

École polytechnique de Louvain

Using Graph Cuts for Anatomical Priors to Segment Pulmonary Embolus

Label Generation for Deep Model Training

Author: **Robin VERSCHUREN**

Supervisors: **Christophe DE VLEESCHOUWER, Benoît MACQ**

Readers: **Victor JOOS DE TER BEERST, Dany RIMEZ**

Academic year 2022–2023

Master [120] in Mathematical Engineering

ABSTRACT

This thesis presents a novel Pulmonary Embolus (PE) detection and segmentation approach by combining classical computer vision and deep learning techniques. Pulmonary Embolism is a life-threatening medical condition characterized by the obstruction of pulmonary arteries, often tricky and time-consuming to diagnose accurately. The study addresses the possibility of generating segmentation of PE candidates as labels for a weakly supervised model. We also explore the potential enhancement in performance by incorporating prior anatomical segmentation in the chest. Implementing a classical algorithm for PE segmentation, our methodology involves generating superpixels and employing graph cuts for prior anatomical segmentation, all within the broader context of generating labels for training a weakly supervised learning framework. Our findings suggest that the presented methodology, while promising, still needs refinement to reduce the number of false positives and improve the segmentation of PE regions. This work forms a foundation for further research to improve PE diagnosis and treatment by enhancing the efficacy of deep learning models without the need for extensively human-annotated CT scan datasets.

ACKNOWLEDGEMENTS

I want to take this opportunity to express my most profound appreciation to my thesis advisors and supervisors, Prof. Benoît Macq and Prof. Christophe De Vleeschouwer, for their support, guidance, and mentorship throughout this research process. Their expertise has been instrumental in shaping this work.

I am deeply grateful to Victor Joos de ter Beerst for serving as my reader and jury member.

I want to acknowledge Nana Yang for her interest in my thesis and participation during meetings.

Special thanks are extended to Dany Rimez, whose dedicated assistance was vital to my research. Despite his own commitments as a Ph.D. student, he was always available to answer my questions, attend meetings, and provide insightful feedback on my thesis. I am immensely grateful for his dedication and support.

Special thanks to my father, Franck Verschuren, not only for his familial support but also for his professional guidance. As a teacher at the Faculty of Medicine of UCLouvain, he has been a critical resource for addressing medical-related questions and providing crucial feedback. His emotional and academic support has been a vital part of this journey.

I express my profound gratitude to my family, Isabelle De Visschere, Alice Verschuren, Antoine Verschuren, Victor Verschuren, and my friend Juliette de Crayencour. Their faith in my abilities and encouragement have provided the necessary motivation and reassurance during challenging times.

I extend my appreciation for artificial intelligence tools that assisted in this thesis. Specifically, ChatGPT from OpenAI aided using a scientific writing style, $\LaTeX$ equations, and Python DocString generation. Also, Grammarly was crucial for sentence refinements and grammar checks.

CONTENTS

1	INTRODUCTION	1
2	BACKGROUND AND CONTEXT	3
2.1	Anatomy and Physiology of Pulmonary Embolism	3
2.1.1	Classification of Pulmonary Embolism	3
2.2	Computed Tomography Angiography (CTA)	4
2.2.1	Scanners	5
2.2.2	Contrast Agent	5
2.2.3	Spectral Scanners	5
2.3	Specific Challenges in Medical Field	6
2.3.1	Limited Availability and Accessibility	6
2.3.2	Interpretability and Explainability for Critical Clinical Decisions . . .	7
2.3.3	Generalizability	8
2.3.4	Ethical Concerns	8
2.4	CAD - An Overview of Tasks	8
2.4.1	Classification	8
2.4.2	Detection	9
2.4.3	Segmentation	9
2.5	Classical Segmentation Algorithms	10
2.5.1	Pulmonary Arteries Segmentation	11
2.5.2	Graph Cuts with alpha-Expansions	13
2.6	Deep Learning Methods	13
2.6.1	Weakly Supervised Models	15
3	RESEARCH QUESTIONS AND HYPOTHESES	19
4	MATERIALS AND METHODOLOGY	21
4.1	Materials	21
4.1.1	Dataset	21
4.1.2	Tools	21

