landmarks for each branch over the artery mask, including the aorta, common iliac artery, external iliac artery, superficial femoral artery, popliteal artery, tibial artery, posterior tibial artery, fibular artery, tibio-peroneal trunk and bypass grafts. All other segments were categorized as secondary branches.

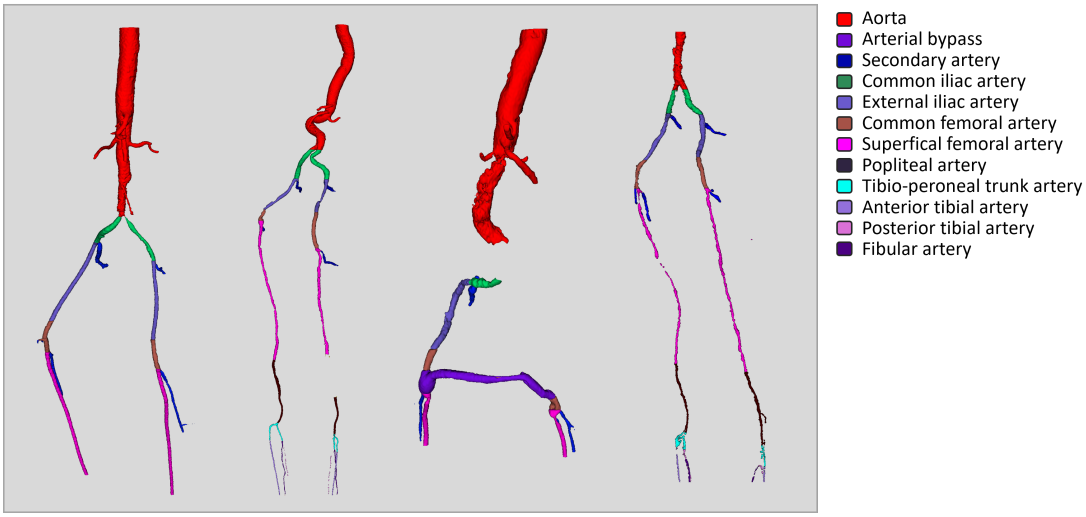


Figure 6.6.: 3D representation of main branches of the arteries of the lower limbs.

6.4.2 Semi-automated branch annotation protocol

A semi-automatic labeling pipeline connected the annotated landmarks over the binary artery mask to assign unique labels to each arterial segment.

In details, we denote the previously segmented binary artery mask $A \in \{0,1\}^{\mathbb{N}^3}$, where $A(v) = 1$ denotes artery voxels, and the expert-provided branch landmarks $b = \{b_p^n, b_d^n\}_{n=1}^N$, where each branch $n \in \{1, \ldots, N\}$ is defined by a set of proximal (b_p^n) and distal (b_d^n) landmarks. In most cases, it results in two landmarks per leg if the branch

is visible. The objective is to produce a branch mask $B \in \{0, 1, \ldots, N\}^{\mathbb{N}^3}$, such that each artery voxel labeled in A is assigned to exactly one branch identity in B.

The procedure is as follows:

1. **Landmark extraction.** For each branch n, all connected components corresponding to its proximal and distal landmarks are extracted from b. In each leg, the proximal and distal landmarks are used to define a spatial bounding box. Specifically, let

$$x^n_{\min} = \min\{x(b^n_p), x(b^n_d)\}, \quad x^n_{\max} = \max\{x(b^n_p), x(b^n_d)\},$$

   and analogously for y and z. The bounding box is then

$$\Omega^n = [x^n_{\min}, x^n_{\max}] \times [y^n_{\min}, y^n_{\max}] \times [z^n_{\min}, z^n_{\max}].$$

2. **Label propagation.** All voxels inside the artery mask A and within the bounding box Ω^n are assigned to branch n:

$$B(v) = n \quad \text{if } A(v) = 1 \text{ and } v \in \Omega^n.$$

   This ensures that only voxels belonging to the artery mask are labeled, and that each branch label corresponds to a subregion of A.

3. **Leg handling.** In cases where four landmarks are provided (two sets of proximal and distal landmarks corresponding to one set in each leg), the algorithm estimates a central dividing sagittal plane Π between paired landmarks. Voxels within Ω^n are then partitioned by Π to preserve left–right assignment.

4. **Iteration.** The process is repeated for all $n = 1, \ldots, N$. The final mask B thus partitions the binary artery mask into branch-specific subregions, with each voxel belonging uniquely to one labeled arterial segment.

5. **Secondary branches.** All artery voxels that did not receive a branch identity in B are assigned to the secondary branch label.

6.4.3 Final ground truth

Because bounding box assignments occasionally exceeded anatomical limits or overlapped with neighboring vessels, manual corrections were applied to the branch mask to verify and refine the accuracy of branch labeling. The images were subsequently reviewed and

corrected by the vascular expert. The overall correction took around 30 minutes per image, adding a total of around 49 hours, which corresponds to 7 full working days.

6.5 Final dataset characteristics

We analyzed the voxel intensity distribution of the global artery, calcifications and stents class across the dataset in Table 6.3 and compared it to the expected HU ranges reported in the literature for CTA [Wright, 2022; Fosbinder, 2011; Bolliger, 2009]. Extreme minimum and maximum values most likely reflect outlier voxels from manual segmentation errors or imaging artifacts. The interquartile range (25th-75th percentiles) closely matches the expected HU distribution for the corresponding tissue classes.

In addition, we assessed the class distribution across the annotated dataset on Table 6.4. There is a significant class imbalance for stents (present in 42.48% of images), and especially for bypass grafts (7.14% of images). Bypass arteries and stents are surgically implanted, and thus are underrepresented compared to native arteries. Moreover, bypass grafts exhibit highly variable anatomical presentations, as they are designed to revascularize occluded arterial segments that may occur at different locations along the arterial tree. This variability, combined with their limited prevalence, makes the learning of these classes especially challenging.

Table 6.3.: Statistics of voxel intensities for each main manually segmented class in PADSET across 113 samples.

	Artery	**Calcifications**	**Stents**
Expected HU	200-400	200-1000+	>1000
Mean	352.32	978.12	1052.43
std	128.98	392.60	561.52
min	-1002.00	-291.00	-857.00
max	3032.00	5234.00	3080.00
Q1 (25th)	263.00	333.00	425.00
Q2 (median)	690.00	900.00	1191.00
Q3 (75th)	942.00	1319.00	3080.00

6.6 Conclusion

In this chapter, we introduced PADSET, a curated dataset of lower-limb CTA scans with high-quality annotations of arterial lumen, individual arterial branches, calcification plaques, and stents. The manual segmentation process, supported by semi-automatic tools and validated through inter-annotator agreement, demonstrated that proximal arteries are easier to consistently delineate than distal branches. This reflects both their larger diameters and the inherent difficulty of annotating small-caliber vessels, where minor boundary inconsistencies can have a disproportionate effect on segmentation accu-

Table 6.4.: Class distribution for each manually segmented images in PADSET across 113 samples for the main classes and across 98 samples for the arterial branches.

Main labels

	Total	*Images*	*Percentage*
Artery	113	113	100%
Calcifications		113	100%
Stents		48	42.48%

Branches

	Total	*Images*	*Percentage*
Aorta	98	98	100%
Collaterals		98	100%
Common iliacs		97	98.98%
External iliacs		97	98.98%
Common femoral		97	98.98%
Superficial		97	98.98%
Popliteal		86	87.76%
TPT		82	83.67%
Anterior Tibial		82	83.67%
Posterior Tibial		81	82.65%
Fibular		82	83.67%
Bypass		7	7.14%

racy. The inclusion of semi-automatic annotation tools improved boundary consistency for calcifications and stents, highlighting their value for generating reproducible ground truth.

7

Automatic Segmentation of lower-limb Arteries on CTA for Pre-surgical Planning of Peripheral Artery Disease

Contents

This chapter demonstrates the clinical application of the methodological innovations introduced earlier on the curated PADSET dataset. We focus on the automatic segmentation of the lower-limb arterial tree with each arterial subsegment, the detection of calcification plaques, and the identification of endoprosthetic stents. Both conventional state-of-the-art methods and our novel approaches are applied and evaluated.

As conventional voxel-based metrics can underestimate performance for small objects with uncertain boundaries, we introduce object-level detection metrics to reflect the identification of calcifications and stents more faithfully.

The segmentation enables automated extraction of anatomical and pathological biomarkers, including vessel diameters, lengths, tortuosity, plaque distribution, and stent locations. Subsequent statistical analysis of these biomarkers characterizes the PADSET population and demonstrates the clinical utility of the dataset by providing actionable insights for surgical strategy selection, reproducible quantification of disease severity, and a foundation for automated stenosis and thrombosis assessment.

This chapter demonstrates that high-quality automatic segmentation can directly inform patient-specific pre-surgical planning in peripheral artery disease.

7.1 Automatic segmentation

7.1.1 Segmentation models and training setup

The CNN-based nnUNet [Isensee, 2021] has consistently set the benchmark for state-of-the-art medical image segmentation for a large set of anatomical structures since its introduction in 2018. Only a handful of methods have surpassed it, typically by narrow margins and at the cost of substantially higher computational complexity [Isensee, 2024]. Although transformer-based architectures have recently gained attention in medical image segmentation, CNNs still dominate in practice. The advantages of CNNs stem from data efficiency, inductive biases toward local patterns, and computational scalability. In contrast, transformer-based models demonstrate their strengths primarily when very large annotated datasets are available and when modeling global context is essential. Moreover, most Vision Transformer (ViT)-based approaches are limited to 2D settings due to their high memory requirements. Consequently, in this research and in medical imaging in general, the limited availability of labeled data and the need for precise boundary delineation generally favor CNNs models [Isensee, 2024; Roy, 2023].

Based on these observations, the nnU-Net was selected as the backbone for developing a robust and efficient automatic segmentation framework for PAD on the PADSET dataset. Its self-configuring nature and computational efficiency make it particularly well-suited for high-resolution CTA images, which require training on very large volumes. Moreover, to justify the effort invested in producing high-quality manual artery segmentations, we also benchmarked several recent models aiming to generalize vessel segmentation across datasets with few to no additional training. Specifically, we evaluated the foundation model VesselFM [Wittmann, 2025], Universeg [Butoi, 2023], and the pre-trained nnU-Net model TotalSegmentator [Wasserthal, 2023], comparing their performance against the nnU-Net model trained on PADSET.

The segmentation experiments are conducted on two annotation sets: the global artery set with the artery, calcifications, and stent classes, and the branch annotation set with the 12 classes of each main arterial segment. Characteristics of each annotation sets are summarized in table 7.1.

For the first annotation set of the global artery segmentation, scans were divided into 76 scans for the training, 20 scans for validation and 17 scans for the independent test set. For the second annotation set of branch segmentation, we divided the 98 annotated scans

into 63 training, 16 validation, and 19 scans for the independent test set, making sure that each label is represented in each set. The branches each correspond to a submask of the global artery mask. Here, the task is considered as a multi-class segmentation directly extracted from the CTA images.

Table 7.1.: Annotation sets description.

	Annotation set 1	**Annotation set 2**
Objective	Global artery	Main branch arteries
Number of classes	3	12
Classes description	global artery, calcifications, stents	Each main arterial segments and bypass
Annotation set size	113	98
Training set	76	63
Validation set	20	16
Test set	17	19

Training was conducted for all experiments on Nvidia A40 PCIe GPUs.

The nnUNet [Isensee, 2021] is a self-configuring deep learning framework designed for biomedical image segmentation. It automatically determines the model parameters based on the dataset's properties, including pre and post processing steps, training parameters, data augmentation techniques and the network architecture. For both annotation sets, the training of the segmentation framework was conducted in a 5-fold cross-validation setup, and then inferred and evaluated on a separate test set.

The intensity normalization was automatically applied by nnUNet; foreground voxel intensities from all training images were collected to compute the 0.5th and 99.5th percentiles. Voxel intensities were clipped to this range, followed by z-score normalization. Patch size was set to 96 $\times$ 320 $\times$ 80 voxels, voxel spacing was resampled to 0.838 $\times$ 0.62 $\times$ 0.828 mm, the training was performed with Adam optimizer and an initial learning rate of 0.01 across 1000 epochs. Data augmentation includes random rotations, scaling, elastic deformation, gamma correction, and mirroring. The network configuration was empirically set to a 3D network with full resolution of the image (3d_fullres), and compared with the 2D configuration (2D), 3D low resolution (3d_lowres), and cascaded networks (3d_cascaded) [Isensee, 2021].

For the loss function, the default nnUNet uses a combination of cross entropy (CE) and Dice loss. In addition, we investigated the use of skeleton-based topology preserving losses, including the clDice loss [Shit, 2021], cbDice loss [Shi, 2024], and the skeleton recall loss (skel_recall) [Kirchhoff, 2024]. In this additional experiment, the compound losses are defined as:

$$\mathcal{L}_{CE+Dice+clDice} = \lambda_1 \times \mathcal{L}_{CE} + \lambda_2 \times \mathcal{L}_{Dice} + \lambda_3 \times \mathcal{L}_{topological\_loss} \tag{7.1}$$

where λ is a weighting factor controlling the contribution of each term. In our experiments, the values of λ_1, λ_2, λ_3 were set respectively to 2, 1, 1 following the implemen-

tation described in [Shi, 2024] for the clDice and cbDice losses, and to 1, 1, 1 for the skeleton recall loss following the implementation in [Kirchhoff, 2024].

We compared our main automatic segmentation protocol with mainstream automatic models that could be directly applied to the raw scans with little to no training required. These models are described below.

VesselFM [Wittmann, 2025] is a foundation model designed specifically for vascular segmentation to generalize across diverse imaging modalities and anatomical regions. Built on a UNet backbone, the model is trained on multiple vessel-centric datasets, enabling robust performance across a wide range of vascular structures. Specifically, domain generalization is achieved through training on three data sources: a collection of annotated datasets from diverse tissue types and modalities, domain-randomized data, and synthetic data generated from a flow matching-based model. Its architecture supports few-shot adaptation to new domains with fine-tuning, making it well-suited for clinical scenarios involving variable imaging protocols. As VesselFM is desgined for global vessel segmentation, it was tested on our in-house dataset for the first annotation set of global artery, exclusively for the artery class with 0-shot, 1-shot, 10-shot, and 76-shot fine-tuning (corresponding to the number of samples used during fine-tuning).

UniverSeg [Butoi, 2023] is a 2D deep learning segmentation framework designed to generalize across a wide range of medical imaging modalities and anatomical structures. It achieves this by training on a large number of medical images, and incorporating cross-attention layers called CrossBlocks. These CrossBlocks align features between a small set of labeled support images and the unlabeled query image during inference. This mechanism enables the model to perform few-shot segmentation tasks to specialize in new anatomies or modalities with only a handful of examples, without requiring further gradient-based training. We also tested UniverSeg only for the global artery segmentation set, exclusively on the artery class, providing either 10, 50, 100, or 500 randomly selected support images.

TotalSegmentator. [Wasserthal, 2023] is a general-purpose anatomical segmentation model trained using the nnUNet framework on a large dataset of contrast-enhanced CT angiography (CTA) images covering multiple organs. With a few classes available for vascular segmentation, the model was evaluated specifically on the aorta and iliac artery classes of the branch annotation set. The model does not provide other labels of interest for this work.

7.1.2 Branch post-processing

Following the initial segmentation obtained with the nnUNet framework, a dedicated post-processing pipeline was applied to refine the identification of each arterial branch.

First, small disconnected components with fewer than five voxels were discarded to eliminate noise. Second, anatomical consistency was enforced by correcting false label transitions along the arterial tree. The algorithm analyzes each axial slice sequentially, from the aorta down to the distal vessels, while ignoring secondary branches. The expected order of branch labels was compared against the observed order of connected components along the superior–inferior axis in each leg. Specifically, the expected order corresponds to the anatomical sequence of arteries from the upper body to the lower extremities, with the fibular and tibial arteries treated as a single anatomical level for simplicity, and excluding bypass arteries (refer to Chapter 2 for anatomical details). Whenever an inconsistency was detected (e.g., a popliteal artery component appearing above a femoral artery), the algorithm compared the sizes of the conflicting components. If the superior component was more than three times larger than the inferior one, the inferior component was considered erroneous and relabeled according to the superior branch. Otherwise, the superior one was relabeled. This rule-based correction ensures that the sequence of arterial branches follows the expected anatomical hierarchy while minimizing the risk of systematic mislabeling.

7.1.3 Evaluation metrics

To assess inter-annotator manual annotations and evaluate automatic segmentation methods, we used a range of overlap, topological, and detection-based metrics that are specifically relevant to this dataset.

Overlap-based metrics. We evaluated the predictions using the Dice score, the Precision, Recall, the Hausdorff Distance (HD) and its (HD95).

Topological-based metrics. These metrics assess the preservation of anatomical continuity and connectivity, crucial in vascular segmentation. The **clDice** [Shit, 2021] metric extends the traditional Dice metric by emphasizing the alignment of tubular structures along their centerlines. To evaluate topological consistency, we computed the mean absolute difference of **Betti numbers** where β_0, β_1, and β_2 respectively represent the number of connected components, loops, and cavities. Additionally, the **Euler characteristic** χ is defined as: $\chi = \beta_0 - \beta_1 + \beta_2$. We used a connectivity of 26 in the foreground and of 6 in the background.

Detection metrics. Segmenting calcifications and stents presents a unique challenge due to their small size and the inherent subjectivity in delineating their exact boundaries. As a result, traditional voxel-wise metrics such as the Dice score can be overly sensitive to minor discrepancies. In clinical applications, the primary objective is to reliably detect the presence and approximate extent of each calcification or stent, rather than achieving perfect boundary accuracy.

To assess detection performance, we employed an object-level matching detection algorithm that compares individual segmented components between the ground truth and predicted segmentation masks. All connected components are extracted from both the ground truth and predicted segmentations. An object from the predicted segmentation is matched to a ground truth object if they share at least one voxel. In cases where multiple objects from one segmentation overlap with the same object in the other segmentation, those objects are merged into a single group to avoid redundant matches. For each matched object pair or group, the volumetric overlap is computed. An object is considered successfully detected if it overlaps by at least 30% (set empirically) with its corresponding ground truth object.

This approach yields object-level counts of true positives (TP), false positives (FP), and false negatives (FN). Using these counts, we compute detection precision and recall, using the same formulas as voxel-based metrics but applied to objects rather than voxels. To distinguish them, we denote these metrics as **D_Pre** and **D_Rec**. The **F1 Score** is the harmonic mean of precision and recall and gives a balanced measure of segmentation accuracy.

7.1.4 Results

Automatic segmentation of the vascular system

Global artery mask: We compared artery segmentation performance of the first annotation set of the global artery on the artery class across different nnUNet configurations, VesselFM, and UniverSeg in Table 7.2 and illustrated in Fig. 7.1.

The 3D cascaded nnUNet achieved the best overall performance, with a Dice score of 0.93 and a clDice of 0.87. The performance was nearly identical to that of the 3D full-resolution nnUNet, which achieved a Dice of 0.93 and a clDice of 0.86. For the 3D network with low resolution, Precision reached 0.93, whereas Recall was lower at 0.73, suggesting a tendency toward under-segmentation. Notably, the 2D nnUNet outperformed the 3D low-resolution variant despite the loss of spatial context, underlining the importance of preserving fine spatial detail, as when resolution is reduced, thin vessels are frequently

overlooked, leading to a drop in recall and topology preservation. Overall, the 3D cascaded and full-resolution networks achieved comparable accuracy, with differences falling within the margin of variability. However, the cascaded approach requires a two-stage pipeline, consisting of an initial 3D low-resolution network followed by a second high-resolution refinement network, which substantially increases training and inference complexity. By contrast, the full-resolution configuration is simpler to deploy and more computationally efficient while retaining equivalent performance. For these reasons, the 3D full-resolution nnUNet was selected for all subsequent experiments.

For UniverSeg, the best results were obtained when using 500 example images, with an average Dice of 0.29 and an HD of 324.90. However, the gain from 100 to 500 examples was marginal, indicating the model's limited capacity to adapt further to our data distribution. The poor segmentation can likely be attributed to its inherently 2D design, as 2D support examples fail to capture the longitudinal continuity of the arterial tree across hundreds or thousands of slices, and the lack of volumetric context leads to discontinuities and missing vessels. This illustrates a limitation of slice-based few-shot approaches for highly structured 3D anatomy such as vascular trees.

VesselFM demonstrated weak performance in the zero-shot setting, characterized by very low Precision (0.03) but high Recall (0.90), indicating severe over-segmentation. Qualitatively, while the model roughly localized the main arterial tree, it frequently mislabeled surrounding structures, such as bones, soft tissues, or background, as arteries. Performance improved progressively with the number of fine-tuning samples (shots), reaching its best values with 76 shots, corresponding to the entire training set. This trend highlights that pretrained vision-language models like VesselFM may capture some vascular priors, but without sufficient task-specific supervision, they introduce substantial false positives. Moreover, achieving clinically acceptable accuracy still requires a large volume of carefully annotated data. This underscores the importance of PADSET and the manual annotation efforts.

We further investigated the effect of different loss functions on the 3D full resolution nnUNet artery segmentation (Table 7.3). Incorporating clDice alongside Dice and CE in the loss improved all metrics compared to the baseline, improving boundary delineation and topological continuity. The final model reached a Dice of 0.94 and a clDice of 0.88, confirming that explicitly enforcing connectivity constraints is beneficial for vascular segmentation, where preserving topology is essential for downstream analysis. However, the integration of the skeleton recall loss and the cbDice loss did not improve the Dice and clDice metrics compared to the baseline. The best HD results were obtained with the cbDice loss (83.87).

Table 7.2.: Automatic segmentation of arteries of the lower limbs using: nnUNet framework with different network configurations, VesselFM with different number of shots, and Universeg with different number of support samples.

Model	Configuration	**Dice ↑**	**clDice ↑**	**HD ↓**	**HD95 ↓**	**Prec. ↑**	**Rec. ↑**	$\beta_0 \downarrow$	$\beta_1 \downarrow$	$\beta_2 \downarrow$	$\chi \downarrow$
nnUNet	*3dfullres*	**0.93**	0.86	103.36	5.59	0.93	**0.93**	**35.06**	67.00	**25.29**	**69.00**
	3dlowres	0.81	0.65	83.18	11.97	**0.94**	0.73	36.82	70.47	27.18	70.47
	3d cascaded	**0.93**	**0.87**	95.85	**4.80**	**0.94**	**0.93**	36.74	**66.53**	25.65	69.18
	2d	0.89	0.79	**93.55**	11.95	0.91	0.88	38.94	66.82	28.47	73.53
VesselFM	*0-shot learning*	0.06	0.04	955.81	953.04	0.03	0.90	5536.71	1896.88	603.71	4243.53
	1 shots	0.24	0.17	769.84	735.89	0.15	0.69	1579.82	471.47	278.94	1413.65
	10 shots	0.27	0.24	413.35	293.75	0.22	0.41	855.24	114.53	32.18	784.53
	76 shots	0.40	0.27	341.36	163.34	0.29	0.77	438.82	164.18	78.29	378.35
UniverSeg	*N=10 samples*	0.11	0.08	259.59	179.95	0.23	0.08	126.00	66.53	27.29	163.47
	N=50	0.25	0.15	348.54	254.65	0.48	0.19	80.88	66.18	27.29	118.71
	N=100	0.29	0.14	331.80	240.76	0.59	0.21	61.47	68.59	27.29	100.76
	N=500	0.29	0.17	324.90	229.71	0.64	0.20	70.73	85.64	24.91	126.36

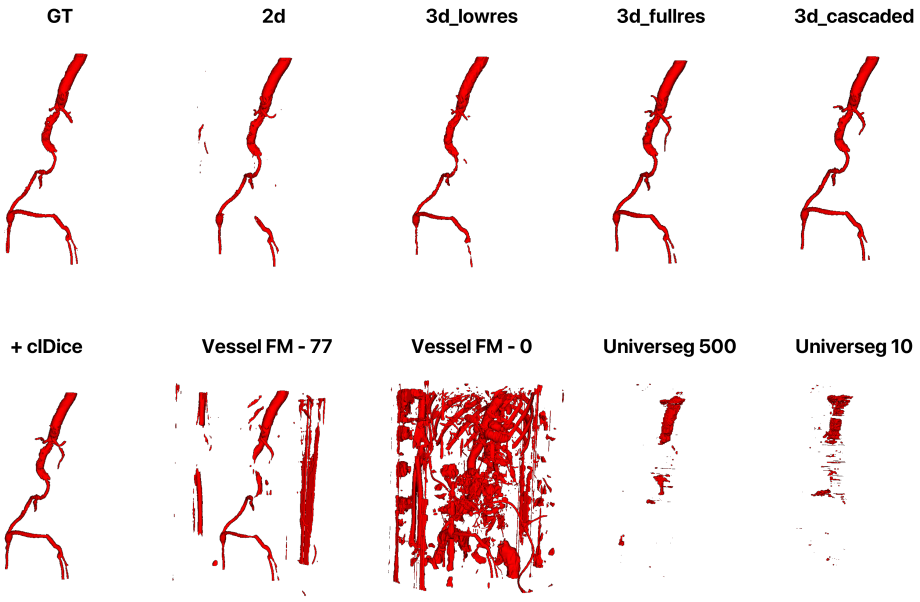


Figure 7.1.: 3D volume representation of artery segmentation of ground truth label (GT), nnUNet with the different network configurations (2d,3d_lowres,3d_fullres,3d_cascaded,3d_fullres with additional clDice loss function), VesselFM (with 76 and 0 shot) and Universeg (with 500 or 10 support images).

Table 7.3.: Automatic segmentation of arteries of the lower limbs using the nnUNet framework with different loss functions: Cross Entropy (CE) + soft Dice baseline compound loss, with the addition of either the skeleton recall loss (skel_recall), the cbDice loss, or the clDice loss.

	Dice ↑	clDice ↑	HD ↓	HD95 ↓	Prec. ↑	Rec. ↑	β_0 ↓	β_1 ↓	β_2 ↓	χ ↓
CE+Dice	0.93	0.86	103.36	5.59	**0.93**	0.93	35.06	67.00	25.29	69.00
+skel_recall	0.90	0.78	142.56	25.56	0.86	**0.95**	104.00	59.18	23.53	137.88
+cbDice	0.93	0.85	**83.87**	6.06	0.92	0.93	**24.41**	66.24	23.88	58.18
+ clDice	**0.94**	**0.88**	93.42	**3.43**	**0.93**	**0.95**	24.65	**59.53**	**22.12**	**54.29**

Main arterial segments: We compared the segmentation performance on the second annotation set of main branch arteries with the nnUNet in a 3D setting with full resolution, and with TotalSegmentator.

We evaluated the performance of TotalSegmentator on the aorta and iliac branches (Table 7.4). The model achieved mean Dice scores of 0.81 and 0.55, and clDice scores of 0.75 and 0.74, for the aorta and iliac arteries, respectively. The discrepancy between Dice and clDice in the iliac arteries suggests that most of the segmentation errors are attributable to the unspecificity of the common and external/internal iliacs or boundary inaccuracies rather than complete detection failures. Nonetheless, a fundamental limitation of TotalSegmentator is its inability to segment peripheral lower-limb arteries, as these structures are not included in its training label set.

For comparison, we report the averaged segmentation performance of the 3D full-resolution nnUNet, both before and after post-processing in Table 7.5, with detailed results per branch provided in Table 7.6. The nnUNet achieved a Dice score of 0.96 for the aorta and 0.79 for the iliac arteries, underscoring the benefit of domain-specific training for these vascular regions. Performance declined in more distal arteries, with the popliteal, tibioperoneal trunk, and posterior tibial arteries all yielding Dice scores below 0.50. This drop aligns with the inter-annotator variability analysis, where distal arteries also showed reduced consistency due to their small caliber, ambiguous boundaries, and frequent pathological alterations. The lowest performance was observed for bypass grafts, with a Dice of 0.19 and a clDice of 0.34. This outcome is consistent with the limited representation of bypasses in the dataset (present in only seven scans) and their high variability in surgical placement and morphology, which hampers model generalization.

In summary, branch-wise segmentation accuracy was highest for large, proximal arteries (aorta, iliac, femoral) and progressively decreased for smaller distal branches, particularly below the knee. These trends closely mirror inter-annotator agreement, suggesting that reduced boundary clarity and greater anatomical variability in distal vessels represent intrinsic challenges for both human annotators and automated models.

Before post-processing, a significant error often occurred with the popliteal artery label present in the superficial femoral artery, as shown in Fig. 7.2. The Post-processing helped

Table 7.4.: Automatic segmentation of arteries of the lower limbs using TotalSegmentator for the segmentation of the Aorta and the Iliacs (common + external) arteries.

Artery	**Dice** ↑	**clDice** ↑	**HD** ↓	**HD95** ↓	**Prec.** ↑	**Rec.** ↑	β_0 ↓	β_1 ↓	β_2 ↓	χ ↓
Aorta	0.81	0.75	418.37	31.15	0.74	0.92	4.50	9.10	7.45	5.65
Iliacs	0.55	0.74	64.81	35.89	0.40	0.90	2.45	2.85	1.40	3.50

correct these discrepancies by enforcing the anatomical order of labels. These corrections were more comprehensive to assess qualitatively rather than quantitatively, as the final results in the appendix Table B.1) show minor improvements compared to before post-processing results. The main improvements were observed in HD and topological metrics (Table 7.5).

Table 7.5.: Averaged automatic segmentation of the main branches of the lower limbs with the nnUNet framework before and after applying post-processing.

	Dice	**clDice**	**HD**	**HD95**	**Precision**	**Recall**	β_0	β_1	β_2	χ
Before	0.62	0.70	188.22	78.65	0.79	0.57	6.05	1.49	0.90	6.30
After	0.62	0.70	**183.86**	78.01	**0.80**	0.57	**5.51**	**1.34**	0.90	**5.80**

Table 7.6.: nnU-Net segmentation performance of each arterial branch.

	Dice ↑	**clDice** ↑	**HD** ↓	**HD95** ↓	**Prec.** ↑	**Rec.** ↑	β_0 ↓	β_1 ↓	β_2 ↓	χ ↓
Aorta	0.96	0.90	402.78	2.85	0.97	0.94	3.85	8.65	6.89	6.20
Second.	0.68	0.61	245.90	52.00	0.75	0.65	5.58	0.85	0.30	5.55
C. iliac	0.79	0.84	28.93	13.05	0.81	0.81	0.60	1.25	1.05	1.10
E. iliac	0.79	0.87	39.25	11.43	0.83	0.79	1.50	1.75	0.30	2.05
C. fem.	0.67	0.72	78.29	63.69	0.71	0.69	1.10	0.45	0.05	1.50
S. fem.	0.75	0.78	156.49	48.16	0.78	0.76	7.80	3.25	1.50	7.15
Popliteal	0.49	0.62	169.40	111.10	0.74	0.40	2.27	0.47	0.06	2.06
TPT	0.45	0.60	35.45	29.02	0.79	0.33	1.53	0.12	0.06	1.47
A. Tib.	0.70	0.80	95.92	21.24	0.88	0.60	3.47	0.36	0.00	3.35
P. Tib.	0.32	0.55	73.15	33.92	0.97	0.20	32.70	0.18	0.06	32.83
Fibular	0.63	0.75	70.45	50.33	0.81	0.54	6.23	0.06	0.00	6.30
Bypass	0.19	0.34	862.67	507.00	0.50	0.12	6.00	0.50	0.50	6.00

Automatic segmentation of anomalies

We reported segmentation performance of the 3d full-resolution nnUNet for calcifications and stents in Table 7.7. The incorporation of the clDice loss function improved performance across all detection, overlap, and topological metrics across both structures. Calcifications were segmented with a Dice score of 0.82 and an F1-score of 0.83, while stents achieved a Dice score of 0.69 and an F1-score of 0.69. Despite the challenges of manual annotation, both structures were segmented with acceptable accuracy. This performance can be attributed in part to the semi-automatic annotation workflow, which reduced inter-annotator variability at structure boundaries and ensured more consistent ground-truth labels. Stent segmentation likely exhibited lower performance due to their limited representation in the dataset, being present in only 42% of the annotated images. The results highlight that even small, high-intensity structures with complex morphology can be reliably captured when supported by consistent annotation and topology-aware loss functions.

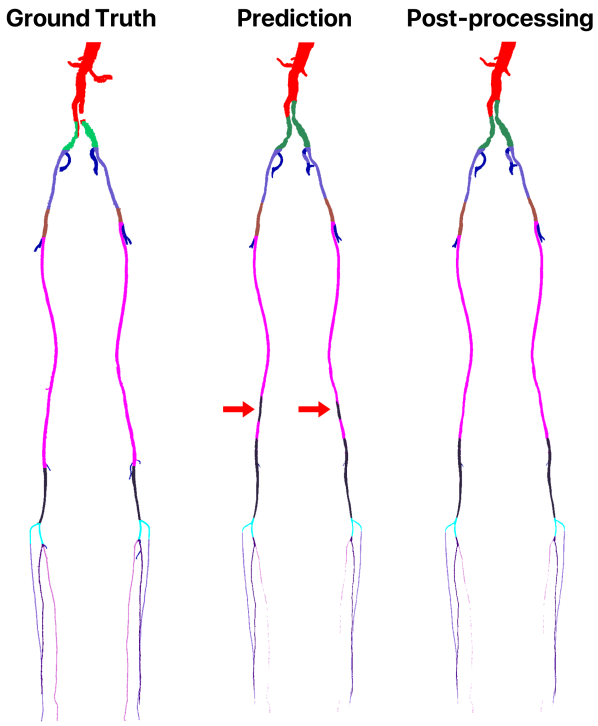


Figure 7.2.: 2D projection representation of segmentation of each arterial branch before and after processing. Post-processing corrections observed inside the dark blue rectangle region, errors in the prediction before post-processing highlighted by red arrow.

Table 7.7.: nnU-Net segmentation performance of the calcification and stent masks with and without incorporating the clDice loss.

		D_Pre	D_Rec	F1	β_0	β_1	β_2	χ	Dice	HD	HD95
Calcif.	-	0.78	0.89	0.82	16.65	27.76	4.06	27.53	0.80	180.04	12.70
	+*clDice*	**0.79**	0.88	**0.83**	**13.06**	**27.34**	4.18	28.59	**0.82**	**175.39**	**6.58**
Stents	-	0.65	0.82	0.65	7.60	85.50	5.50	82.00	0.68	137.14	74.90
	+*clDice*	**0.66**	**0.85**	**0.69**	**6.00**	90.50	5.70	85.60	**0.69**	162.73	109.03

7.1.5 Application of methodological contributions

Due to the high resolution and large size of the PADSET images (average volume size of $512 \times 512 \times 1645$ voxels, with voxel spacing of $0.81 \times 0.81 \times 0.77$ mm), training a 3D full-resolution nnU-Net across 1000 epochs with 5-fold cross-validation is extremely time-consuming, averaging more than 15 hours per fold. As the proposed methodological contributions introduce additional complexity to the pipeline, running this full protocol for every methodological configuration would therefore be computationally prohibitive, requiring several months of GPU time.

To enable a systematic evaluation of the methodological contributions under realistic time constraints, we adopted a reduced training regime of 250 epochs. While this shorter training schedule may not fully saturate performance, it provides a consistent and computationally tractable framework for controlled comparisons between methods. Importantly, all baselines and proposed variants were trained under identical conditions, ensuring fairness in the evaluation.

The full 1000-epoch setting was already established in the previous experiments to demonstrate the upper-bound performance of the baseline nnU-Net. Here, the focus is instead on relative gains introduced by the methodological contributions, which remain meaningful even under shorter training.

SoftMorph

We evaluated the SoftMorph morphological filters introduced in Chapter 4, as final layers of a 3D full-resolution nnU-Net applied to PADSET for calcification plaque segmentation. The results are summarized in Table 7.8.

We report only D_pre, D_rec, and F1, as these directly reflect detection and classification performance for calcifications. Other overlap-based metrics (such as Dice) did not provide additional insight in this context. Furthermore, morphological filtering did not yield improvements for arteries or stents, so we focus here exclusively on calcification detection, where the impact was the most meaningful. Here, the model was trained exclusively to predict calcification with the usual multi-class segmentation of arteries and stents.

Without morphological filtering, the network achieved a detection precision (D_pre) of 0.72 and an F1 score of 0.75. Incorporating morphological operations as final layers improved the results. The closing and opening filters increased both precision and F1, with the opening filter reaching 0.79 precision and 0.79 F1. The erosion filter yielded the

best trade-off, achieving the highest precision (0.91) and F1 (0.81), although at the cost of a reduced recall (0.84 compared to 0.96 without filtering).

This trade-off highlights the effect of morphological filtering as the erosion filter eliminated small false-positive predictions (boosting precision), it can also remove subtle true positives, thereby lowering recall. From a clinical standpoint, the higher precision is particularly valuable, as reducing false positives can increase the reliability of automated detection of calcified plaques.

Table 7.8.: Detection results of PADSET with different SoftMorph last layers. (*) denote statistically significant improvement compared to the baseline.

Final Layer	D_pre ↑	D_rec ↑	F1 ↑
Baseline	0.72	**0.96**	0.75
Closing	0.75	0.92	0.77
Opening	0.79	0.92	0.79
Erosion	**0.91***	0.84	**0.81**

Hausdorff Distance Losses

We trained the 3D full-resolution nnU-Net using two baselines (Dice loss and CE+Dice loss) and compared them with the proposed Hausdorff losses. Results are summarized in Table 7.9. Training was limited to 250 epochs, which proved sufficient to optimize artery and calcification segmentation, but was insufficient to fully converge on the third label (stents). Consequently, stent segmentation performance remained suboptimal across most configurations, with Dice values below 0.30 except for the CE+Dice and the CE+LAH_{sym} losses.

For arteries, all Hausdorff-based losses consistently improved distance-based metrics compared to their corresponding baselines, with the exception of CE+LAH, which degraded performance. The best results were obtained with CE+LAH_{sym}, achieving a Dice score of 0.92 (vs. 0.89 for the baseline), and a substantial reduction in Hausdorff distance (HD: 118.7 vs. 337.9; HD95: 8.0 vs. 214.5). In addition, clDice improved from 0.64 to 0.83, indicating gains not only in overlap but also in topological consistency. Qualitatively, the example shown in Fig. 7.3 demonstrates that while the other losses encouraged mainly the delineation of larger vessels, the LH and LAH_{sym} losses enabled to capture the whole arterial tree down to the smaller vessels, explaining the significant increase in clDice.

Results for calcifications followed a similar trend, although improvements were less consistent. The LH loss alone drastically reduced Dice to 0.40 and degraded all metrics. LAH and LAH_{sym} provided moderate improvements over the Dice baseline, with LAH also improving the detection metrics (F1: 74.55 vs. 67.73). Only CE+LAH_{sym} achieved

performance comparable to CE+Dice, with slight improvements in HD95 (30.1 vs. 35.7), clDice (0.79 vs. 0.78), and detection metrics (F1: 81.31 vs. 79.07). Overall, while distance losses did not substantially boost calcification overlap metrics, they contributed to small improvements in boundary accuracy and detection metrics.

As noted, 250 epochs were insufficient to optimize the stent class, due to its small representation and complex appearance. Nevertheless, CE+LAH_{sym} outperformed the baseline across all metrics (Dice: 0.64 vs. 0.62; HD95: 47.6 vs. 235.1; clDice: 0.67 vs. 0.62, F1: 55.85 vs. 38.00), suggesting that the loss can still extract more reliable spatial information even under limited training.

Across structures, CE+LAH_{sym} provided the most consistent improvements, especially for the artery mask, where it achieved significant gains in both overlap and distance-based metrics. These results indicate that the averaged symmetric Hausdorff loss is the most effective configuration for enhancing boundary-sensitive vascular segmentation in PADSET, though additional training epochs would likely be required to fully assess its impact on stent segmentation. Moreover, all Hausdorff distances were set with 2 erosion iterations. Further studies on the optimal number of iterations tailored to this dataset should be conducted.

Table 7.9.: 3D nnU-Net segmentation performances on PADSET comparing the baselines Dice and CE+Dice losses with the proposed HD losses for the artery, calcifications and stents masks. (*) denote statistically significant improvements compared to the baseline using the Wilcoxon rank test.