4.1.3	Hardware	22
4.2	Enlarged Overview of the Framework	22
4.2.1	Two-Stage Framework and Algorithm Overview	23
4.3	Stage 1: Methodology and Intermediate Results	25
4.3.1	Preprocessing	25
4.3.2	Superpixels Generation	27
4.3.3	Energy Minimization and Graph Cuts	30
4.3.4	Data Fidelity	32
4.3.5	Graph Construction	42
4.3.6	Alpha-expansion	45
4.3.7	PE Localization Using Prior-Anatomical-Segmentation	50
4.4	Evaluation Metrics	52
5	RESULTS	55
5.1	Quantitative Results	55
5.1.1	Detection	55
5.1.2	Segmentation	57
5.2	Visual Analysis of Pulmonary Embolus Delineation	59
6	DISCUSSIONS	65
6.1	Limitations and Future Improvement	68
6.2	Additional Research Question	69
7	CONCLUSION	71
	ACRONYMS	73
	GLOSSARY	75
	BIBLIOGRAPHY	77

1 INTRODUCTION

Pulmonary Embolism represents a critical medical challenge - a life-threatening vascular condition marked by the obstruction of pulmonary arteries due to blood clots (i.e., Embolus, PE), commonly originating from the leg veins. This medical emergency imposes a significant health threat, with an annual incidence rate of one in 1,000 individuals in the United States [18]. Obstruction caused by these clots impairs blood and oxygen flow to vital organs, potentially leading to fatal consequences if not treated promptly and accurately.

Diagnosing Pulmonary Embolism, a task entrusted primarily to radiologists, is often a complex and time-consuming process [29], where rapid identification and accurate interpretation are the pillars of effective treatment. Given the complex nature of Pulmonary Embolism diagnosis and the criticality of time in the process, Computer-Aided Diagnosis (CAD) software has been gaining traction as a potential solution. The advent of CAD could offer an increase in radiologists' self-confidence, reducing the chances of misdiagnosing Pulmonary Embolism - a not uncommon issue due to overlapping symptoms with other medical conditions [23].

While classical algorithm-based CAD designs have been studied, the recent advancements in deep learning methodologies have demonstrated promising and robust performance outcomes [9]. Yet, the efficiency of these methodologies is often linked to the availability of annotated scans, a resource typically scarce and costly in the medical field. Some frameworks mitigate this issue by utilizing inaccurate labels generated by classical computer vision algorithms. However, these algorithms are neither readily available nor reusable.

This thesis aims to address this research gap by proposing an innovative solution: generating pixel-wise labels using classical algorithms for training weakly supervised models to segment PEs. Two primary research questions guide this study: Can we generate segmented PE candidates that can be utilized as labels for a weakly supervised model? Secondly, can we enhance performance by incorporating prior anatomical segmentation in the chest?

Our methodological approach aims to develop a classical algorithm for PE segmentation that could serve as inaccurate labels. Our proposed technique involves generating superpixels, employ-

ing graph cuts for anatomical segmentation as prior knowledge, and prospecting for PE candidates in the pulmonary arteries, all within the broader framework of weakly supervised learning.

The potential implications of this research are profound. By leveraging labels generated from classical algorithms, we could enhance the efficacy of deep learning models without requiring extensive, fully human-annotated CT scan datasets, thereby promoting more efficient and cost-effective diagnosis of PEs.

The structure of this thesis unfolds as follows: Chapter 2 delves into existing deep learning frameworks and classical algorithms. Chapter 3 presents our research questions. Chapter 4 delineates our methodology, detailing the framework description and label generation using superpixels, accompanied by anatomical segmentation with graph cuts. The ensuing chapter shares our findings, and the final chapter engages in discussions, drawing comparisons with existing methods.

2 BACKGROUND AND CONTEXT

This section provides a comprehensive review of several important topics related to the subject matter of this thesis. We begin by exploring the anatomy and physiology of Pulmonary Embolism, then by examining the diagnostic techniques used in its detection, explicitly focusing on Computed Tomography Angiograms (CTA). Additionally, we delve into the challenges encountered when applying machine learning algorithms to medical images and discuss the growing interest in utilizing deep learning methods for computer-aided detection in this domain.