		Dice ↑	HD ↓	HD95 ↓	clDice ↑	D_pre ↑	D_rec ↑	F1 ↑
Artery	*Dice (baseline)*	0.86	442.86	330.09	0.61	-	-	-
	LH	0.85	377.87	237.56	0.58	-	-	-
	LAH	0.86	369.41	251.33	0.61	-	-	-
	LAH sym	**0.87**	**300.16**	**189.62***	**0.63**	-	-	-
	CE + Dice (baseline)	0.89	337.86	214.45	0.64	-	-	-
	CE + LH	0.91	129.64*	15.01*	0.78*	-	-	-
	CE + LAH	0.87	340.26	223.77	0.63	-	-	-
	CE + LAH sym	**0.92***	**118.71***	**8.02***	**0.83***	-	-	-
Calcif.	*Dice (baseline)*	0.65	273.03	134.52	0.68	64.93	62.55	67.73
	LH	0.40	442.09	193.90	0.43	30.86	83.89	43.40
	LAH	0.69	**181.12**	78.16	0.71	**78.76***	**73.54**	**74.55**
	LAH sym	**0.75***	193.57	**39.97***	0.76	57.02	72.11	62.02
	CE + Dice (baseline)	**0.78**	**170.51**	35.71	0.78	77.50	82.83	79.07
	CE + LH	0.61	191.26	56.22	0.60	57.03	72.11	62.02
	CE + LAH	0.72	171.54	49.15	0.76	70.29	84.19	75.73
	CE + LAH sym	**0.78**	215.57	**30.06**	**0.79**	**80.31**	**83.68**	**81.31**
Stents	*CE + Dice (baseline)*	0.62	459.81	235.12	0.62	**63.85**	33.61	38.00
	CE + LAH sym	**0.64**	**214.56**	**47.56***	**0.67**	**63.86**	**68.47***	**55.85**

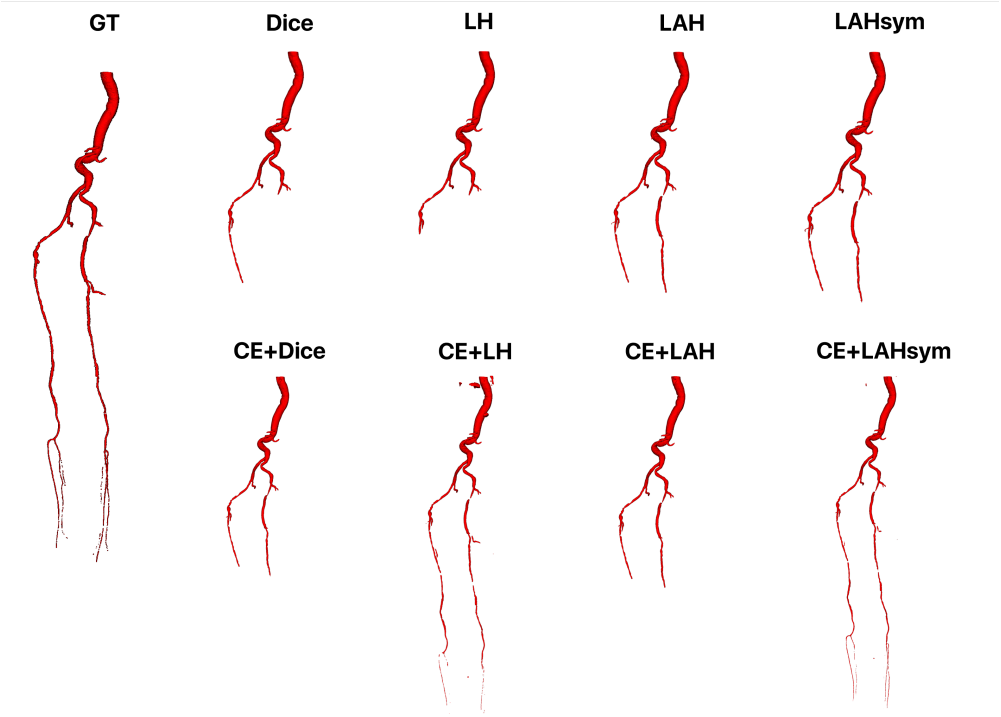


Figure 7.3.: 3D volume representation of artery segmentation of ground truth label (GT), and nnUNet full resolution with the different loss functions.

7.2 Estimation of pathological features.

The final segmentation volumes provide quantitative information on both anatomical and pathological characteristics of each patient. These measurements are directly relevant for presurgical planning in peripheral artery disease (PAD) and can be mapped to standardized scoring systems such as GLASS or TASC II.

7.2.1 Anatomical features

Anatomical descriptors capture patient-specific vascular geometry and are essential for treatment planning, including stent placement and surgical access routes. They also allow objective assessment of arterial narrowing. We report example metrics such as artery length, diameter, and tortuosity for each vascular branch in Table 7.10, all extracted using VesselVio [Bumgarner, 2022].

Specifically, the **length** computed for each segment is defined as the sequence of connected centerline voxels between two branch points or endpoints. Length is calculated as the sum of Euclidean distances between consecutive centerline voxels.

Table 7.10.: Average artery diameter (mm) , length (mm) and tortuosity per branch across PADSET.

	Aorta	C. i.	E. i.	C. f.	S. f.	Popliteal	TPT	A. t.	P.t.	Fibular
Diameter	8.73	4.58	4.04	4.87	3.23	3.26	2.31	1.76	1.67	1.61
Length	14.71	16.98	36.38	20.96	61.53	55.45	33.40	82.53	101.96	91.29
Tortuosity	1.26	1.23	1.17	1.10	1.09	1.03	1.67	1.04	1.05	1.04

Vessel radii are estimated from the Euclidean distance between each centerline voxel and the surrounding background. A cubic search region is expanded until at least four background voxels are identified, and the mean of the four smallest distances is taken to reduce noise and anisotropy bias. Because a standard Euclidean Distance Transform systematically overestimates the true vessel radius (as vessel walls lie on voxel borders rather than centers), a corrected distance lookup table (LUT) is applied. In this LUT, distances to face-connected neighbors are reduced by half the voxel resolution, while edge- and corner-connected distances remain unchanged. This correction yields radius estimates that more accurately represent the true lumen boundary. The vessel diameter is obtained by doubling the corrected radius.

Tortuosity is quantified as the ratio of centerline length to the Euclidean distance between segment endpoints. A perfectly straight vessel has a tortuosity of 1, with higher values indicating increasing curvature.

We can also extract the number and location of **secondary branch**, specifically by extracting the number of connected components for this label, along with their exact slice position, and matching them to the main branch locations. For example, (Fig. 7.4) exhibits seven secondary branches: one in each iliac artery (corresponding to the internal iliac), two collaterals arising from the right common femoral artery, and three branches in the superficial femoral arteries (corresponding to the deep femoral).

7.2.2 Clinical features

Clinical descriptors capture disease burden and previous interventions, both of which are essential for presurgical planning.

We computed the average calcification burden for each vascular branch by calculating the ratio of calcification volume to arterial volume within each segment in Table 7.11. Fig. 7.5.a. shows the distribution of calcification volumes for *patient_001*. Each bar represents the number of segmented calcifications falling within a given volume range. For instance, the majority of calcifications have small volumes (<50 mm^3), whereas a few larger lesions are also present (of 300 and >400 mm^3). This demonstrates that the segmentation allows us to quantify lesion burden not only in terms of presence but also distribution and size. Fig. 7.5.b. shows the distribution of total calcification

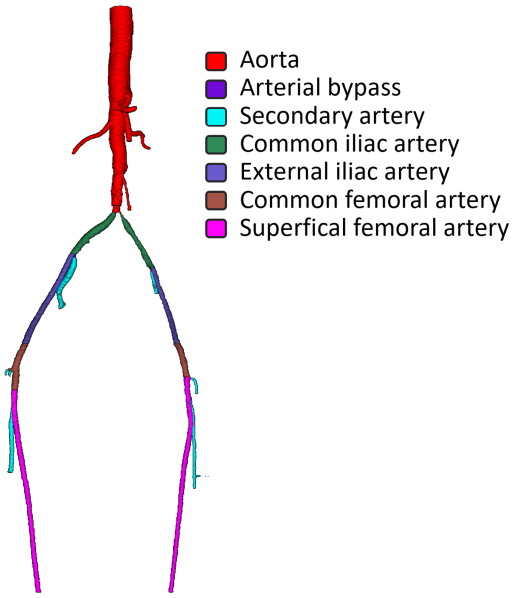


Figure 7.4.: Example anatomy of *patient_001*, with secondary artery branches highlighted in cyan.

Table 7.11.: Averaged ratio of calcified artery per branch and probability of occurrence of stents for each branch across PADSET. **C.i .**: Common iliac, **E.i .**: External iliac, **C.f .**: Common femoral, **S.f .**: Superficial femoral, **A.t .**: Anterior tibial, **P.t .**: Posterior tibial.

	Aorta	**C. i.**	**E. i.**	**C. f.**	**S. f.**	**Popliteal**	**TPT**	**A. t.**	**P.t.**	**Fibular**
Calcified	0.07	0.25	0.20	0.15	0.12	0.06	0.15	0.12	0.06	0.09
Stent	0.12	0.10	0.17	0.11	0.01	0.14	0.01	0.00	0.01	0.00

volumes across all patients in the dataset. For each patient, the cumulative volume of all segmented calcifications was computed, and the resulting density is displayed. This illustrates the heterogeneity of arterial calcification in PAD and highlights the ability of the segmentation approach to quantify patient-level disease burden.

In addition, stent locations were systematically retrieved from the segmentation masks. Across the annotated dataset, 42% of scans contained at least one stent, with patients having on average 1.67 stents. The frequency of stent placement was further reported per arterial branch in Table 7.11, highlighting common intervention sites.

Together, these features demonstrate how the proposed framework supports automated, branch-specific quantification of vascular anatomy and pathology, thereby enabling integration into standardized imaging-based scoring systems and presurgical decision support.

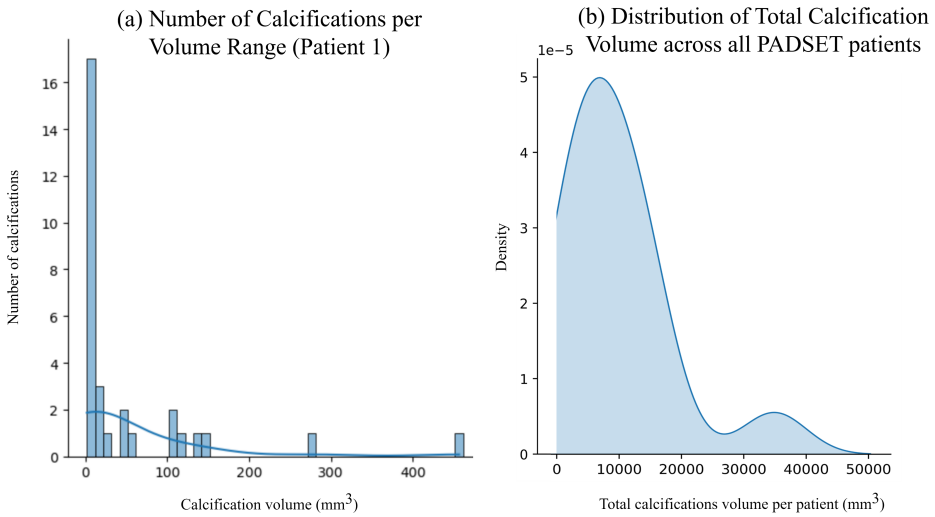


Figure 7.5.: **(a)** Distribution of calcification volumes for *patient_001*, x-axis shows the volume of individual calcifications (mm^3), y-axis indicates the number of calcifications within each volume range. **(b)** Distribution of the total calcification volume across patients in PADSET, x-axis represents the total segmented calcification volume per patient (mm^3), y-axis shows the corresponding density.

7.3 Conclusion

In this chapter, we focused on benchmarking and advancing segmentation methodologies for the segmentation of the lower-limb arteries on the PADSET dataset introduced previously. Importantly, benchmarking against general-purpose pre-trained models revealed poor performance in this domain, underlining the need for a dedicated dataset and models tailored to the lower-limb arterial tree, which has been largely neglected in existing work.

We then evaluated segmentation performance on PADSET using both baseline nnU-Net models and our methodological contributions. Results showed that classes with lower prevalence and poorer annotation consistency, such as distal arteries, remain the most challenging to segment. Despite these challenges, the nnUNet baseline achieves high overall performance for the major arterial structures. Nevertheless, the introduction of the topology-preserving loss clDice, of the Hausdorff losses and SoftMorph morphological filters improved segmentation outcomes in clinically meaningful ways: calcification detection was substantially enhanced with SoftMorph, while artery segmentation benefited from losses that prioritize topological preservation and boundary accuracy.

Finally, we demonstrated the clinical utility of segmentation outputs by extracting quantitative anatomical and pathological descriptors, including artery length, diameter, tortuosity, calcification burden, and stent frequency. These features capture patient-specific pathology and can be directly mapped to standardized scoring systems. Beyond re-

producing existing clinical scores, they also enable consistent quantification of vascular disease, providing clinical insights for pre-surgical planning and revascularization strategy selection.

Overall, this chapter bridges advanced segmentation methodologies with the clinical context of PAD. By explicitly labeling individual branches, integrating plaque and stent detection, and extracting standardized vascular features, our framework provides a reproducible and quantitative basis for presurgical decision support. This work fills a critical gap in the automated analysis of the full lower-limb arterial tree, where no prior dataset or specialized models existed.

In future work, we aim to extend this framework towards improved segmentation of small distal branches, particularly in the below-the-knee region, and to build upon the presented features to enable automatic detection and quantification of stenosis and thrombosis lesions. Together, these advances would establish a comprehensive, automated pipeline for patient-specific PAD assessment and intervention planning.

8 Conclusion

Contents

This final chapter synthesizes the contributions of this thesis and outlines perspectives for future work. We first summarize the methodological and clinical advancements achieved, then highlight opportunities to extend this work, emphasizing avenues that leverage the curated dataset, refined segmentations, and novel analysis pipelines to enhance further diagnosis, surgical planning, and patient-specific evaluation in PAD.

8.1 Summary of the Main Contributions

PAD is a prevalent disease characterized by the obstruction of arteries in the lower limbs. Revascularization strategies are chosen based on the anatomical and pathological characteristics observed on CTA scans. However, its analysis is time-consuming, complex and prone to operator-dependent variability. The main goal of this thesis was to help analyze PAD-relevant features on CTA using recent advances in DL-based automatic segmentation to accelerate and reinforce the decision-making process for the presurgical planning of PAD. Specifically, the work addressed:

Two methodological innovations to improve the segmentation of tubular and small structures: We introduced a novel morphological strategy for deep learning–based segmentation that explicitly incorporates differentiable morphological operators into

convolutional neural networks to promote topological consistency. This strategy, referred to as SoftMorph (Chapter 4), can improve the segmentation of tubular structures. We introduced Regional Hausdorff Distance losses in Chapter 5 to enable distance-aware optimization of segmentation. Both approaches significantly improved the segmentation of tubular and discrete structures.

The translation of the methodological innovations into a real clinical context: We deployed these methods on PADSET, an in-house CTA dataset of patients with PAD in Chapter 7. We curated PADSET with annotations of PAD-relevant structures in the whole lower-limb arterial tree using semi-automatic tools and refinement methods. We benchmarked automatic segmentation methods and validated the clinical relevance of our methodological contributions, achieving high segmentation accuracy. We leveraged segmentation outputs for the automatic retrieval of anatomical and pathological descriptors directly relevant to revascularization strategy planning, providing a reproducible and objective basis for decision-making, reducing reliance on subjective visual assessment.

The work demonstrated that automatic segmentation is not only technically feasible but also provides clinically relevant information, providing a foundation for standardized, quantitative, and patient-specific vascular assessment. We summarize in more detail each contribution below.

8.1.1 SoftMorph: Differentiable Probablistic Morphological Operators for Image Segmentation

In Chapter 4, we introduced SoftMorph, a novel framework to integrate morphological operations directly into CNNs. While original morphological filters are not differentiable and therefore do not support the gradient-based optimization used in CNNs, previous works have proposed to integrate specific operations through various methods. However, these methods can produce homotopy errors or are hard to translate to complex operations. To address these gaps, SoftMorph proposes a generic method to accurately extend any binary morphological operator with any structuring element into a differentiable and probabilistic equivalent.

Specifically, we proposed a definition of probabilistic morphological filters as the expectation of the filter over the probability of generating each possible binary configuration, that can be expressed as a multi-linear polynomial deduced from its truth table. We then presented a family of quasi-probabilistic operators for intractable truth tables, based on fuzzy logic. The quasi-probabilistic operators represent a direct translation of the Boolean expression into a differentiable expression. They approximate the probabilistic filter while maintaining the computational complexity of the original binary one. We

validated SoftMorph in a binary context to ensure that the proposed operators accurately correspond to their binary counterparts, and assessed their backpropagation capability. Finally, we demonstrated the utility of SoftMorph through various applications, notably in segmentation tasks by integrating SoftMorph operators as final layers of a CNN or inside the loss function. Our experiments revealed that its integration into CNNs could improve the topological preservation of segmentation prediction for tubular structures.

Despite these advances, some limitations can be noted. While the quasi-probabilistic operators exactly replicate the complexity of the Boolean formula underlying the operation, certain formulas, such as skeletonization, are intrinsically very complex and can lead to a large increase in CNN training time. Moreover, tuning the structuring element and selecting the fuzzy operator type is not straightforward and usually requires empirical testing to adapt to the dataset. In the same way, there is no clear guideline on which morphological layer will improve segmentation results. More work is therefore needed to establish intrinsic rules for parameter selection based on dataset characteristics. One possible improvement would be to extend the approach to non-flat (fuzzy) structuring elements and to allow the learning of new morphological operations by directly optimizing both the logical formula and the structuring element with fuzzy operators.

Nonetheless, SoftMorph constitutes a generic framework applicable to any segmentation problem where a morphological operation may be required by bridging the gap between rigid mathematical morphology and CNNs. It offers a new paradigm for shape-aware medical image analysis, offering the possibility to define novel operations tailored to the requirements of specific tasks.

8.1.2 Regional Hausdorff Distance Losses for Medical Image Segmentation

Chapter 5 presents the second methodological contribution, introducing three formulations of regional Hausdorff Distance losses. The HD is a boundary-sensitive metric widely employed in medical image analysis to assess segmentation quality, particularly for small or geometrically intricate anatomical structures. Despite its importance, the HD is rarely directly optimized during CNN training. Previous attempts have proven unstable, often relying on auxiliary losses, and do not fully address the differentiability of the underlying Euclidean distance function, limiting reliable gradient backpropagation during CNN optimization. Although differentiable distance transform methods exist, they are prone to numerical instability or require soft binarization of segmentation outputs. To overcome these challenges, we developed a family of regional HD losses, implemented via differentiable erosion-based distance transforms derived from a direct application of the SoftMorph framework.

In detail, we introduced a novel morphological erosion-based differentiable distance transform formulation to compute the signed, unsigned, and positive distance maps in a differentiable manner and directly applicable to probability maps. Based on the formulation of the differentiable positive distance transform, we derive three well-established variations of the HD to propose smooth formulations of the Hausdorff, Modified Hausdorff and the Symmetric Averaged Hausdorff Distance losses. We assessed the effect of connectivity hyperparameters and the applicability of the losses through various 2D and 3D medical segmentation tasks. We demonstrated that these losses enable CNNs to effectively minimize boundary discrepancies without requiring auxiliary losses and with only 2 erosion iterations, showing consistent improvements in distance-based metrics while maintaining or sometimes improving the Dice score.

However, no single formulation consistently outperformed the others, which complicates the choice of the optimal loss function. Additionally, increasing the number of erosion iterations in the averaged HD losses led to numerical instabilities and substantially increased computational cost. Future work should investigate strategies to mitigate these instabilities, such as exploring alternative numbers of erosion iterations, alternative normalization terms, or directly regressing the distance transform using the CNN.

Despite these limitations, this chapter establishes a methodological framework for distance-aware optimization, broadening the range of loss functions available for medical image segmentation beyond traditional overlap-based losses.

8.1.3 Clinical application

The third contribution translated the methodological advances into a realistic and clinically relevant setting for PAD. For this purpose, we curated PADSET, a dedicated dataset of lower-limb CTA scans with comprehensive annotations of the full arterial tree, calcification plaques, stents, and branch-level identities, presented in Chapter 6. Annotations were performed using a semi-automatic workflow in close collaboration with vascular surgeons, ensuring reproducibility and consistency while reducing inter-annotator variability.

In Chapter 7, we systematically investigated the performance of the 3D nnU-Net framework, exploring multiple configurations to identify optimal settings for high-resolution, large-volume PADSET images. In addition, we benchmarked other automatic segmentation tools that required minimal or no training to assess their ability to generalize to this specific anatomical context. We further integrated skeleton-based topology preserving losses to enforce topological preservation, where the clDice contributed to improved segmentation of arteries, calcifications and stents. Methodological contributions such as

SoftMorph and regional Hausdorff losses enhanced the detection of calcifications and refined arterial topology, confirming their practical value on real clinical data.

Despite these advances, several limitations remain. Distal arteries remain challenging to segment due to their small caliber, imaging artifacts, and the difficulty of producing fully consistent manual annotations, which can result in small vessel discontinuities that may be misinterpreted as occlusions rather than segmentation errors. Bypass grafts also pose a challenge because of their low representation in the dataset and high anatomical variability. Furthermore, there is currently no external dataset available for independent validation of the methods. Finally, the full potential of the proposed methodological contributions could not be assessed under maximal training configurations due to the high resolution and large size of the PADSET images, which require substantial computational resources.

Nonetheless, this work establishes a robust foundation for automated lower-limb arterial segmentation and pathology quantification in PAD. It demonstrates the feasibility of integrating advanced neural architectures and topology-aware losses into clinically meaningful workflows, providing a strong basis for future developments in imaging-based decision support and personalized pre-surgical planning.

8.2 Perspectives and Future Applications

While the methods and clinical applications developed in this thesis provide a comprehensive framework for automated segmentation and pathology quantification in PAD, they also open new avenues for further research. The approaches presented highlight both opportunities and challenges to refine morphological operations, enhance vessel continuity, and leverage segmentation outputs for higher-level clinical tasks. In the following, we outline several promising directions that could extend and build upon this foundation.

8.2.1 Learning morphological operation

Deep morphological neural networks (DMNNs) have emerged to replace convolutional layers in CNNs by morphological operations [Mondal, 2022; Hirata, 2021]. In some of these applications, the flat or non-flat structuring element is optimized [Nakashizuka, 2009; Shih, 2019; Shen, 2022], or the appropriate sequence of erosion and dilation is also learned [Shen, 2022; Mondal, 2020; Masci, 2013] based on the target data and task. However, the operations are always selected between erosion and dilation operations only.

A natural extension of the SoftMorph framework presented in Chapter 4 is to move beyond manually defined morphological operations and toward learning optimal logical rules and structuring elements directly from data. Currently, SoftMorph relies on pre-defined logical formulas to implement operations such as dilation, erosion, or skeletonization. While it allows precise control, it limits flexibility and may not always capture dataset-specific nuances. Future work could formalize a differentiable optimization process in which the Boolean logic itself becomes a trainable parameter. Each candidate Boolean formula could be parameterized as a probabilistic combination of truth table entries or as a differentiable logical function using fuzzy operators, allowing the network to select the operation that minimizes a loss function. This approach is inherently related to learning the structuring element, since it is determined by the logical expression itself. This extension could introduce the definition of completely new morphological operations learned for a specific data or task beyond traditional sequences of erosion and dilation.

This line of work is highly relevant as it bridges the gap between human-designed morphological intuition and data-driven optimization, offering the possibility of dataset-adaptive, task-specific morphological filtering. Moreover, learning these operators could directly address the limitations observed in this thesis, such as spurious disconnections in small vessels or inconsistent labeling at bifurcations, without requiring extensive manual tuning.

8.2.2 Improved segmentation for artery continuity

Despite the overall high performance of the 3D nnU-Net framework on PADSET and the integration of the topology-preserving clDice loss [Shit, 2021], small-caliber distal arteries and tortuous vessels remain prone to spurious disconnections and small holes in the segmentation mask. These errors, while minor in voxel count, can critically affect downstream applications such as branch-specific feature extraction, thrombosis detection, and surgical planning, where precise connectivity and anatomical continuity are essential.

To address these challenges, future work could explore morphological and graph-based post-processing strategies specifically designed to correct disconnected components without violating true anatomical occlusions. Traditional methods have explored similar ideas using geodesic minimum spanning trees to infer vascular topology from angiographic images [Moriconi, 2018], demonstrating how quantitative features such as location, direction, scale, and bifurcations can be used to reconstruct continuous vessel trees. Recent works use walk algorithms to reconnect broken vessel segments [Qiu, 2023], but can enforce reconnection of broken vessels at thrombosis locations where the network should not be reconnected.

More recent deep learning–based methods explicitly aim to restore vessel continuity. For instance, [Cavicchioli, 2025] proposed D2-RD-UNet, which applies a centerline-based post-processing step to reconnect fragmented hepatic vessels while also providing local radius-based branch evaluation. Similarly, [Carneiro-Esteves, 2024] introduced a learned post-processing model capable of restoring connectivity.

Graph-based approaches provide a complementary solution by representing the segmented vasculature as a node-and-edge structure, where nodes correspond to centerline points and edges encode anatomical connectivity. In this framework, spurious disconnections appear as missing edges, which can be selectively reconstructed based on vessel diameter, orientation, and local connectivity priors. Modern deep learning approaches, including graph neural networks (GNNs), extend these concepts to capture the global arterial connectivity [Shin, 2019]. In [Esmaeilzadeh, 2025], a loss function is designed to penalize both false positive and false negative connections by optimizing a graph connectivity metric.

The application of these methods in the context of PAD could enhance the clinical reliability of automated segmentation, ensuring that derived metrics such as branch length, tortuosity, and diameter are accurate and free from artifacts caused by mistakenly disconnected voxels. It also strengthens the utility of the segmentation outputs for higher-level applications, including automated detection of stenosis and thrombosis or computational hemodynamic modeling, for instance.

8.2.3 Branch Identification

In Chapter 7, we approached the task of branch identification as a multi-task segmentation problem performed directly on CTA scans. While effective for major vessels, this approach remains challenging for underrepresented branches in the dataset and for small-caliber vessels. An alternative strategy could rely on artery segmentation as the image on which the branches are then segmented. However, our experiments suggest that anatomical context beyond the vessel lumen is often required to reliably identify specific branches, particularly in regions with complex bifurcations.

Several methods for branch recognition have been proposed in other vascular territories, such as the circle of Willis [Yang, 2024; Essadik, 2022; Nader, 2023]. Most of these approaches rely on segmentation networks, while some build upon a preliminary segmentation and exploit geometric or topological cues. For instance, [Nader, 2023] extracts a 3D graph from the vessel mask to detect bifurcation point coordinates, selects local 3D patches in the image around these coordinates, and classifies each bifurcation patch using a dedicated classifier. Similarly, [Essadik, 2022] derives geometric descriptors of bifurcations (e.g., vessel diameter, length, and branching angles), applies dimensionality

reduction, and trains a classifier on the resulting features to identify each bifurcation. In [Riffaud, 2022], aortic branches are identified by extracting geometric and topological features from an artery segmentation and matching them to known reference branches, from the celiac artery to the iliac ones, using similarity measures.

Other traditional methods have explored the explicit modeling of vascular connectivity and anatomical labeling through graph representations. For example, approaches for the Circle of Willis have represented the vasculature as a rooted attributed relational graph, where bifurcations are vertices characterized on a Riemannian manifold. Anatomical labels are then assigned using a maximum a posteriori estimate that combines local bifurcation features with prior knowledge of the global topology [Bogunović, 2013]. Another strategy jointly optimizes segmentation and labeling by constructing an overcomplete vascular graph and selecting the subset of edges that best represent the true vasculature. This is formulated as an integer program, balancing image evidence and connectivity priors, and has been shown to outperform standard sequential segmentation-labeling pipelines [Robben, 2014; Robben, 2016]. These works illustrate the value of explicitly leveraging graph-based models and topological priors for robust branch identification.

Such strategies have not yet been systematically applied to the lower-limb arterial tree. A key difficulty arises from the pathological characteristics of PAD as occlusions and calcifications frequently disrupt vessel continuity, complicating graph construction. Moreover, not all branch distinctions arise from bifurcations. This is the case of the superficial femoral and popliteal arteries, which form a continuous segment. If these challenges are overcome, branch recognition could leverage geometric and topological features to assign each segment its most probable label within expected anatomical ranges. This would provide a robust alternative or complement to direct multi-task learning, particularly for underrepresented branches where annotated training data remain scarce.

8.2.4 Automated detection of stenosis and thrombosis

Building upon the comprehensive segmentation outputs obtained in Chapter 7, a natural extension is the automated detection of arterial stenosis and thrombosis, which are clinically critical for guiding revascularization strategies in PAD patients.

In [Dai, 2021], a parallel efficient network (p-EffNet) was used to classify the degree of stenosis of lower extremity arteries on 2D axial slices of CTA. The model operates on smaller, manually defined regions of interest (ROIs) cropped from the full image using a prior segmentation. Our automatic segmentation can serve as an automatic way to extract ROIs to apply this method. In [Jin, 2022], radiomic features were extracted from an initial artery segmentation and plaque detection for the coronary arteries on CTA.

The radiomic features were used to train classifiers for plaque classification and stenosis grading. This method could be similarly tested for PAD.

Otherwise, the approach could leverage the branch-wise labeled arterial masks to quantify local vessel diameters along each segment. By computing the diameter at each centerline voxel or along small overlapping windows, the relative reduction in lumen can be assessed compared to proximal or reference segment diameter, allowing the identification of stenosis when diameter reductions exceed clinically relevant thresholds (between 50–70% of diameter reduction). Occlusions due to thrombosis can be inferred as abrupt interruptions in the artery mask, especially when the precise segmentation continuity is preserved, as the central challenge arises from false positives in small-caliber vessels, where segmentation errors or minor disconnections could mimic stenoses or occlusions.

An alternative or complementary strategy involves direct deep learning detection of stenosis and thrombosis on raw CTA volumes, bypassing segmentation errors entirely. CNNs or vision transformer-based architectures could be trained to predict localized diameter reduction or occlusion probability maps from the native image, similar to [Zhang, 2020] in coronary artery disease or [Qiu, 2022] for intracranial arteries, where stenosis is detected directly from CTA without explicit lumen segmentation. In this scenario, the branch-level segmentation masks could still provide anatomical priors to constrain predictions, improving localization and reducing false positives in areas prone to artifacts, such as distal arteries or calcified segments.

This research direction can provide quantitative indicators of lesion severity and support surgical planning, improving reproducibility, and providing a foundation for downstream predictive modeling of disease progression and treatment outcomes.

8.2.5 Detection and prognostic modeling using Imaging features

As PAD remains underdiagnosed and undertreated [Criqui, 2015], several studies have tried to employ AI models to diagnose the disease or to predict the likely course and outcomes of patients [Goffart, 2025c; Goffart, 2025b]. While these approaches have shown encouraging performance for predicting complications such as mortality, limb adverse events, major bleeding, or progression to CLTI, they often rely on limited or manually curated input features. In particular, detailed image-based vascular descriptors are not currently used, which restricts the models' ability to fully capture patient-specific anatomical risk factors.

The segmentation outputs generated by our framework provide a rich set of anatomical and pathological features, including vessel diameters, lengths, tortuosity, plaque distribution, and stent locations. These imaging-derived descriptors can be directly used as input for predictive models, for example, to estimate patient-specific outcomes such as survival, risk of restenosis, or post-operative complications. Implementing models that rely solely on these automatically extracted imaging features would already constitute a practical and clinically useful application of the proposed framework.

As an extension, these imaging features could be combined with non-imaging clinical variables (e.g., age, comorbidities, smoking status, previous interventions) to build more comprehensive predictive models. Approaches such as gradient-boosted trees, random forests, or deep learning architectures capable of handling heterogeneous data could integrate these multimodal features. This combined strategy is likely to improve predictive performance and enable a complete, patient-specific risk assessment.

Appendix Chapter 4

A

A.1 Thinning formula

A.1.1 2D

The morphological thinning on 2D images proposed by Wagner *et al.* [Wagner, 2020] is divided into 4 sub-iterations, one for each North, East, South, and West directions. It consists of the following Boolean function for the North direction defined on the $k = 8+1$ neighborhood as :

$$
\begin{aligned}
F1 &= \neg Y_6 \wedge \neg Y_7 \wedge (Y_2 \oplus Y_5 \oplus \neg Y_1) \wedge (Y_5 \oplus Y_8) \wedge (Y_5 \oplus Y_4) \\
F2 &= (Y_6 \oplus Y_8) \wedge (Y_2 \oplus Y_8) \wedge (Y_1 \oplus \neg Y_8) \wedge \neg Y_4 \wedge \neg Y_7 \\
F3 &= Y_2 \wedge Y_1 \wedge \neg Y_7 \\
F4 &= \neg Y_6 \wedge Y_1 \wedge Y_4 \\
F5 &= Y_2 \wedge Y_1 \wedge Y_4 \\
F6 &= \neg Y_6 \wedge \neg Y_2 \wedge \neg Y_5 \wedge Y_4 \wedge Y_7 \\
F &= \neg Y_3 \wedge (F1 \vee F2 \vee F3 \vee F4 \vee F5 \vee F6).
\end{aligned}
$$

The Boolean formula must be rotated for the 3 other directions.

A.1.2 3D

Boolean formula defining the sub-iterative North 3D thinning operation on a $k = 26 + 1$ neighborhood :

$$F1 = \neg Y_0 \wedge \neg Y_1 \wedge \neg Y_2 \wedge \neg Y_3 \wedge \neg Y_5 \wedge \neg Y_6 \wedge \neg Y_7 \wedge \neg Y_8 \wedge Y_{25} \wedge (Y_9 \vee Y_{10} \vee Y_{11} \vee Y_{12} \vee Y_{13} \vee Y_{14} \vee Y_{15} \vee Y_{16} \vee Y_{17} \vee Y_{18} \vee Y_{19} \vee Y_{20} \vee Y_{21} \vee Y_{22} \vee Y_{23} \vee Y_{24})$$

$$F2 = \neg Y_8 \wedge Y_{25} \wedge ((Y_3 \vee Y_4 \vee Y_5 \vee Y_6 \vee Y_7 \vee \neg Y_{10}) \vee (Y_1 \vee Y_2 \vee Y_3 \vee Y_4 \vee Y_5 \vee \neg Y_{16}) \vee (Y_0 \vee Y_1 \vee Y_2 \vee Y_3 \vee Y_7 \vee \neg Y_{14}) \vee (Y_0 \vee Y_1 \vee Y_5 \vee Y_6 \vee Y_7 \vee \neg Y_{12}))$$

$$F3 = \neg Y_8 \wedge Y_{25} \wedge ((Y_5 \vee Y_6 \vee Y_7 \vee \neg Y_{10} \vee \neg Y_{12}) \vee (Y_5 \vee Y_4 \vee Y_3 \vee \neg Y_{10} \vee \neg Y_{16}) \vee (Y_1 \vee Y_2 \vee Y_3 \vee \neg Y_{16} \vee \neg Y_{14}) \vee (Y_0 \vee Y_1 \vee Y_7 \vee \neg Y_{14} \vee \neg Y_{12}))$$

$$F4 = \neg Y_1 \wedge \neg Y_3 \wedge \neg Y_5 \wedge \neg Y_7 \wedge \neg Y_8 \wedge Y_{25} \wedge ((Y_0 \vee Y_4 \vee Y_6 \vee \neg Y_{11} \vee \neg Y_2) \vee (Y_2 \vee Y_4 \vee Y_6 \vee \neg Y_0 \vee \neg Y_9) \vee (Y_0 \vee Y_4 \vee Y_2 \vee \neg Y_6 \vee \neg Y_{15}) \vee (Y_0 \vee Y_6 \vee Y_2 \vee \neg Y_4 \vee \neg Y_{13}))$$

$$F5 = \neg Y_0 \wedge \neg Y_1 \wedge \neg Y_2 \wedge \neg Y_3 \wedge \neg Y_4 \wedge \neg Y_5 \wedge \neg Y_6 \wedge \neg Y_7 \wedge \neg Y_8 \wedge \neg Y_{25} \wedge ((Y_{13} \vee Y_{14} \vee Y_{15} \vee Y_{21} \vee Y_{22} \vee Y_{23} \vee \neg Y_{18} \vee (Y_9 \vee Y_{10} \vee Y_{11} \vee Y_{12} \vee Y_{16} \vee Y_{17} \vee Y_{19} \vee Y_{20} \vee Y_{24})) \vee (Y_{13} \vee Y_{12} \vee Y_{11} \vee Y_{21} \vee Y_{20} \vee Y_{19} \vee \neg Y_{24} \vee (Y_9 \vee Y_{10} \vee Y_{14} \vee Y_{15} \vee Y_{16} \vee Y_{17} \vee Y_{18} \vee Y_{22} \vee Y_{23})) \vee (Y_9 \vee Y_{10} \vee Y_{11} \vee Y_{17} \vee Y_{18} \vee Y_{19} \vee \neg Y_{22} \vee (Y_{15} \vee Y_{14} \vee Y_{13} \vee Y_{12} \vee Y_{16} \vee Y_{23} \vee Y_{21} \vee Y_{20} \vee Y_{24})) \vee (Y_9 \vee Y_{16} \vee Y_{15} \vee Y_{24} \vee Y_{17} \vee Y_{23} \vee \neg Y_{20} \vee (Y_{14} \vee Y_{10} \vee Y_{11} \vee Y_{12} \vee Y_{13} \vee Y_{18} \vee Y_{19} \vee Y_{22} \vee Y_{21})))$$

$$F6 = \neg Y_0 \wedge \neg Y_1 \wedge \neg Y_2 \wedge \neg Y_3 \wedge \neg Y_4 \wedge \neg Y_5 \wedge \neg Y_6 \wedge \neg Y_7 \wedge \neg Y_8 \wedge \neg Y_{25} \wedge ((Y_{14} \vee Y_{15} \vee Y_{16} \vee Y_{22} \vee Y_{23} \vee Y_{24} \vee \neg Y_{18} \vee \neg Y_{20}) \vee (Y_{14} \vee Y_{12} \vee Y_{13} \vee Y_{22} \vee Y_{20} \vee Y_{21} \vee \neg Y_{18} \vee \neg Y_{24}) \vee (Y_{10} \vee Y_{11} \vee Y_{12} \vee Y_{18} \vee Y_{19} \vee Y_{20} \vee \neg Y_{24} \vee \neg Y_{22}) \vee (Y_9 \vee Y_{10} \vee Y_{16} \vee Y_{24} \vee Y_{17} \vee Y_{18} \vee \neg Y_{22} \vee \neg Y_{20}))$$

$$F = \neg(F1 \vee F2 \vee F3 \vee F4 \vee F5 \vee F6)$$

This formula must be rotated in the other 5 directions for complete thinning operation.

A.2 Generated dataset

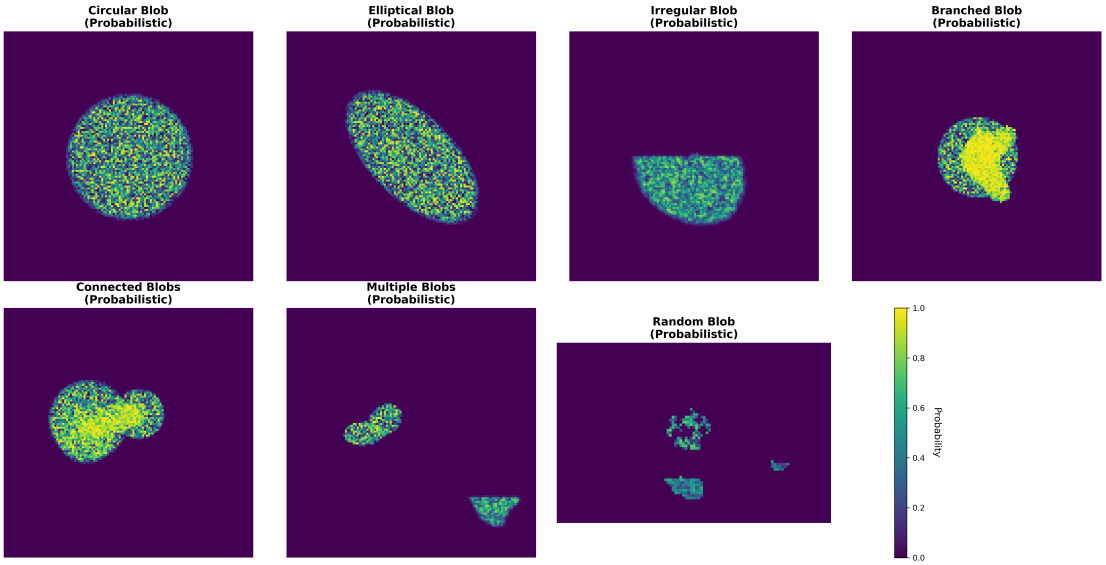


Figure A.1.: Example blobs from the controlled randomly generated dataset.