2.1 ANATOMY AND PHYSIOLOGY OF PULMONARY EMBOLISM

Pulmonary embolism and deep venous thrombosis are interrelated stages of the same disease. Thrombi, Embolus, or blood clots [4] typically form in the deep veins of the calf (i.e., the back of your lower leg, behind your shin bone) and can propagate into the proximal veins, from where they are more likely to embolize. It has been observed that about 79% of patients diagnosed with PE have evidence of thrombosis in their legs. It is likely that the thrombus has already left the legs and embolized in the remaining 31% of patients [50]. This displacement of thrombosis is graphically depicted in fig. 2.1. In cases of acute pulmonary embolism, anatomical obstruction is undoubtedly the primary cause of problems, although the release of vasoactive and bronchoactive agents can also impair ventilation-perfusion matching. For this thesis, we will exclusively focus on anatomical obstructions.

2.1.1 CLASSIFICATION OF PULMONARY EMBOLISM

There are three types of pulmonary embolism, acute, subacute, and chronic, which depend on the temporal pattern of the disease [53].

Acute PE is characterized by a sudden obstruction of an artery due to a blood clot. It is considered a medical emergency that requires immediate treatment. **Subacute** PE is a less severe condition that develops over several days or weeks. **Chronic** PE refers to cases where the disease develops over several months or years. This long-term condition can still cause permanent damage to the lungs.

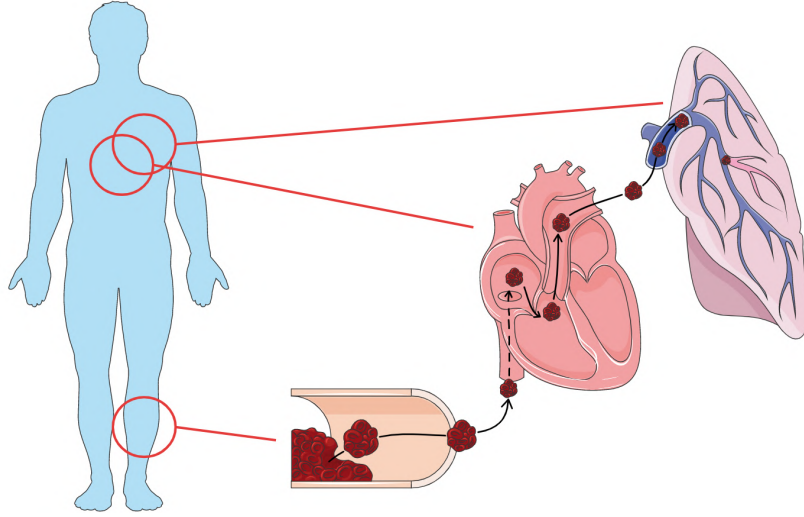


Figure 2.1: Pathophysiology of Thrombosis and Pulmonary Embolism. Parts of the figure were drawn using pictures from Servier Medical Art. Servier Medical Art by Servier is licensed under a Creative Commons Attribution 3.0 Unported License (<https://creativecommons.org/licenses/by/3.0/>).

In this thesis, we will not differentiate between the different types of PE. Instead, our focus will be on segmenting any presence of PE, regardless of its specific type.

2.2 COMPUTED TOMOGRAPHY ANGIOGRAPHY (CTA)

The present thesis does not aim to provide an in-depth analysis of the selection criteria for the various medical techniques employed for disease diagnosis, given the complexity of this topic that requires the knowledge of healthcare professionals and researchers. However, among the different techniques used for pulmonary embolism diagnosis, some notable ones are: The **Pulmonary angiogram**, a catheterization-based method, which used to be the gold standard for diagnosing PE, but was later replaced by CT scans [33]; the **D-dimer** test, which involves blood analysis and is utilized to exclude PE on patients obtaining a negative result [22]; the **Ventilation-perfusion (V/Q) scan**, a nuclear medicine imaging test that uses a small amount of radioactive material to determine airflow and blood flow in the lungs [49]; and the **Capnography** method, which is portable, non-invasive, and cost-effective, relies on exhaled CO₂ measurements [55]. However, the most effective technique, and thus, considered a gold method for diagnosing acute and chronic PE, is the **Computed Tomography Angiography (CTA)** scan. It is worth noting that this scan is typically only performed on patients with an intermediate or high probability of having PE. Thus, for confirming the suspicion of pulmonary embolism [22]. Furthermore, CTA seems to be the

preferred imaging modality for acquiring data that serves as input for CAD (computer-aided detection) machine learning methods.

In the subsequent sections, we will examine the different types of scanners utilized to acquire these images and the significance of contrast agents in CTA scans.

2.2.1 SCANNERS

The identification of the generation of the scanner is a crucial factor to consider, as its capabilities vary depending on the technological advancements implemented in each generation. In particular, the detection of embolus may be affected by the resolution and sensitivity of the scanner. For instance, older first-generation scanners, with a 5mm resolution, may miss up to one-third of embolus in contrast to newer third-generation scanners, with a 1mm resolution [15]. Therefore, selecting the appropriate scanner generation should be carefully considered to ensure accurate diagnosis and treatment.

2.2.2 CONTRAST AGENT

To obtain conclusive images using helical CT (computed tomography), radiologists resort to pulmonary angiography, also known as CTA scan, which is considered the gold standard for confirming a suspicion of pulmonary embolism [22]. During the CTA scan, an intravenous contrast agent containing iodine is injected into the patient, followed by scanning after a specified delay to optimize contrast in the lungs. Due to the contrast agent's mix with the blood, the Hounsfield Units (HU) values of the blood change, whereas blood clots remain in a solid state and are thus not impacted by the contrast agent. The mixed blood appears bright on the scan due to X-rays reactions with iodine, whereas the embolus appears darker. The embolus's detectability improvements are shown in fig. 2.2.

2.2.3 SPECTRAL SCANNERS

Spectral CT scanners are a recent advancement in CTA technology, such as the one utilized at Cliniques Universitaires Saint-Luc. These scanners capture multi-spectral images of patients, which provide additional information to improve image quality without increasing radiation exposure [43]. This method is beneficial in detecting pulmonary embolus, where traditional CTA scans may not provide clear visualization of the embolus. An example of this improvement is illustrated in the figure below, where the spectral image allows to represent the pulmonary perfusion with the distribution of the contrast medium in color scale. The purple triangular area at the right lung base (arrow) highlights a perfusion defect. Rereading the CT scan in spectral mode allows us to rectify this patient's first diagnosis and identify a peripheral pulmonary embolus (arrow).

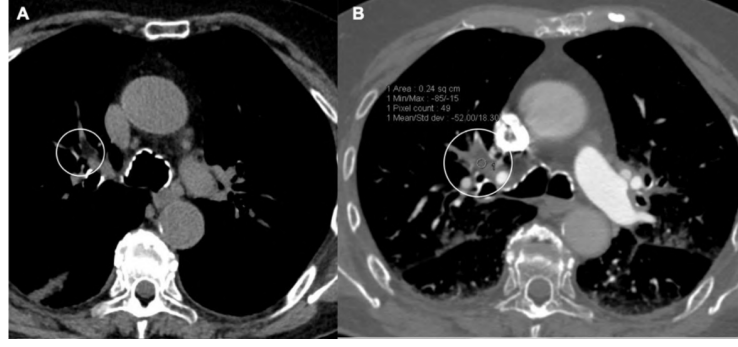


Figure 2.2: Caption. Image cropped from Zantonelli, G et al. in *Acute Pulmonary Embolism: Prognostic Role of Computed Tomography Pulmonary Angiography (CTPA)*[61] and is licensed under a Creative Commons Attribution 4.0 (CC BY) License (<https://creativecommons.org/licenses/by/4.0/>).