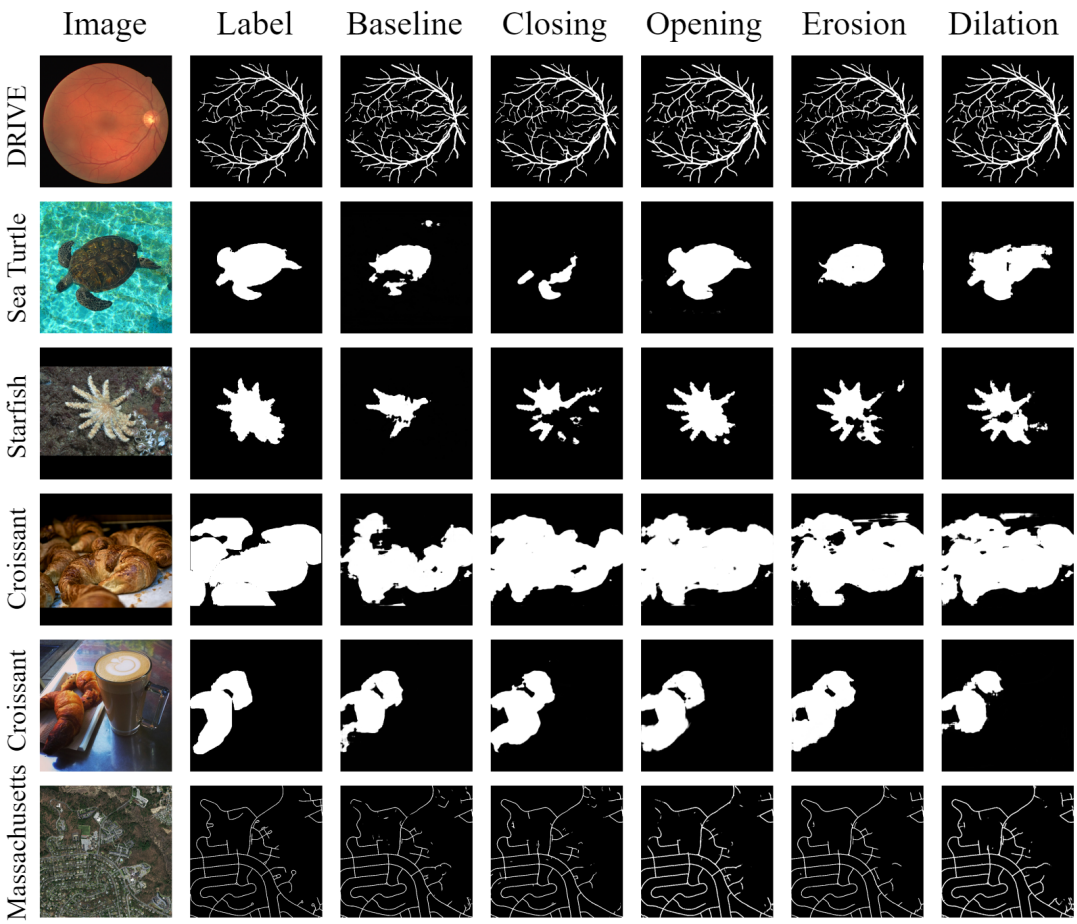


Figure A.2.: Qualitative segmentation results from the DRIVE, Sea Turtle, Starfish, Croissant, and Massachusetts datasets. This figure compares the real image (Image), the ground truth segmentation mask (Label), the segmentation output from the U-Net without any morphological layers (Baseline), and the U-Net outputs with double final morphological layers using different operators (Closing, Opening, Erosion, Dilation).

A.3 Additional qualitative results

We provide additional visual results of segmentation outputs on the 2D datasets, where a SoftMorph morphological operator is applied twice as final layers of a U-Net in Fig. A.2.

Appendix Chapter 7

B

Table B.1.: nnU-Net segmentation performance of each arterial branch after post-processing.

	Dice ↑	clDice ↑	HD ↓	HD95 ↓	Prec. ↑	Rec. ↑	β_0 ↓	β_1 ↓	β_2 ↓	χ ↓
Aorta	0.96	0.90	402.90	**2.04**	**0.98**	0.94	5.0	8.65	6.89	6.25
Second.	0.68	**0.62**	245.90	52.01	0.75	0.65	5.80	0.85	0.30	5.55
C. iliac	0.79	0.84	25.29	8.80	0.81	0.81	0.65	1.25	1.05	0.95
E. iliac	**0.80**	0.87	**38.71**	**11.21**	0.83	0.79	1.55	1.75	0.30	2.10
C. fem.	**0.68**	**0.73**	**41.62**	**27.54**	0.73	**0.70**	**0.90**	0.45	0.05	**1.20**
S. fem.	0.75	**0.79**	**143.29**	**46.05**	0.77	**0.77**	**4.55**	**2.15**	1.50	**6.00**
Popliteal	0.48	0.62	170.96	146.98	**0.79**	0.38	**1.74**	**0.32**	**0.05**	**1.58**
TPT	0.45	**0.59**	35.44	29.01	0.79	0.33	1.52	0.12	0.06	1.47
A. Tib.	0.71	0.80	95.92	21.23	0.88	**0.61**	3.47	0.35	0.00	3.35
P. Tib.	0.32	0.54	73.15	33.92	0.97	0.20	32.71	0.18	0.06	32.82
Fibular	0.63	**0.76**	70.44	50.33	**0.82**	**0.55**	6.24	0.06	0.00	6.29
Bypass	0.18	0.34	862.67	507.00	0.50	0.11	**2.00**	0.50	0.50	**2.00**

Table B.2.: 3D nnU-Net segmentation performances on PADSET comparing the baselines Dice and CE+Dice losses with the proposed HD losses for the artery, calcifications and stents masks with standard deviations.

		Dice ↑	HD ↓	HD95 ↓	clDice ↑	D_pre ↑	D_rec ↑	F1 ↑
Artery	*Dice*	0.86 ± 0.07	442.86 ± 313.58	330.09 ± 266.83	0.61 ± 0.22	-	-	-
	LH	0.85 ± 0.08	377.87 ± 261.50	237.56 ± 226.47	0.58 ± 0.20	-	-	-
	LAH	0.86 ± 0.07	369.41 ± 322.46	251.33 ± 266.47	0.61 ± 0.21	-	-	-
	LAH sym	0.87 ± 0.08	300.16 ± 281.34	189.62 ± 237.64	0.63 ± 0.21	-	-	-
	CE + Dice	0.89 ± 0.05	337.86 ± 299.73	214.45 ± 255.73	0.64 ± 0.20	-	-	-
	CE + LH	0.91 ± 0.04	129.64 ± 41.10	15.01 ± 16.45	0.78 ± 0.06	-	-	-
	CE + LAH	0.87 ± 0.07	340.26 ± 309.26	223.77 ± 265.96	0.63 ± 0.22	-	-	-
	CE + LAH sym	0.92 ± 0.04	118.71 ± 38.75	8.02 ± 8.30	0.83 ± 0.05	-	-	-
Calcif.	*Dice*	0.65 ± 0.13	273.03 ± 76.86	134.52 ± 51.13	0.68 ± 0.12	64.93 ± 13.11	62.55 ± 14.43	67.73 ± 11.58
	LH	0.40 ± 0.24	442.09 ± 257.17	193.90 ± 248.47	0.43 ± 0.26	30.86 ± 18.95	83.89 ± 27.30	43.40 ± 22.18
	LAH	0.69 ± 0.18	181.12 ± 85.38	78.16 ± 80.59	0.71 ± 0.19	78.76 ± 12.42	73.54 ± 20.10	74.55 ± 15.35
	LAH sym	0.75 ± 0.14	193.57 ± 94.77	39.97 ± 55.80	0.76 ± 0.15	57.02 ± 18.55	72.11 ± 20.91	62.02 ± 18.89
	CE + Dice	0.78 ± 0.12	170.51 ± 99.51	35.71 ± 54.25	0.78 ± 0.13	77.50 ± 14.17	82.83 ± 12.67	79.07 ± 10.57
	CE + LH	0.61 ± 0.17	191.26 ± 94.71	56.22 ± 62.96	0.60 ± 0.20	57.03 ± 18.55	72.11 ± 20.91	62.02 ± 18.89
	CE + LAH	0.72 ± 0.13	171.54 ± 100.26	49.15 ± 63.38	0.76 ± 0.13	70.29 ± 14.12	84.19 ± 11.64	75.73 ± 10.65
	CE + LAH sym	0.78 ± 0.13	215.57 ± 190.98	30.06 ± 54.27	0.79 ± 0.13	80.31 ± 11.88	83.68 ± 11.55	81.31 ± 9.58
Stents	*CE + Dice*	0.62 ± 0.28	459.81 ± 340.76	235.12 ± 366.31	0.62 ± 0.31	63.85 ± 35.18	33.61 ± 30.82	38.00 ± 24.90
	CE + LAH sym	0.64 ± 0.29	214.56 ± 188.09	47.56 ± 76.18	0.67 ± 0.32	63.86 ± 35.18	68.47 ± 38.07	55.85 ± 35.49

Bibliography

[Abraham, 2019] Nabila Abraham and Naimul Mefraz Khan. “A novel focal tversky loss function with improved attention u-net for lesion segmentation”. In: *2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019)*. IEEE. 2019, pp. 683–687 (cit. on p. 27).

[AbuRahma, 2013] AF AbuRahma and JE Campbell. *Overview of peripheral arterial disease of the lower extremity. Noninvasive vascular diagnosis*. 2013 (cit. on p. 12).

[Amutha, 2011] S. Amutha, D.R Ramesh Babu, M. Ravi Shankar, and N. Harish Kumar. “Mammographic image enhancement using modified mathematical morphology and Bi-orthogonal wavelet”. In: *2011 IEEE International Symposium on IT in Medicine and Education*. Vol. 1. Dec. 2011, pp. 548–553 (cit. on p. 43).

[Anaya-Isaza, 2021] Andrés Anaya-Isaza, Leonel Mera-Jiménez, and Martha Zequera-Diaz. “An overview of deep learning in medical imaging”. In: *Informatics in Medicine Unlocked* 26 (2021), p. 100723 (cit. on p. 2).

[Angulo, 2003] Jesús Angulo and Jean Serra. “Automatic analysis of DNA microarray images using mathematical morphology”. In: *Bioinformatics* 19.5 (Mar. 2003), pp. 553–562 (cit. on p. 43).

[Angulo, 2008] Jesus Angulo. “A Mathematical Morphology Approach to Cell Shape Analysis”. en. In: *Progress in Industrial Mathematics at ECMI 2006*. Vol. 12. Series Title: Mathematics in Industry. Berlin, Heidelberg, 2008, pp. 543–547 (cit. on p. 43).

[Angulo, 2010] Jesús Angulo. “Pseudo-morphological Image Diffusion Using the Counter-Harmonic Paradigm”. en. In: *Advanced Concepts for Intelligent Vision Systems*. Berlin, Heidelberg, 2010, pp. 426–437 (cit. on p. 45).

[Antonelli, 2022] Michela Antonelli, Annika Reinke, Spyridon Bakas, et al. “The medical segmentation decathlon”. In: *Nature communications* 13.1 (2022), p. 4128 (cit. on pp. 57, 78).

[Anwer, 2025] Ali MOA Anwer, Hacer Karacan, Muhammed Rabee, Levent Enver, and Gonca Cabuk. “Automated Artery Detection and Stenosis Classification in CTA Using Deep Learning for Peripheral Arterial Disease”. In: *IEEE Access* (2025) (cit. on pp. 35, 88, 89).

[Aouad, 2022] Theodore Aouad and Hugues Talbot. “Binary Morphological Neural Network”. In: *2022 IEEE International Conference on Image Processing (ICIP)*. 2022, pp. 3276–3280 (cit. on p. 46).

[Ara, 2019] Lena Ara, Xiao Luo, Alan Sawchuk, and Dave Rollins. “Automate the peripheral arterial disease prediction in lower extremity arterial doppler study using machine learning and neural networks”. In: *Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*. 2019, pp. 130–135 (cit. on p. 35).

[Asano, 2004] A. Asano, Y. Kobayashi, Chie Muraki, and M. Muneyasu. “Optimization of gray scale morphological opening for noise removal in texture images”. en. In: *The 2004 47th Midwest Symposium on Circuits and Systems, 2004. MWSCAS ’04.* Vol. 1. Hiroshima, Japan, 2004, 1_313–1_316 (cit. on p. 43).

[Attias, 2018] David Attias and Nicolas Lellouche. *Cardiologie vasculaire*. VG éditions, 2018 (cit. on p. 18).

[Azad, 2024] Reza Azad, Ehsan Khodapanah Aghdam, Amelie Rauland, et al. “Medical image segmentation review: The success of u-net”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2024) (cit. on p. 25).

[Bagheri Rajeoni, 2023] Alireza Bagheri Rajeoni, Breanna Pederson, Daniel G Clair, Susan M Lessner, and Homayoun Valafar. “Automated measurement of vascular calcification in femoral endarterectomy patients using deep learning”. In: *Diagnostics* 13.21 (2023), p. 3363 (cit. on pp. 2, 35).

[Bai, 2017] Min Bai and Raquel Urtasun. “Deep watershed transform for instance segmentation”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 5221–5229 (cit. on p. 73).

[Bernard, 2018] Olivier Bernard, Alain Lalande, Clement Zotti, et al. “Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved?” In: *IEEE Transactions on Medical Imaging* 37.11 (2018), pp. 2514–2525 (cit. on p. 78).

[Birkett, 2021] Stefan T Birkett, Amy E Harwood, Edward Caldow, et al. “A systematic review of exercise testing in patients with intermittent claudication: A focus on test standardisation and reporting quality in randomised controlled trials of exercise interventions”. In: *PLoS One* 16.5 (2021), e0249277 (cit. on p. 16).

[Björkegren, 2022] Johan LM Björkegren and Aldons J Lusis. “Atherosclerosis: recent developments”. In: *Cell* 185.10 (2022), pp. 1630–1645 (cit. on p. 14).

[Bloch, 1995] Isabelle Bloch and Henri Maitre. “Fuzzy mathematical morphologies: A comparative study”. In: *Pattern Recognition* 28.9 (Sept. 1995), pp. 1341–1387 (cit. on pp. 42, 45, 56).

[Bloch, 2017] Isabelle Bloch. “Chapter 3 - Fuzzy skeleton and skeleton by influence zones: a review”. In: *Skeletonization*. 2017, pp. 71–87 (cit. on p. 43).

[Bloch, 2021] Isabelle Bloch, Samy Blusseau, Ramón Pino Pérez, Élodie Puybareau, and Guillaume Tochon. “On Some Associations Between Mathematical Morphology and Artificial Intelligence”. In: *Discrete Geometry and Mathematical Morphology*. 2021, pp. 457–469 (cit. on pp. 43, 47).

[Bogunović, 2013] Hrvoje Bogunović, José María Pozo, Rubén Cárdenes, Luis San Román, and Alejandro F Frangi. “Anatomical labeling of the circle of willis using maximum a posteriori probability estimation”. In: *IEEE transactions on medical imaging* 32.9 (2013), pp. 1587–1599 (cit. on p. 126).

[Bolliger, 2009] Stephan A Bolliger, Lars Oesterhelweg, Danny Spendlove, Steffen Ross, and Michael J Thali. “Is differentiation of frequently encountered foreign bodies in corpses possible by Hounsfield density measurement?” In: *Journal of forensic sciences* 54.5 (2009), pp. 1119–1122 (cit. on p. 97).

[Bollinger, 1981] A Bollinger, K Breddin, H Hess, et al. “Semiquantitative assessment of lower limb atherosclerosis from routine angiographic images”. In: *Atherosclerosis* 38.3-4 (1981), pp. 339–346 (cit. on p. 20).

[Buckley, 2017] Chris Buckley and Ellen M Voorhees. “Evaluating evaluation measure stability”. In: *Acm Sigir Forum*. Vol. 51. 2. ACM New York, NY, USA. 2017, pp. 235–242 (cit. on p. 29).

[Bui, 2019] Toan Duc Bui, Li Wang, Jian Chen, et al. “Multi-task Learning for Neonatal Brain Segmentation Using 3D Dense-Unet with Dense Attention Guided by Geodesic Distance”. In: *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*. Cham: Springer International Publishing, 2019, pp. 243–251 (cit. on p. 73).

[Bumgarner, 2022] Jacob R. Bumgarner and Randy J. Nelson. “Open-source analysis and visualization of segmented vasculature datasets with VesselVio”. In: *Cell Reports Methods* 2.4 (2022), p. 100189 (cit. on p. 113).

[Bunch, 1977] Philip C Bunch, John F Hamilton, Gary K Sanderson, and Arthur H Simmons. “A free response approach to the measurement and characterization of radiographic observer performance”. In: *Application of optical instrumentation in medicine VI*. Vol. 127. SPIE. 1977, pp. 124–135 (cit. on p. 29).

[Butoi, 2023] Victor Ion Butoi*, Jose Javier Gonzalez Ortiz*, Tianyu Ma, et al. “UniverSeg: Universal Medical Image Segmentation”. In: *International Conference on Computer Vision* (2023) (cit. on pp. 100, 102).

[Caliva, 2019] Francesco Caliva, Claudia Iriondo, Alejandro Morales Martinez, Sharmila Majumdar, and Valentina Pedoia. *Distance Map Loss Penalty Term for Semantic Segmentation*. 2019. arXiv: `1908.03679 [eess.IV]` (cit. on pp. 34, 72).

[Cao, 2022] Hu Cao, Yueyue Wang, Joy Chen, et al. “Swin-unet: Unet-like pure transformer for medical image segmentation”. In: *European conference on computer vision*. Springer. 2022, pp. 205–218 (cit. on p. 32).

[Carneiro-Esteves, 2024] Sophie Carneiro-Esteves, Antoine Vacavant, and Odyssée Merveille. "Restoring connectivity in vascular segmentations using a learned post-processing model". In: *International Workshop on Topology-and Graph-Informed Imaging Informatics*. Springer. 2024, pp. 55–65 (cit. on p. 125).

[Cavicchioli, 2025] Matteo Cavicchioli, Andrea Moglia, Guillaume Garret, et al. "D2-RD-UNet: A dual-stage dual-class framework with connectivity correction for hepatic vessels segmentation". In: *Computers in Biology and Medicine* 195 (2025), p. 110530 (cit. on p. 125).

[Chen, 2015] Weifu Chen, Jianrong Xu, and Bernard Chiu. "Fast segmentation of the femoral arteries from 3D MR images: A tool for rapid assessment of peripheral arterial disease". In: *Medical physics* 42.5 (2015), pp. 2431–2448 (cit. on p. 35).

[Chen, 2019] Yo-Chuan Chen, Yi-Chen Lin, Ching-Ping Wang, et al. "Coronary artery segmentation in cardiac CT angiography using 3D multi-channel U-net". In: *arXiv preprint arXiv:1907.12246* (2019) (cit. on p. 31).

[Chen, 2021] Jieneng Chen, Yongyi Lu, Qihang Yu, et al. "Transunet: Transformers make strong encoders for medical image segmentation". In: *arXiv preprint arXiv:2102.04306* (2021) (cit. on p. 32).

[Chen, 2024] Jieneng Chen, Jieru Mei, Xianhang Li, et al. "TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers". In: *Medical Image Analysis* 97 (2024), p. 103280 (cit. on p. 32).

[Cheung, 2020] Mark Cheung, John Shi, Oren Wright, et al. "Graph signal processing and deep learning: Convolution, pooling, and topology". In: *IEEE Signal Processing Magazine* 37.6 (2020), pp. 139–149 (cit. on p. 25).

[Chierici, 2025] Andrea Chierici, Fabien Lareyre, Antonio Iannelli, et al. "Applications of artificial intelligence in liver cancer: A scoping review". In: *Artificial Intelligence in Medicine* (2025), p. 103244 (cit. on p. 7).

[Chowdhury, 2017] Mohammed M Chowdhury, Gregory C Makris, Jason M Tarkin, et al. "Lower limb arterial calcification (LLAC) scores in patients with symptomatic peripheral arterial disease are associated with increased cardiac mortality and morbidity". In: *PloS one* 12.9 (2017), e0182952 (cit. on p. 20).

[Çiçek, 2016] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. "3D U-Net: learning dense volumetric segmentation from sparse annotation". In: *International conference on medical image computing and computer-assisted intervention*. Springer. 2016, pp. 424–432 (cit. on p. 25).

[Cimminiello, 2010] Claudio Cimminiello, Claudio Borghi, Serge Kownator, et al. "Prevalence of peripheral arterial disease in patients at non-high cardiovascular risk. Rationale and design of the PANDORA study". In: *BMC cardiovascular disorders* 10.1 (2010), p. 35 (cit. on p. 15).

[Clough, 2020] James R Clough, Nicholas Byrne, Ilkay Oksuz, et al. “A topological loss function for deep-learning based image segmentation using persistent homology”. In: *IEEE transactions on pattern analysis and machine intelligence* 44.12 (2020), pp. 8766–8778 (cit. on pp. 2, 33).

[Collins, 2007] Ros Collins, G Cranny, J Burch, et al. “A systematic review of duplex ultrasound, magnetic resonance angiography and computed tomography angiography for the diagnosis and assessment of symptomatic, lower limb peripheral arterial disease.” In: *Health Technology Assessment (Winchester, England)* 11.20 (2007), pp. iii–184 (cit. on p. 17).