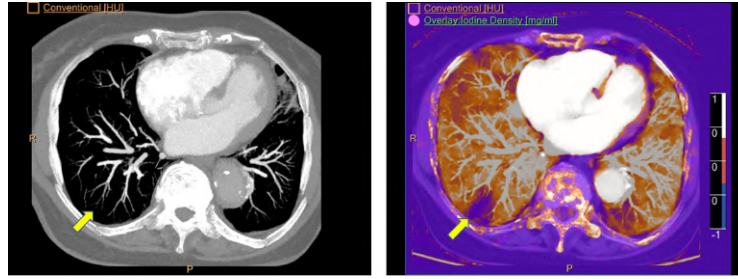


Figure 2.3: Comparing normal CT acquisition (left) with spectral-CTA acquisition (right) - Images are from the case study in Philips press communication [5]

2.3 SPECIFIC CHALLENGES IN MEDICAL FIELD

2.3.1 LIMITED AVAILABILITY AND ACCESSIBILITY

In the field of machine learning methods applied to medical data, obtaining large, diverse, and high-quality medical datasets can be challenging due to privacy concerns, legal regulations, and data-sharing policies. This is known as the limited availability and accessibility of high-quality medical data [3]. Most studies encountered in the literature are developed and validated using private datasets [31]. To address legal regulations and data-sharing policies, novel approaches for developing improved models have emerged. One such approach, as illustrated in fig. 2.4, is federated learning [60], which operates on the principle of bringing the model to the data rather than the converse. This method allows multiple hospitals to work collaboratively toward improving a common model while ensuring that sensitive data remains secure and unavailable to unautho-

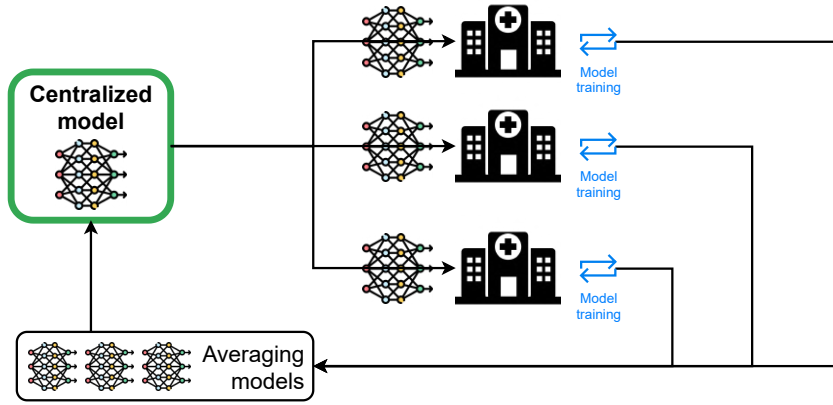


Figure 2.4: Secure Federated Learning in the Medical Application

rized parties. This strategy has the potential to offer significant benefits over classical methods of model development and training.

2.3.2 INTERPRETABILITY AND EXPLAINABILITY FOR CRITICAL CLINICAL DECISIONS

A challenge of relying on machine learning methods for medical decision-making is that the radiologist has to double-check the predicted output and see whether it is trustworthy. This can be a difficult task as deep learning methods can be seen as black boxes where intermediate results are difficult to interpret and explain. Ma *et al.* introduced a multi-task deep learning framework in

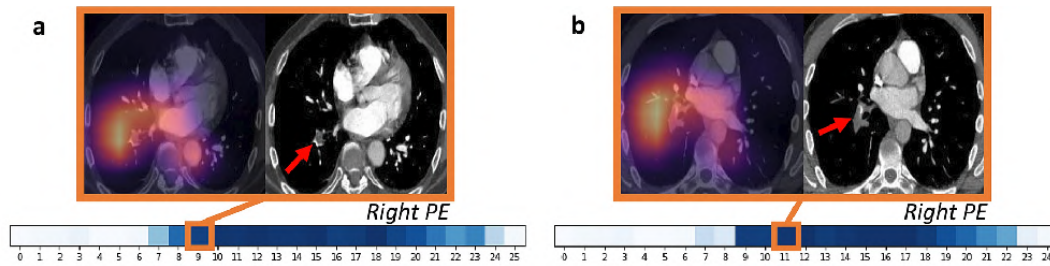


Figure 2.5: Explainability of two true positive PE scans using Grad-CAM. PE is shown with the red arrow. - from Ma *et al.* in *A multitask deep learning approach for pulmonary embolism detection and identification* [29] and is licensed under a Creative Commons Attribution 4.0 (CC BY) License (<https://creativecommons.org/licenses/by/4.0/>).