[Committee, 2015] TASC Steering Committee*, Michael R Jaff, Christopher J White, et al. “An update on methods for revascularization and expansion of the TASC lesion classification to include below-the-knee arteries: a supplement to the Inter-Society Consensus for the Management of Peripheral Arterial Disease (TASC II)”. In: *Journal of Endovascular Therapy* 22.5 (2015), pp. 663–677 (cit. on pp. 2, 20).

[Conte, 2019] Michael S Conte, Andrew W Bradbury, Philippe Kolh, et al. “Global vascular guidelines on the management of chronic limb-threatening ischemia”. In: *European Journal of Vascular and Endovascular Surgery* 58.1 (2019), S1–S109 (cit. on p. 20).

[Copi, 2016] Irving M. Copi, Carl Cohen, and Kenneth McMahon. *Introduction to Logic*. 14th ed. New York: Routledge, July 2016 (cit. on p. 49).

[Criqui, 1996] Michael H Criqui, Julie O Denenberg, Cameron E Bird, et al. “The correlation between symptoms and non-invasive test results in patients referred for peripheral arterial disease testing”. In: *Vascular medicine* 1.1 (1996), pp. 65–71 (cit. on p. 14).

[Criqui, 2015] Michael H Criqui and Victor Aboyans. “Epidemiology of peripheral artery disease”. In: *Circulation research* 116.9 (2015), pp. 1509–1526 (cit. on pp. 1, 127).

[Crum, 2006] W.R. Crum, O. Camara, and D.L.G. Hill. “Generalized Overlap Measures for Evaluation and Validation in Medical Image Analysis”. In: *IEEE Transactions on Medical Imaging* 25.11 (2006), pp. 1451–1461 (cit. on p. 72).

[Dai, 2021] Lisong Dai, Quan Zhou, Hongmei Zhou, et al. “Deep learning-based classification of lower extremity arterial stenosis in computed tomography angiography”. In: *European journal of radiology* 136 (2021), p. 109528 (cit. on pp. 35, 126).

[Dang, 2022] Vien Ngoc Dang, Francesco Galati, Rosa Cortese, et al. “Vessel-CAPTCHA: An efficient learning framework for vessel annotation and segmentation”. In: *Medical Image Analysis* 75 (2022), p. 102263 (cit. on p. 30).

[Dangi, 2019] Shusil Dangi, Cristian A Linte, and Ziv Yaniv. “A distance map regularized CNN for cardiac cine MR image segmentation”. In: *Medical physics* 46.12 (2019), pp. 5637–5651 (cit. on p. 73).

[Deng, 2009] Jia Deng, Wei Dong, Richard Socher, et al. “Imagenet: A large-scale hierarchical image database”. In: *2009 IEEE conference on computer vision and pattern recognition*. Ieee. 2009, pp. 248–255 (cit. on p. 30).

[Dombi, 1982] J. Dombi. “A general class of fuzzy operators, the demorgan class of fuzzy operators and fuzziness measures induced by fuzzy operators”. In: *Fuzzy Sets and Systems* 8.2 (1982), pp. 149–163 (cit. on p. 53).

[Dong, 2022] Caixia Dong, Songhua Xu, and Zongfang Li. “A novel end-to-end deep learning solution for coronary artery segmentation from CCTA”. In: *Medical Physics* 49.11 (2022), pp. 6945–6959 (cit. on p. 30).

[Dosovitskiy, 2020] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, et al. “An image is worth 16x16 words: Transformers for image recognition at scale”. In: *arXiv preprint arXiv:2010.11929* (2020) (cit. on p. 32).

[Du, 2020] Getao Du, Xu Cao, Jimin Liang, Xueli Chen, and Yonghua Zhan. “Medical image segmentation based on U-net: A review.” In: *Journal of Imaging Science & Technology* 64.2 (2020) (cit. on p. 25).

[Dubuisson, 1994] M-P Dubuisson and Anil K Jain. “A modified Hausdorff distance for object matching”. In: *Proceedings of 12th international conference on pattern recognition*. Vol. 1. IEEE. 1994, pp. 566–568 (cit. on p. 76).

[EL Jurdi, 2021] Rosana EL Jurdi, Caroline Petitjean, Paul Honeine, Veronika Cheplygina, and Fahed Abdallah. “A Surprisingly Effective Perimeter-based Loss for Medical Image Segmentation”. In: *Proceedings of the Fourth Conference on Medical Imaging with Deep Learning*. Vol. 143. Proceedings of Machine Learning Research. Conference dates: 07–09 Jul. July 2021, pp. 158–167 (cit. on p. 72).

[Esmaeilzadeh, 2025] Elyar Esmaeilzadeh, Ehsan Garaaghaji, Farzad Hallaji Azad, and Doruk Oner. “CAPE: Connectivity-Aware Path Enforcement Loss for Curvilinear Structure Delineation”. In: *arXiv preprint arXiv:2504.00753* (2025) (cit. on p. 125).

[Essadik, 2022] Ibtissam Essadik, Anass Nouri, Raja Touahni, Romain Bourcier, and Florent Autrusseau. “Automatic classification of the cerebral vascular bifurcations using dimensionality reduction and machine learning”. In: *Neuroscience Informatics* 2.4 (2022), p. 100108 (cit. on p. 125).

[Falcetta, 2025] Daniele Falcetta, Vincenzo Marciano, Kaiyuan Yang, et al. “VesselVerse: A dataset and collaborative framework for vessel annotation”. In: *MICCAI 2025, 28th International Conference on Medical Image Computing and Computer Assisted Intervention, 23-27 September 2025, Daejon, Republic of Korea*. Ed. by Springer. © Springer. Personal use of this material is permitted. The definitive version of this paper was published in MICCAI 2025, 28th International Conference on Medical Image Computing and Computer Assisted Intervention, 23-27 September 2025, Daejon, Republic of Korea and is available at : Daejon, 2025 (cit. on p. 94).

[Falk, 2006] Erling Falk. “Pathogenesis of atherosclerosis”. In: *Journal of the American College of cardiology* 47.8S (2006), pp. C7–C12 (cit. on p. 14).

[Feehs, 1987] Richard J. Feehs and Gonzalo R. Arce. “Multidimensional Morphological Edge Detection”. In: *Visual Communications and Image Processing II*. Vol. 0845. Oct. 1987, pp. 285–292 (cit. on p. 43).

[Fontaine, 1954] R Fontaine. “Surgical treatment of peripheral circulation disorders”. In: *Helv. Chir. Acta.* 21 (1954), pp. 499–533 (cit. on p. 19).

[Fosbinder, 2011] Robert Fosbinder and Denise Orth. *Essentials of radiologic science*. Lippincott Williams & Wilkins, 2011 (cit. on p. 97).

[Fowkes, 2013] F Gerald R Fowkes, Diana Rudan, Igor Rudan, et al. “Comparison of global estimates of prevalence and risk factors for peripheral artery disease in 2000 and 2010: a systematic review and analysis”. In: *The lancet* 382.9901 (2013), pp. 1329–1340 (cit. on p. 15).

[Franchi, 2020] Gianni Franchi, Amin Fehri, and Angela Yao. “Deep morphological networks”. In: *Pattern Recognition* 102 (June 2020), p. 107246 (cit. on pp. 43, 45–47).

[Frank, 2019] Ulrich Frank, Sigrid Nikol, Jill Belch, et al. “ESVM Guideline on peripheral arterial disease”. In: *Vasa* (2019) (cit. on p. 13).

[Gardner, 1991] ANDREW W Gardner, James S Skinner, BRADFORD W Cantwell, and L Kent Smith. “Progressive vs single-stage treadmill tests for evaluation of claudication.” In: *Medicine and science in sports and exercise* 23.4 (1991), pp. 402–408 (cit. on p. 16).

[Gardner, 2007] Andrew W Gardner, Polly S Montgomery, and Azhar Afaq. “Exercise performance in patients with peripheral arterial disease who have different types of exertional leg pain”. In: *Journal of Vascular Surgery* 46.1 (2007), pp. 79–86 (cit. on p. 14).

[Gerhard-Herman, 2017] Marie D Gerhard-Herman, Heather L Gornik, Coletta Barrett, et al. “2016 AHA/ACC guideline on the management of patients with lower extremity peripheral artery disease: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines”. In: *Journal of the American College of Cardiology* 69.11 (2017), e71–e126 (cit. on p. 1).

[Gharaibeh, 2019] Yazan Gharaibeh, David Prabhu, Chaitanya Kolluru, et al. “Coronary calcification segmentation in intravascular OCT images using deep learning: application to calcification scoring”. In: *Journal of Medical Imaging* 6.4 (2019), pp. 045002–045002 (cit. on p. 36).

[Ghodrati, 2022] Vahid Ghodrati, Yair Rivenson, Ashley Prosper, et al. “Automatic segmentation of peripheral arteries and veins in ferumoxytol-enhanced MR angiography”. In: *Magnetic Resonance in Medicine* 87.2 (2022), pp. 984–998 (cit. on pp. 36, 88, 89).

[Goffart, 2025a] Sebastien Goffart, Andréa Chierici, Lisa Guzzi, et al. “Artificial Intelligence to enhance future clinical trials in Vascular Surgery”. In: *Annals of Vascular Surgery* 111 (2025), pp. 331–335 (cit. on pp. 7, 35).

[Goffart, 2025b] Sebastien Goffart, Hervé Delingette, Andrea Chierici, et al. “Artificial Intelligence Techniques for Prognostic and Diagnostic Assessments in Peripheral Artery Disease: A Scoping Review”. In: *Angiology* (2025), p. 00033197241310572 (cit. on pp. 7, 127).

[Goffart, 2025c] Sébastien Goffart, Odette Hart, Fabien Lareyre, et al. “Deep Learning Strategies for Predicting Amputation Risk in Patients With Peripheral Artery Disease”. In: *Journal of Vascular Surgery* 81.6 (2025), e295–e296 (cit. on pp. 7, 127).

[Gordon, 1972] Tavia Gordon and William B Kannel. “Predisposition to atherosclerosis in the head, heart, and legs: the Framingham study”. In: *Jama* 221.7 (1972), pp. 661–666 (cit. on p. 13).

[Graziani, 2007] L Graziani, A Silvestro, V Bertone, et al. “Vascular involvement in diabetic subjects with ischemic foot ulcer: a new morphologic categorization of disease severity”. In: *European Journal of Vascular and Endovascular Surgery* 33.4 (2007), pp. 453–460 (cit. on p. 20).

[Guidi, 2022] Lucas Guidi, Fabien Lareyre, Arindam Chaudhuri, et al. “Automatic measurement of vascular calcifications in patients with aorto-iliac occlusive disease to predict the risk of re-intervention after endovascular repair”. In: *Annals of Vascular Surgery* 83 (2022), pp. 10–19 (cit. on p. 35).

[Guo, 2021] Changlu Guo, Márton Szemenyei, Yugen Yi, et al. “Sa-unet: Spatial attention u-net for retinal vessel segmentation”. In: *2020 25th international conference on pattern recognition (ICPR)*. IEEE. 2021, pp. 1236–1242 (cit. on p. 32).

[Guzzi, 2024] Lisa Guzzi, Maria A. Zuluaga, Fabien Lareyre, et al. “Differentiable Soft Morphological Filters for Medical Image Segmentation”. In: *MICCAI 2024 - Medical Image Computing and Computer Assisted Intervention*. Marrakesh, Morocco, Oct. 2024 (cit. on pp. 6, 42, 65, 76).

[Guzzi, 2025a] Lisa Guzzi, Maria A Zuluaga, Fabien Lareyre, et al. “Automatic Segmentation of Lower-Limb Arteries on CTA for Pre-surgical Planning of Peripheral Artery Disease”. In: *MICCAI-AMAI2025 - 4th Workshop on Applications of Medical Artifical Intelligence*. Daejeon, South Korea, Sept. 2025 (cit. on p. 6).

[Guzzi, 2025b] Lisa Guzzi, Maria A Zuluaga, Riccardo Taiello, et al. “Regional Hausdorff Distance Losses for Medical Image Segmentation”. In: *MLMI 2025 - 16th International Workshop on Machine Learning in Medical Imaging (In conjunction with MICCAI 2025)*. Daejeon, South Korea, Sept. 2025 (cit. on p. 6).

[Hardman, 2011] Rulon L Hardman, Jorge E Lopera, Rex A Cardan, Clayton K Trimmer, and Shellie C Josephs. “Common and rare collateral pathways in aortoiliac occlusive disease: a pictorial essay”. In: *American Journal of Roentgenology* 197.3 (2011), W519–W524 (cit. on p. 13).

[Hardman, 2014] Rulon L Hardman, Omid Jazaeri, J Yi, M Smith, and Rajan Gupta. "Overview of classification systems in peripheral artery disease". In: *Seminars in interventional radiology*. Vol. 31. 04. Thieme Medical Publishers. 2014, pp. 378–388 (cit. on p. 20).

[Hatamizadeh, 2021] Ali Hatamizadeh, Vishwesh Nath, Yucheng Tang, et al. "Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images". In: *International MICCAI brainlesion workshop*. Springer. 2021, pp. 272–284 (cit. on p. 32).

[Hatamizadeh, 2022] Ali Hatamizadeh, Yucheng Tang, Vishwesh Nath, et al. "Unetr: Transformers for 3d medical image segmentation". In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 2022, pp. 574–584 (cit. on p. 32).

[He, 2021] Yufan He, Dong Yang, Holger Roth, Can Zhao, and Daguang Xu. "Dints: Differentiable neural network topology search for 3d medical image segmentation". In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021, pp. 5841–5850 (cit. on p. 31).

[Heller, 2019] Nicholas Heller, Niranjan Sathianathen, Arveen Kalapara, et al. "The kits19 challenge data: 300 kidney tumor cases with clinical context, ct semantic segmentations, and surgical outcomes". In: *arXiv preprint arXiv:1904.00445* (2019) (cit. on p. 89).

[Heller, 2021] Nicholas Heller, Fabian Isensee, Klaus H Maier-Hein, et al. "The state of the art in kidney and kidney tumor segmentation in contrast-enhanced CT imaging: Results of the KiTS19 challenge". In: *Medical image analysis* 67 (2021), p. 101821 (cit. on p. 89).

[Hilbert, 2020] Adam Hilbert, Vince I Madai, Ela M Akay, et al. "BRAVE-NET: fully automated arterial brain vessel segmentation in patients with cerebrovascular disease". In: *Frontiers in artificial intelligence* 3 (2020), p. 552258 (cit. on pp. 30, 31).

[Hilbert, 2022] Adam Hilbert, Jana Rieger, Vince I Madai, et al. "Anatomical labeling of intracranial arteries with deep learning in patients with cerebrovascular disease". In: *Frontiers in Neurology* 13 (2022), p. 1000914 (cit. on p. 2).

[Hirata, 2021] Nina S. T. Hirata and George A. Papakostas. "On Machine-Learning Morphological Image Operators". en. In: *Mathematics* 9.16 (Jan. 2021), p. 1854 (cit. on pp. 47, 123).

[Hirsch, 2001] Alan T Hirsch, Michael H Criqui, Diane Treat-Jacobson, et al. "Peripheral arterial disease detection, awareness, and treatment in primary care". In: *Jama* 286.11 (2001), pp. 1317–1324 (cit. on p. 1).

[Hirsch, 2006] Alan T Hirsch, Ziv J Haskal, Norman R Hertzer, et al. “ACC/AHA 2005 practice guidelines for the management of patients with peripheral arterial disease (lower extremity, renal, mesenteric, and abdominal aortic) a collaborative report from the American Association for Vascular Surgery/Society for Vascular Surgery,* Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, Society of Interventional Radiology, and the ACC/AHA Task Force on Practice Guidelines (writing committee to develop guidelines for the management of patients with peripheral arterial disease): endorsed by the American Association of Cardiovascular and Pulmonary Rehabilitation; National Heart, Lung, and Blood Institute; Society for Vascular Nursing; TransAtlantic Inter-Society Consensus; and Vascular Disease Foundation”. In: *circulation* 113.11 (2006), e463–e654 (cit. on pp. 1, 13).

[Høyer, 2013] Christian Høyer, Jes Sandermann, and Lars J Petersen. “The toe-brachial index in the diagnosis of peripheral arterial disease”. In: *Journal of vascular surgery* 58.1 (2013), pp. 231–238 (cit. on p. 16).

[Huttenlocher, 1993] D.P. Huttenlocher, G.A. Klanderman, and W.J. Rucklidge. “Comparing images using the Hausdorff distance”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15.9 (1993), pp. 850–863 (cit. on p. 72).

[Imran, 2025] Muhammad Imran, Jonathan R Krebs, Vishal Balaji Sivaraman, et al. “Multi-class segmentation of aortic branches and zones in computed tomography angiography: The aortaseg24 challenge”. In: *arXiv preprint arXiv:2502.05330* (2025) (cit. on pp. 30, 34, 88, 89).

[Isensee, 2021] Fabian Isensee, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein. “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation”. In: *Nature methods* 18.2 (2021), pp. 203–211 (cit. on pp. 2, 25, 31, 63, 81, 94, 100, 101).

[Isensee, 2024] Fabian Isensee, Tassilo Wald, Constantin Ulrich, et al. “nnu-net revisited: A call for rigorous validation in 3d medical image segmentation”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2024, pp. 488–498 (cit. on pp. 25, 34, 100).

[Janowitz, 1991] Warren R Janowitz, Arthur S Agatston, and Manuel Viamonte Jr. “Comparison of serial quantitative evaluation of calcified coronary artery plaque by ultrafast computed tomography in persons with and without obstructive coronary artery disease”. In: *The American journal of cardiology* 68.1 (1991), pp. 1–6 (cit. on p. 20).

[Jin, 2022] Xin Jin, Yuze Li, Fei Yan, et al. “Automatic coronary plaque detection, classification, and stenosis grading using deep learning and radiomics on computed tomography angiography images: a multicenter multi-vendor study”. In: *European radiology* 32.8 (2022), pp. 5276–5286 (cit. on p. 126).

[Joskowicz, 2019] Leo Joskowicz, D Cohen, N Caplan, and Jacob Sosna. “Inter-observer variability of manual contour delineation of structures in CT”. In: *European radiology* 29.3 (2019), pp. 1391–1399 (cit. on p. 94).

[Jurdi, 2021] Rosana EL Jurdi, Caroline Petitjean, Paul Honeine, Veronika Cheplygina, and Fahed Abdallah. “A Surprisingly Effective Perimeter-based Loss for Medical Image Segmentation”. en. In: *Proceedings of the Fourth Conference on Medical Imaging with Deep Learning*. Aug. 2021, pp. 158–167 (cit. on pp. 43, 46, 48).

[Karimi, 2019] Davood Karimi and Septimiu E Salcudean. “Reducing the hausdorff distance in medical image segmentation with convolutional neural networks”. In: *IEEE Transactions on medical imaging* 39.2 (2019), pp. 499–513 (cit. on pp. 34, 72, 73, 79, 80).

[Kervadec, 2021] Hoel Kervadec, Jihene Bouchtiba, Christian Desrosiers, et al. “Boundary loss for highly unbalanced segmentation”. In: *Medical Image Analysis* 67 (2021), p. 101851 (cit. on pp. 34, 72).

[Kesari, 2025] Anshika Kesari, Satyajit Maurya, Mohammad Tufail Sheikh, Rakesh Kumar Gupta, and Anup Singh. “Large blood vessel segmentation in quantitative DCE-MRI of brain tumors: A Swin UNETR approach”. In: *Magnetic Resonance Imaging* 118 (2025), p. 110342 (cit. on p. 33).

[Kimori, 2011] Yoshitaka Kimori. “Mathematical morphology-based approach to the enhancement of morphological features in medical images”. en. In: *Journal of Clinical Bioinformatics* 1.1 (Dec. 2011), p. 33 (cit. on p. 43).

[Kirchhoff, 2024] Yannick Kirchhoff, Maximilian R Rokuss, Saikat Roy, et al. “Skeleton recall loss for connectivity conserving and resource efficient segmentation of thin tubular structures”. In: *European Conference on Computer Vision*. Springer. 2024, pp. 218–234 (cit. on pp. 33, 101, 102).

[Klein, 2012] Almar Klein, J Adam van der Vliet, Luuk J Oostveen, et al. “Automatic segmentation of the wire frame of stent grafts from CT data”. In: *Medical image analysis* 16.1 (2012), pp. 127–139 (cit. on p. 36).

[Krasin, 2017] Ivan Krasin, Tom Duerig, Neil Alldrin, et al. “OpenImages: A public dataset for large-scale multi-label and multi-class image classification.” In: (2017). Dataset available from `https://storage.googleapis.com/openimages/web/index.html` (cit. on p. 57).

[Krishnapuram, 1992] Raghu Krishnapuram and Sundeep Gupta. “Morphological methods for detection and classification of edges in range images”. en. In: *Journal of Mathematical Imaging and Vision* 2.4 (Dec. 1992), pp. 351–375 (cit. on p. 43).

[Kruse, 2017] Rombout R Kruse, Denise E Doomernik, Kasper V Maltha, et al. “Collateral artery pathways of the femoral and popliteal artery”. In: *journal of surgical research* 211 (2017), pp. 45–52 (cit. on p. 13).

[Kullo, 2016] Iftikhar J. Kullo and Thom W. Rooke. “Peripheral Artery Disease”. In: *New England Journal of Medicine* 374.9 (2016), pp. 861–871 (cit. on pp. 1, 15).

[Kumar, 2024] Siddharth Krishna Kumar. *GD doesn't make the cut: Three ways that non-differentiability affects neural network training*. arXiv:2401.08426 [cs]. Jan. 2024 (cit. on p. 43).

[Lam, 1992] L. Lam, S.-W. Lee, and C.Y. Suen. "Thinning methodologies-a comprehensive survey". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14.9 (1992), pp. 869–885 (cit. on p. 56).

[Landis, 1977] J Richard Landis and Gary G Koch. "The measurement of observer agreement for categorical data". In: *biometrics* (1977), pp. 159–174 (cit. on p. 91).

[Lareyre, 2021] Fabien Lareyre, Cédric Adam, Marion Carrier, and Juliette Raffort. "Automated segmentation of the human abdominal vascular system using a hybrid approach combining expert system and supervised deep learning". In: *Journal of Clinical Medicine* 10.15 (2021), p. 3347 (cit. on pp. 2, 30).

[Lareyre, 2023] Fabien Lareyre, Kak Khee Yeung, Lisa Guzzi, et al. "Artificial intelligence in vascular surgical decision making". In: *Seminars in Vascular Surgery*. Vol. 36. 3. Elsevier. 2023, pp. 448–453 (cit. on p. 7).

[Lareyre, 2025] Fabien Lareyre, Lisa Guzzi, Bahaa Nasr, et al. "Imaging characterization of peripheral artery disease: a scoping review on current classifications and new insights brought by Artificial Intelligence". In: *EJVES Vascular Forum*. Elsevier. 2025 (cit. on pp. 7, 20).

[LeCun, 1989] Yann LeCun, Bernhard Boser, John Denker, et al. "Handwritten digit recognition with a back-propagation network". In: *Advances in neural information processing systems* 2 (1989) (cit. on p. 24).

[LeCun, 1998] Yann LeCun and Yoshua Bengio. "Convolutional networks for images, speech, and time series". In: *The handbook of brain theory and neural networks* (1998) (cit. on p. 24).

[LeCun, 2002] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. "Gradient-based learning applied to document recognition". In: *Proceedings of the IEEE* 86.11 (2002), pp. 2278–2324 (cit. on p. 24).

[LeCun, 2015] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning". en. In: *Nature* 521.7553 (May 2015), pp. 436–444 (cit. on pp. 26, 43).

[Lee, 2019] David W Lee and Matthew A Cavender. *Guidelines for peripheral vascular disease: where is the evidence?* 2019 (cit. on p. 1).

[Lee, 2020] Juhwan Lee, Yazan Gharaibeh, Chaitanya Kolluru, et al. "Segmentation of coronary calcified plaque in intravascular OCT images using a two-step deep learning approach". In: *IEEE Access* 8 (2020), pp. 225581–225593 (cit. on p. 36).

[Li, 2024] Angran Li, Mingzhu Sun, and Zengshuo Wang. "TD Swin-UNet: Texture-Driven Swin-UNet with Enhanced Boundary-Wise Perception for Retinal Vessel Segmentation". In: *Bioengineering* 11.5 (2024), p. 488 (cit. on p. 33).

[Libby, 2021] Peter Libby. “The changing landscape of atherosclerosis”. In: *Nature* 592.7855 (2021), pp. 524–533 (cit. on p. 14).

[Lidayová, 2017] Kristína Lidayová, Hans Frimmel, Chunliang Wang, Ewert Bengtsson, and Orjan Smedby. “Chapter 12 - Skeleton-based fast, fully automated generation of vessel tree structure for clinical evaluation of blood vessel systems”. In: *Skeletonization*. Jan. 2017, pp. 345–382 (cit. on p. 43).

[Lim, 2025] Kyoung Yoon Lim, Jae Eun Ko, Yoo Na Hwang, Sang Goo Lee, and Sung Min Kim. “TransRAUNet: a deep neural network with reverse attention module using HU windowing augmentation for robust liver vessel segmentation in full resolution of CT images”. In: *Diagnostics* 15.2 (2025), p. 118 (cit. on p. 32).

[Liu, 2021] Ze Liu, Yutong Lin, Yue Cao, et al. “Swin transformer: Hierarchical vision transformer using shifted windows”. In: *Proceedings of the IEEE/CVF international conference on computer vision*. 2021, pp. 10012–10022 (cit. on p. 32).

[Livne, 2019] Michelle Livne, Jana Rieger, Orhun Utku Aydin, et al. “A U-Net deep learning framework for high performance vessel segmentation in patients with cerebrovascular disease”. In: *Frontiers in neuroscience* 13 (2019), p. 97 (cit. on pp. 30, 31).

[Ma, 2021] Jun Ma, Jianan Chen, Matthew Ng, et al. “Loss odyssey in medical image segmentation”. In: *Medical image analysis* 71 (2021), p. 102035 (cit. on p. 26).

[Macchi, 1996] Claudio Macchi, Fabrizio Giannelli, Francesca Cecchi, et al. “Collateral circulation in occlusion of lower limbs arteries: an anatomical study and statistical research in 35 old subjects.” In: *ITALIAN JOURNAL OF ANATOMY AND EMBRYOLOGY* 101 (1996), pp. 89–96 (cit. on p. 13).

[Masci, 2013] Jonathan Masci, Jesús Angulo, and Jürgen Schmidhuber. “A Learning Framework for Morphological Operators Using Counter–Harmonic Mean”. en. In: *Mathematical Morphology and Its Applications to Signal and Image Processing*. Berlin, Heidelberg, 2013, pp. 329–340 (cit. on pp. 45, 47, 123).

[Maurya, 2022] Akansh Maurya, Kunal Dashrath Patil, Rohan Padhy, Kalluri Ramakrishna, and Ganapathy Krishnamurthi. “Parse challenge 2022: Pulmonary arteries segmentation using swin u-net transformer (swin unetr) and u-net”. In: *arXiv preprint arXiv:2208.09636* (2022) (cit. on p. 33).

[McBane, 2024] Robert D McBane, Dennis H Murphree, David Liedl, et al. “Artificial intelligence of arterial Doppler waveforms to predict major adverse outcomes among patients evaluated for peripheral artery disease”. In: *Journal of the American Heart Association* 13.3 (2024), e031880 (cit. on p. 35).

[McDermott, 2015] Mary McGrae McDermott. “Lower extremity manifestations of peripheral artery disease: the pathophysiologic and functional implications of leg ischemia”. In: *Circulation research* 116.9 (2015), pp. 1540–1550 (cit. on p. 1).

[Mellouli, 2017] Dorra Mellouli, Tarek M. Hamdani, Mounir Ben Ayed, and Adel M. Alimi. “Morph-CNN: A Morphological Convolutional Neural Network for Image Classification”. en. In: *Neural Information Processing*. 2017, pp. 110–117 (cit. on pp. 45, 47).

[Mellouli, 2019] Dorra Mellouli, Tarek M. Hamdani, Javier J. Sanchez-Medina, Mounir Ben Ayed, and Adel M. Alimi. “Morphological Convolutional Neural Network Architecture for Digit Recognition”. In: *IEEE Transactions on Neural Networks and Learning Systems* 30.9 (2019), pp. 2876–2885 (cit. on pp. 45, 47).

[Menten, 2023] Martin J. Menten, Johannes C. Paetzold, Veronika A. Zimmer, et al. “A Skeletonization Algorithm for Gradient-Based Optimization”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Oct. 2023, pp. 21394–21403 (cit. on pp. 33, 43, 44, 46, 47, 59, 62, 65).

[Mills Sr, 2014] Joseph L Mills Sr, Michael S Conte, David G Armstrong, et al. “The society for vascular surgery lower extremity threatened limb classification system: risk stratification based on wound, ischemia, and foot infection (WIfI)”. In: *Journal of vascular surgery* 59.1 (2014), pp. 220–234 (cit. on p. 16).

[Mnih, 2013] Volodymyr Mnih. “Machine Learning for Aerial Image Labeling”. PhD thesis. University of Toronto, 2013 (cit. on p. 57).

[Moccia, 2018] Sara Moccia, Elena De Momi, Sara El Hadji, and Leonardo S Mattos. “Blood vessel segmentation algorithms—review of methods, datasets and evaluation metrics”. In: *Computer methods and programs in biomedicine* 158 (2018), pp. 71–91 (cit. on p. 30).

[Mohammadi, 2019] Saba Mohammadi, Mahdi Mohammadi, Vahab Dehlaghi, and Arash Ahmadi. “Automatic segmentation, detection, and diagnosis of abdominal aortic aneurysm (AAA) using convolutional neural networks and hough circles algorithm”. In: *Cardiovascular engineering and technology* 10.3 (2019), pp. 490–499 (cit. on p. 30).

[Mondal, 2019] Ranjan Mondal, Pulak Purkait, Sanchayan Santra, and Bhabatosh Chanda. “Morphological Networks for Image De-raining”. en. In: *Discrete Geometry for Computer Imagery*. 2019, pp. 262–275 (cit. on pp. 43, 46, 47).

[Mondal, 2020] Ranjan Mondal, Moni Shankar Dey, and Bhabatosh Chanda. “Image Restoration by Learning Morphological Opening-Closing Network”. en. In: *Mathematical Morphology - Theory and Applications* 4.1 (Jan. 2020), pp. 87–107 (cit. on pp. 47, 123).

[Mondal, 2022] Ranjan Mondal, Sanchayan Santra, Soumendu Sundar Mukherjee, and Bhabatosh Chanda. *Morphological Network: How Far Can We Go with Morphological Neurons?* en. arXiv:1901.00109 [cs, stat]. Dec. 2022 (cit. on pp. 47, 123).

[Moradi, 2019] Shakiba Moradi, Mostafa Ghelich Oghli, Azin Alizadehasl, et al. "MFP-Unet: A novel deep learning based approach for left ventricle segmentation in echocardiography". In: *Physica Medica* 67 (2019), pp. 58–69 (cit. on p. 72).

[Moriconi, 2018] Stefano Moriconi, Maria A Zuluaga, H Rolf Jäger, et al. "Inference of cerebrovascular topology with geodesic minimum spanning trees". In: *IEEE transactions on medical imaging* 38.1 (2018), pp. 225–239 (cit. on p. 124).

[Mou, 2021] Lei Mou, Yitian Zhao, Huazhu Fu, et al. "CS2-Net: Deep learning segmentation of curvilinear structures in medical imaging". In: *Medical image analysis* 67 (2021), p. 101874 (cit. on p. 31).

[Mukhopadhyay, 2000] Susanta Mukhopadhyay and Bhabatosh Chanda. "A multiscale morphological approach to local contrast enhancement". en. In: *Signal Processing* 80.4 (Apr. 2000), pp. 685–696 (cit. on p. 43).

[Müller, 2022] Dominik Müller, Iñaki Soto-Rey, and Frank Kramer. "Towards a guideline for evaluation metrics in medical image segmentation". In: *BMC Research Notes* 15.1 (2022), p. 210 (cit. on p. 27).

[Myronenko, 2018] Andriy Myronenko. "3D MRI brain tumor segmentation using autoencoder regularization". In: *International MICCAI brainlesion workshop*. Springer. 2018, pp. 311–320 (cit. on p. 31).

[Nader, 2023] Rafic Nader, Romain Bourcier, and Florent Autrusseau. "Using deep learning for an automatic detection and classification of the vascular bifurcations along the Circle of Willis". In: *Medical image analysis* 89 (2023), p. 102919 (cit. on p. 125).

[Nakashizuka, 2009] Makoto Nakashizuka. "A Design Method for Morphological Filters with Approximations of Min/Max Operators". In: *Proceedings : APSIPA ASC 2009 : Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference*. Oct. 2009, pp. 841–844 (cit. on pp. 43, 45, 123).

[Nardelli, 2018] Pietro Nardelli, Daniel Jimenez-Carretero, David Bermejo-Pelaez, et al. "Pulmonary artery–vein classification in CT images using deep learning". In: *IEEE transactions on medical imaging* 37.11 (2018), pp. 2428–2440 (cit. on p. 30).

[Navarro, 2019] Fernando Navarro, Suprosanna Shit, Ivan Ezhov, et al. "Shape-Aware Complementary-Task Learning for Multi-organ Segmentation". In: *Machine Learning in Medical Imaging*. 2019, pp. 620–627 (cit. on p. 73).

[Netter, 2022] Frank H Netter. *Netter Atlas of Human Anatomy: Classic Regional Approach-Ebook*. Elsevier health sciences, 2022 (cit. on p. 13).

[Nguyen, 2021] Nam Hoang Nguyen. "U-Net based skeletonization and bag of tricks". en. In: *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. Montreal, BC, Canada, Oct. 2021, pp. 2105–2109 (cit. on pp. 43, 46, 59).

[Ni, 2020] Jiajia Ni, Jianhuang Wu, Haoyu Wang, et al. “Global channel attention networks for intracranial vessel segmentation”. In: *Computers in biology and medicine* 118 (2020), p. 103639 (cit. on p. 31).

[Nogueira, 2021] Keiller Nogueira, Jocelyn Chanussot, Mauro Dalla Mura, and Jefersson A. Dos Santos. “An Introduction to Deep Morphological Networks”. In: *IEEE Access* 9 (2021), pp. 114308–114324 (cit. on p. 43).

[Nomura, 2005] Shigueo Nomura, Keiji Yamanaka, Osamu Katai, Hiroshi Kawakami, and Takayuki Shiose. “A novel adaptive morphological approach for degraded character image segmentation”. In: *Pattern Recognition* 38.11 (Nov. 2005), pp. 1961–1975 (cit. on p. 43).

[Nordanstig, 2024] Joakim Nordanstig, Christian-Alexander Behrendt, Iris Baumgartner, et al. “Editor’s choice–European Society for Vascular Surgery (ESVS) 2024 clinical practice guidelines on the management of asymptomatic lower limb peripheral arterial disease and intermittent claudication”. In: *European Journal of Vascular and Endovascular Surgery* 67.1 (2024), pp. 9–96 (cit. on pp. 1, 2, 13–20).

[Norgren, 2007] Lars Norgren, William R Hiatt, John A Dormandy, et al. “Inter-society consensus for the management of peripheral arterial disease (TASC II)”. In: *Journal of vascular surgery* 45.1 (2007), S5–S67 (cit. on p. 2).

[Novák, 1999] V. Novák, I. Perfilieva, and J. Mockor. *Mathematical Principles of Fuzzy Logic*. The Springer International Series in Engineering and Computer Science. 1999 (cit. on p. 53).

[Oh, 1998] Jinsung Oh and L.F. Chaparro. “Adaptive fuzzy morphological filtering of images”. In: *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP ’98 (Cat. No.98CH36181)*. Vol. 5. ISSN: 1520-6149. May 1998, 2901–2904 vol.5 (cit. on p. 43).

[Oktay, 2018] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, et al. “Attention u-net: Learning where to look for the pancreas”. In: *arXiv preprint arXiv:1804.03999* (2018) (cit. on p. 31).

[Ouyang, 2024] Xi Ouyang, Dongdong Gu, Xuejian Li, et al. “Towards a general computed tomography image segmentation model for anatomical structures and lesions”. In: *Communications Engineering* 3.1 (2024), p. 143 (cit. on pp. 2, 31).

[Padilla, 2020] Rafael Padilla, Sergio L Netto, and Eduardo AB Da Silva. “A survey on performance metrics for object-detection algorithms”. In: *2020 international conference on systems, signals and image processing (IWSSIP)*. IEEE. 2020, pp. 237–242 (cit. on p. 29).

[Paetzold, 2019] J. Paetzold, Oliver Schoppe, Rami Al-Maskari, et al. “Transfer learning from synthetic data reduces need for labels to segment brain vasculature and neural pathways in 3D”. In: *International Conference on Medical Imaging with Deep Learning–Extended Abstract Track*. Apr. 2019 (cit. on p. 57).

[Palàgyi, 1998] Kàlmàn Palàgyi and Attila Kuba. “A 3D 6-subiteration thinning algorithm for extracting medial lines”. In: *Pattern Recognition Letters* 19.7 (May 1998), pp. 613–627 (cit. on p. 56).

[Pan, 2021] Li-Syuan Pan, Chia-Wei Li, Shun-Feng Su, et al. “Coronary artery segmentation under class imbalance using a U-Net based architecture on computed tomography angiography images”. In: *Scientific Reports* 11.1 (2021), p. 14493 (cit. on p. 31).

[Panichev, 2019] Oleg Panichev and Alona Voloshyna. “U-Net Based Convolutional Neural Network for Skeleton Extraction”. en. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Long Beach, CA, USA: IEEE, June 2019, pp. 1186–1189 (cit. on pp. 43, 46).

[Pedregosa, 2011] F. Pedregosa, G. Varoquaux, A. Gramfort, et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830 (cit. on p. 92).

[Peters, 1995] R.A. Peters. “A new algorithm for image noise reduction using mathematical morphology”. In: *IEEE Transactions on Image Processing* 4.5 (May 1995). Conference Name: IEEE Transactions on Image Processing, pp. 554–568 (cit. on p. 43).

[Pham, 2000] Dzung L Pham, Chenyang Xu, and Jerry L Prince. “Current methods in medical image segmentation”. In: *Annual review of biomedical engineering* 2.1 (2000), pp. 315–337 (cit. on p. 24).

[Pham, 2021] Duc Duy Pham, Gurbandurdy Dovletov, and Josef Pauli. “A differentiable convolutional distance transform layer for improved image segmentation”. In: *Pattern Recognition: 42nd DAGM German Conference, DAGM GCPR 2020, Tübingen, Germany, September 28–October 1, 2020, Proceedings 42*. Springer. 2021, pp. 432–444 (cit. on pp. 73, 78, 79).

[Phellan, 2017] Renzo Phellan, Alan Peixinho, Alexandre Falcão, and Nils D Forkert. “Vascular segmentation in TOF MRA images of the brain using a deep convolutional neural network”. In: *International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*. Springer. 2017, pp. 39–46 (cit. on p. 30).

[Pihlak, 2021] René Pihlak and Andri Riid. “Morphological Cross Entropy Loss for Improved Semantic Segmentation of Small and Thin Objects”. In: *Procedia Computer Science* 192 (2021). Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 25th International Conference KES2021, pp. 582–591 (cit. on p. 48).

[Qazi, 2022] Emmad Qazi, Joerg Wilting, Neeral R Patel, et al. “Arteries of the lower limb—embryology, variations, and clinical significance”. In: *Canadian Association of Radiologists Journal* 73.1 (2022), pp. 259–270 (cit. on p. 13).

[Qiu, 2022] Jinming Qiu, Guanru Tan, Yan Lin, et al. “Automated detection of intracranial artery stenosis and occlusion in magnetic resonance angiography: a preliminary study based on deep learning”. In: *Magnetic Resonance Imaging* 94 (2022), pp. 105–111 (cit. on p. 127).

[Qiu, 2023] Yuehui Qiu, Zihan Li, Yining Wang, et al. “Corsegrec: a topology-preserving scheme for extracting fully-connected coronary arteries from ct angiography”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2023, pp. 670–680 (cit. on pp. 2, 34, 124).

[Radl, 2022] Lukas Radl, Yuan Jin, Antonio Pepe, et al. “AVT: Multicenter aortic vessel tree CTA dataset collection with ground truth segmentation masks”. In: *Data in brief* 40 (2022), p. 107801 (cit. on pp. 88, 89).

[Reitsma, 2017] Marissa B Reitsma, Nancy Fullman, Marie Ng, et al. “Smoking prevalence and attributable disease burden in 195 countries and territories, 1990–2015: a systematic analysis from the Global Burden of Disease Study 2015”. In: *The Lancet* 389.10082 (2017), pp. 1885–1906 (cit. on p. 15).

[Renard, 2020] Félix Renard, Soulaimane Guedria, Noel De Palma, and Nicolas Vuillerme. “Variability and reproducibility in deep learning for medical image segmentation”. In: *Scientific Reports* 10.1 (2020), p. 13724 (cit. on p. 94).

[Ribera, 2019] Javier Ribera, David Güera, Yuhao Chen, and Edward J. Delp. *Locating Objects Without Bounding Boxes*. 2019. arXiv: 1806.07564 [cs.CV] (cit. on p. 72).

[Riffaud, 2022] Sébastien Riffaud, Gwladys Ravon, Thibault Allard, et al. “Automatic branch detection of the arterial system from abdominal aortic segmentation”. In: *Medical & Biological Engineering & Computing* 60.9 (2022), pp. 2639–2654 (cit. on p. 126).

[Rizzetto, 2020] Francesco Rizzetto, Francesca Calderoni, Cristina De Mattia, et al. “Impact of inter-reader contouring variability on textural radiomics of colorectal liver metastases”. In: *European radiology experimental* 4 (2020), pp. 1–12 (cit. on p. 72).

[Robben, 2014] David Robben, Engin Türetken, Stefan Sunaert, et al. “Simultaneous segmentation and anatomical labeling of the cerebral vasculature”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2014, pp. 307–314 (cit. on p. 126).

[Robben, 2016] David Robben, Engin Türetken, Stefan Sunaert, et al. “Simultaneous segmentation and anatomical labeling of the cerebral vasculature”. In: *Medical image analysis* 32 (2016), pp. 201–215 (cit. on p. 126).

[Rogowska, 2009] Jadwiga Rogowska. “Overview and fundamentals of medical image segmentation”. In: *Handbook of medical image processing and analysis* (2009), pp. 73–90 (cit. on p. 24).

[Román, 2021] Julio César Mello Román, Vicente R. Fretes, Carlos G. Adorno, et al. “Panoramic Dental Radiography Image Enhancement Using Multiscale Mathematical Morphology”. en. In: *Sensors* 21.9 (Jan. 2021), p. 3110 (cit. on p. 43).

[Ronneberger, 2015] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-net: Convolutional networks for biomedical image segmentation”. In: *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer. 2015, pp. 234–241 (cit. on pp. 25, 26, 63, 79).

[Rosenfeld, 1968] Azriel Rosenfeld and John L Pfaltz. “Distance functions on digital pictures”. In: *Pattern recognition* 1.1 (1968), pp. 33–61 (cit. on p. 72).

[Roy, 2023] Saikat Roy, Gregor Koehler, Constantin Ulrich, et al. “Mednext: transformer-driven scaling of convnets for medical image segmentation”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2023, pp. 405–415 (cit. on p. 100).

[Rutherford, 1997] Robert B Rutherford, J Dennis Baker, Calvin Ernst, et al. “Recommended standards for reports dealing with lower extremity ischemia: revised version”. In: *Journal of vascular surgery* 26.3 (1997), pp. 517–538 (cit. on p. 19).

[Sabrowsky-Hirsch, 2021] Bertram Sabrowsky-Hirsch, Stefan Thumfart, Wolfgang Fenz, et al. “Automatic Segmentation of the Abdominal Aorta and Stent-Grafts”. In: *Journal of Image and Graphics* 9.3 (2021) (cit. on p. 36).

[Said, 2006] E.H. Said, D.E.M. Nassar, G. Fahmy, and H.H. Ammar. “Teeth segmentation in digitized dental X-ray films using mathematical morphology”. In: *IEEE Transactions on Information Forensics and Security* 1.2 (June 2006). Conference Name: IEEE Transactions on Information Forensics and Security, pp. 178–189 (cit. on p. 43).

[Salehi, 2017] Seyed Sadegh Mohseni Salehi, Deniz Erdogmus, and Ali Gholipour. “Tversky loss function for image segmentation using 3D fully convolutional deep networks”. In: *International workshop on machine learning in medical imaging*. Springer. 2017, pp. 379–387 (cit. on p. 27).

[Santoro, 2018] L Santoro, Andrea Flex, A Nesci, et al. “Association between peripheral arterial disease and cardiovascular risk factors: role of ultrasonography versus ankle-brachial index.” In: *European Review for Medical & Pharmacological Sciences* 22.10 (2018) (cit. on p. 14).

[Santos, 2025] Rui Santos, Rui Castro, Rúben Baeza, et al. “Segmentation of coronary calcifications with a domain knowledge-based lightweight 3D convolutional neural network”. In: *Computers in Biology and Medicine* 196 (2025), p. 110798 (cit. on p. 36).

[Schonfeld, 1991] D. Schonfeld and J. Goutsias. “Optimal morphological pattern restoration from noisy binary images”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13.1 (Jan. 1991). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 14–29 (cit. on p. 43).

[Serhal, 2018] Ali Serhal, Ioannis Koktzoglou, Pascale Aouad, et al. “Cardiovascular magnetic resonance imaging of aorto-iliac and ilio-femoral vascular calcifications using proton density-weighted in-phase stack of stars”. In: *Journal of Cardiovascular Magnetic Resonance* 20.1 (2018), p. 51 (cit. on p. 14).

[Sermesant, 2021] Maxime Sermesant, Herve Delingette, Hubert Cochet, Pierre Jais, and Nicholas Ayache. “Applications of artificial intelligence in cardiovascular imaging”. In: *Nature Reviews Cardiology* 18.8 (2021), pp. 600–609 (cit. on p. 2).

[Serra, 1994] Jean Serra, Pierre Soille, and Max A. Viergever. *Mathematical Morphology and Its Applications to Image Processing*. Vol. 2. Computational Imaging and Vision. 1994 (cit. on p. 42).

[Sharma, 2010] Neeraj Sharma and Lalit M Aggarwal. “Automated medical image segmentation techniques”. In: *Journal of medical physics* 35.1 (2010), pp. 3–14 (cit. on p. 24).

[Shelhamer, 2017] Evan Shelhamer, Jonathan Long, and Trevor Darrell. “Fully convolutional networks for semantic segmentation”. In: *IEEE transactions on pattern analysis and machine intelligence* 39.4 (2017), pp. 640–651 (cit. on p. 30).

[Shen, 2017] Dinggang Shen, Guorong Wu, and Heung-Il Suk. “Deep learning in medical image analysis”. In: *Annual review of biomedical engineering* 19.1 (2017), pp. 221–248 (cit. on p. 24).

[Shen, 2022] Yucong Shen, Frank Y. Shih, Xin Zhong, and I-Cheng Chang. “Deep Morphological Neural Networks”. en. In: *International Journal of Pattern Recognition and Artificial Intelligence* 36.12 (Sept. 2022), p. 2252023 (cit. on pp. 45, 47, 123).

[Shi, 2024] Pengcheng Shi, Jiesi Hu, Yanwu Yang, et al. “Centerline boundary dice loss for vascular segmentation”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2024, pp. 46–56 (cit. on pp. 33, 101, 102).

[Shih, 2017] Frank Y. Shih. *Image Processing and Mathematical Morphology: Fundamentals and Applications*. Boca Raton: CRC Press, Jan. 2017 (cit. on pp. 42, 55).

[Shih, 2019] Frank Y. Shih, Yucong Shen, and Xin Zhong. “Development of Deep Learning Framework for Mathematical Morphology”. In: *International Journal of Pattern Recognition and Artificial Intelligence* 33.06 (June 2019), p. 1954024 (cit. on pp. 45, 123).

[Shin, 2019] Seung Yeon Shin, Soochahn Lee, Il Dong Yun, and Kyoung Mu Lee. “Deep vessel segmentation by learning graphical connectivity”. In: *Medical image analysis* 58 (2019), p. 101556 (cit. on p. 125).

[Shit, 2021] Suprosanna Shit, Johannes C. Paetzold, Anjany Sekuboyina, et al. "clDice - a Novel Topology-Preserving Loss Function for Tubular Structure Segmentation". English. In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2021, pp. 16555–16564 (cit. on pp. 2, 29, 33, 43, 46, 48, 57, 59, 64, 78, 101, 103, 124).

[Siddique, 2020] Nahian Siddique, Paheding Sidike, Colin Elkin, and Vijay Devabhaktuni. "U-Net and its variants for medical image segmentation: theory and applications". In: *arXiv preprint arXiv:2011.01118* (2020) (cit. on p. 25).

[Sigvant, 2007] Birgitta Sigvant, Katarina Wiberg-Hedman, David Bergqvist, et al. "A population-based study of peripheral arterial disease prevalence with special focus on critical limb ischemia and sex differences". In: *Journal of vascular surgery* 45.6 (2007), pp. 1185–1191 (cit. on p. 15).

[Sinha, 1992] Divyendu Sinha and Edward R. Dougherty. "Fuzzy mathematical morphology". In: *Journal of Visual Communication and Image Representation* 3.3 (Sept. 1992), pp. 286–302 (cit. on pp. 42, 45).

[Song, 2019] Peige Song, Diana Rudan, Yajie Zhu, et al. "Global, regional, and national prevalence and risk factors for peripheral artery disease in 2015: an updated systematic review and analysis". In: *The Lancet Global Health* 7.8 (2019), e1020–e1030 (cit. on pp. 1, 15).

[Song, 2022] Along Song, Lisheng Xu, Lu Wang, et al. "Automatic coronary artery segmentation of CCTA images with an efficient feature-fusion-and-rectification 3D-UNet". In: *IEEE journal of biomedical and health informatics* 26.8 (2022), pp. 4044–4055 (cit. on p. 30).

[Staal, 2004] J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, and B. Van Ginneken. "Ridge-Based Vessel Segmentation in Color Images of the Retina". en. In: *IEEE Transactions on Medical Imaging* 23.4 (Apr. 2004), pp. 501–509 (cit. on pp. 57, 78).

[Stringhini, 2019] Rômulo Marconato Stringhini, Daniel Welfer, Marcos Cordeiro d'Ornellas, and Daniel Fernando Tello Gamarra. "A Mathematical Morphology-Based Filter for Noise Reduction and Detail Preservation in Low-Dose Dental CT Images". eng. In: *Studies in Health Technology and Informatics* 264 (Aug. 2019), pp. 253–257 (cit. on p. 43).

[Stucki, 2023] Nico Stucki, Johannes C Paetzold, Suprosanna Shit, Bjoern Menze, and Ulrich Bauer. "Topologically faithful image segmentation via induced matching of persistence barcodes". In: *International Conference on Machine Learning*. PMLR. 2023, pp. 32698–32727 (cit. on pp. 2, 33).

[Taha, 2015] Abdel Aziz Taha and Allan Hanbury. "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool". In: *BMC medical imaging* 15 (2015), pp. 1–28 (cit. on pp. 27, 72).

[Taha, 2018] Ahmed Taha, Pechin Lo, Junning Li, and Tao Zhao. “Kid-net: convolution networks for kidney vessels segmentation from ct-volumes”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2018, pp. 463–471 (cit. on p. 30).

[Takahashi, 1984] Mutsumasa Takahashi, Yukinori Koga, Hiromasa Bussaka, and Masayuki Miyawaki. “The value of digital subtraction angiography in peripheral vascular diseases”. In: *The British Journal of Radiology* 57.674 (1984), pp. 123–132 (cit. on p. 17).

[Tan, 2021] Wenjun Tan, Luyu Zhou, Xiaoshuo Li, et al. “Automated vessel segmentation in lung CT and CTA images via deep neural networks”. In: *Journal of X-ray science and technology* 29.6 (2021), pp. 1123–1137 (cit. on p. 30).

[Tetteh, 2020] Giles Tetteh, Velizar Efremov, Nils D Forkert, et al. “Deepvesselnet: Vessel segmentation, centerline prediction, and bifurcation detection in 3-d angiographic volumes”. In: *Frontiers in Neuroscience* 14 (2020), p. 592352 (cit. on p. 31).

[Tukey, 1977] John Wilder Tukey et al. *Exploratory data analysis*. Vol. 2. Springer, 1977 (cit. on p. 94).

[Van der Walt, 2014] Stefan Van der Walt, Johannes L Schönberger, Juan Nunez-Iglesias, et al. “scikit-image: image processing in Python”. In: *PeerJ* 2 (2014), e453 (cit. on p. 59).

[Wagner, 2020] Martin G. Wagner. “Real-Time Thinning Algorithms for 2D and 3D Images using GPU processors”. In: *Journal of real-time image processing* 17.5 (Oct. 2020), pp. 1255–1266 (cit. on pp. 56, 129).

[Wang, 2009] Josiah Wang, Katja Markert, and Mark Everingham. “Learning Models for Object Recognition from Natural Language Descriptions”. In: *Proceedings of the British Machine Vision Conference*. 2009 (cit. on p. 57).

[Wang, 2020] Yan Wang, Xu Wei, Fengze Liu, et al. “Deep distance transform for tubular structure segmentation in ct scans”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020, pp. 3833–3842 (cit. on p. 73).

[Wasserthal, 2023] Jakob Wasserthal, Hanns-Christian Breit, Manfred T Meyer, et al. “TotalSegmentator: robust segmentation of 104 anatomic structures in CT images”. In: *Radiology: Artificial Intelligence* 5.5 (2023), e230024 (cit. on pp. 100, 102).

[Weng, 2023] Ziqiao Weng, Jiancheng Yang, Dongnan Liu, and Weidong Cai. “Topology repairing of disconnected pulmonary airways and vessels: baselines and a dataset”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2023, pp. 382–392 (cit. on p. 2).

[Wijnand, 2021] Joep GJ Wijnand, Devin Zarkowsky, Bian Wu, et al. “The global limb anatomic staging system (GLASS) for CLTI: improving inter-observer agreement”. In: *Journal of Clinical Medicine* 10.16 (2021), p. 3454 (cit. on p. 2).

[Wittmann, 2025] Bastian Wittmann, Yannick Wattenberg, Tamaz Amiranashvili, Suprosanna Shit, and Bjoern Menze. "vesselFM: A Foundation Model for Universal 3D Blood Vessel Segmentation". In: *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR)*. June 2025, pp. 20874–20884 (cit. on pp. 100, 102).

[Wolterink, 2019] Jelmer M Wolterink, Tim Leiner, and Ivana Išgum. "Graph convolutional networks for coronary artery segmentation in cardiac CT angiography". In: *International Workshop on Graph Learning in Medical Imaging*. Springer. 2019, pp. 62–69 (cit. on p. 30).

[Wright, 2022] Fred Wynn Wright. *Radiology of the chest and related conditions*. CRC Press, 2022 (cit. on p. 97).

[Xia, 2006] Yong Xia, Dagan Feng, and Rongchun Zhao. "Morphology-based multifractal estimation for texture segmentation". In: *IEEE Transactions on Image Processing* 15.3 (Mar. 2006). Conference Name: IEEE Transactions on Image Processing, pp. 614–623 (cit. on p. 43).

[Xiao, 2018] Xiao Xiao, Shen Lian, Zhiming Luo, and Shaozi Li. "Weighted res-unet for high-quality retina vessel segmentation". In: *2018 9th international conference on information technology in medicine and education (ITME)*. IEEE. 2018, pp. 327–331 (cit. on p. 30).

[Xie, 2020] Meiyan Xie, Yunzhi Li, Yunzhe Xue, et al. "Vessel lumen segmentation in carotid artery ultrasounds with the U-Net convolutional neural network". In: *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE. 2020, pp. 2680–2684 (cit. on p. 31).

[Xue, 2020] Yuan Xue, Hui Tang, Zhi Qiao, et al. "Shape-aware organ segmentation by predicting signed distance maps". In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 34. 2020, pp. 12565–12572 (cit. on p. 73).

[Yang, 2019] Su Yang, Jihoon Kweon, and Young-Hak Kim. "Major vessel segmentation on x-ray coronary angiography using deep networks with a novel penalty loss function". In: (2019) (cit. on p. 72).

[Yang, 2024] Kaiyuan Yang, Fabio Musio, Yihui Ma, et al. "Benchmarking the cow with the topcow challenge: Topology-aware anatomical segmentation of the circle of willis for cta and mra". In: *ArXiv* (2024), arXiv–2312 (cit. on pp. 34, 125).

[yong, 2010] Huang yong, Tao yin, and Liu huijuan. "Noise image restoration based on mathematical morphology". In: *The 2nd International Conference on Information Science and Engineering*. ISSN: 2160-1291. Dec. 2010, pp. 3840–3843 (cit. on p. 43).

[Young, 2019] Jessica C Young, Nicole Jadue Paul, Turkan Banu Karatas, et al. "Cigarette smoking intensity informs outcomes after open revascularization for peripheral artery disease". In: *Journal of Vascular Surgery* 70.6 (2019), pp. 1973–1983 (cit. on p. 15).

[Yu, 2019] Wei Yu, Bin Fang, Yongqing Liu, et al. "Liver vessels segmentation based on 3d residual U-NET". In: *2019 IEEE international conference on image processing (ICIP)*. IEEE. 2019, pp. 250–254 (cit. on p. 31).

[Yucheng, 2009] Liu Yucheng and Liu Yubin. “An Algorithm of Image Segmentation Based on Fuzzy Mathematical Morphology”. In: *2009 International Forum on Information Technology and Applications*. Vol. 2. May 2009, pp. 517–520 (cit. on p. 43).

[Yushkevich, 2016] Paul A Yushkevich, Yang Gao, and Guido Gerig. “ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images”. In: *2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE. 2016, pp. 3342–3345 (cit. on p. 91).

[Zhang, 2020] Dong Zhang, Guang Yang, Shu Zhao, et al. “Direct quantification of coronary artery stenosis through hierarchical attentive multi-view learning”. In: *IEEE transactions on medical imaging* 39.12 (2020), pp. 4322–4334 (cit. on p. 127).

[Zhang, 2022] Minghui Zhang, Guang-Zhong Yang, and Yun Gu. “Differentiable topology-preserved distance transform for pulmonary airway segmentation”. In: *arXiv preprint* (2022) (cit. on p. 73).

[Zhang, 2023] Zhixing Zhang, Ziwei Zhao, Dong Wang, et al. “Topology-preserving automatic labeling of coronary arteries via anatomy-aware connection classifier”. In: *International conference on medical image computing and computer-assisted intervention*. Springer. 2023, pp. 759–769 (cit. on pp. 2, 34).

[Zhang, 2024] Yishuo Zhang and Albert CS Chung. “Retinal vessel segmentation by a transformer-u-net hybrid model with dual-path decoder”. In: *IEEE Journal of Biomedical and Health Informatics* 28.9 (2024), pp. 5347–5359 (cit. on p. 32).

[Zhao Yu-qian, 2005] Zhao Yu-qian, Gui Wei-hua, Chen Zhen-cheng, Tang Jing-tian, and Li Ling-yun. “Medical Images Edge Detection Based on Mathematical Morphology”. en. In: *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*. Shanghai, China, 2005, pp. 6492–6495 (cit. on p. 43).

[Zhao, 2015] Binsheng Zhao, Lawrence H Schwartz, Mark G Kris, and Gregory J Riely. “Coffee-break lung ct collection with scan images reconstructed at multiple imaging parameters”. In: *(No Title)* (2015) (cit. on p. 89).

[Zhou, 2025] Feixiang Zhou, Zhuangzhi Gao, He Zhao, et al. “GLCP: Global-to-Local Connectivity Preservation for Tubular Structure Segmentation”. In: *arXiv preprint arXiv:2507.21328* (2025) (cit. on p. 34).

[Zhu, 2022] Ying Zhu, Liwei Chen, Wenjie Lu, Yongjun Gong, and Ximing Wang. “The application of the nnU-Net-based automatic segmentation model in assisting carotid artery stenosis and carotid atherosclerotic plaque evaluation”. In: *Frontiers in physiology* 13 (2022), p. 1057800 (cit. on pp. 2, 31).

[Zhu, 2024] Ruiyun Zhu, Masahiro Oda, Yuichiro Hayashi, Takayuki Kitasaka, and Kensaku Mori. “Semi-supervised Tubular Structure Segmentation with Cross Geometry and Hausdorff Distance Consistency”. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2024*. Cham: Springer Nature Switzerland, 2024, pp. 612–622 (cit. on pp. 73, 78, 79).

[Zou, 2025] Mingyu Zou, Shuxiang Guo, Bingzhi Shen, et al. “Segmentation of Calcified Lesions in Lower Limb Artery CTA”. In: *2025 IEEE International Conference on Mechatronics and Automation (ICMA)*. IEEE. 2025, pp. 874–879 (cit. on p. 36).

[Zulfiqar, 2026] Manahil Zulfiqar, Maciej Stanuch, Sylvia Vagena, Fragiska Sigala, and Andrzej Skalski. “Automated segmentation of peripheral arteries in 3D CT data based on a Single-Depth vascular network”. In: *Biomedical Signal Processing and Control* 112 (2026), p. 108410 (cit. on pp. 35, 88, 89).