[29] for the detection of pulmonary embolus (PE) and employed Grad-CAM (Gradient-weighted Class Activation Mapping) as an explainable method to visualize the regions of the CT scan slices that influenced the final prediction (refer to fig. 2.5). They also utilized the same technique to explain the results in her thesis [52]. Grad-CAM provides a heatmap to highlight the regions that

contribute the most to the model’s prediction, thereby improving the model’s interpretability and facilitating the validation by the radiologist.

2.3.3 GENERALIZABILITY

Deep learning models are often trained on specific medical datasets, which may not generalize well to new data from different sources or populations. As a result, extensive testing and validation are necessary to ensure the model’s generalizability and robustness. It is crucial to carefully read the specifications and descriptions when using publicly available datasets to achieve this. Relevant questions include the population characteristics, the number of different machines used, the number of contributing hospitals, and the preprocessing steps employed, among others. Such details are essential for assessing the applicability and reliability of the model on new data.

2.3.4 ETHICAL CONCERNS

Using deep learning methods on medical data raises ethical concerns regarding privacy, security, and bias, as this data is sensitive and personal. Implementing proper safeguards and regulations is crucial to ensure the ethical use of medical data. It is often mentioned in the papers presenting new public datasets what were the ethical concerns. For instance, in the CAD-PE dataset [16], they note that all patients’ health information was removed to protect their privacy. In the FUMPE dataset [31], they mention that patients had to sign a document agreeing to share their anonymized scans. The two datasets mentioned above are CTA scans for applying PE detections.

2.4 CAD - AN OVERVIEW OF TASKS

Computer-aided diagnosis (CAD) for pulmonary embolism involves several techniques such as classification, multi-class classification, detection, and segmentation. Each of these techniques has its unique characteristics and applications in CAD. The comparison of these techniques is presented below and illustrated in fig. 2.6:

2.4.1 CLASSIFICATION

Classification is a technique that assigns a label or class to an input image or scans based on predefined categories. In pulmonary embolism, **classification** can differentiate between normal and abnormal samples. When multiple predefined classes are assigned to an input sample, the task is known as **multi-class classification**. For example, the *RSNA STR Pulmonary Embolism Detection Challenge* [11, 12] was a multi-class classification task where the labels are detailed in table 2.1.

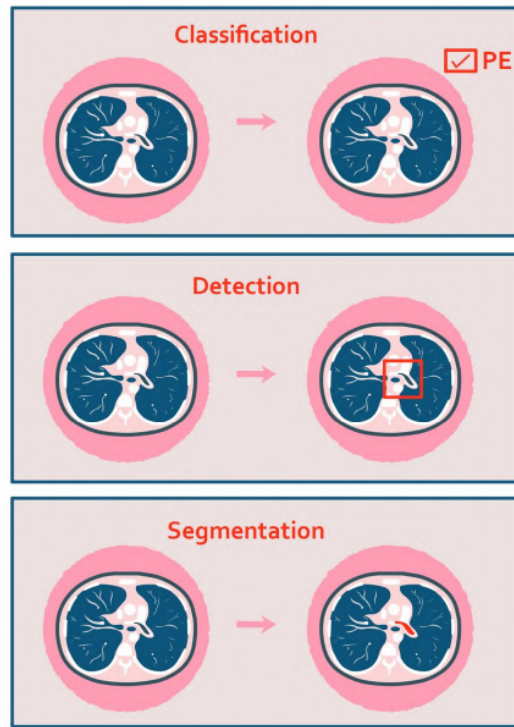


Figure 2.6: Main computer visions task for PE CAD. From Soffer, S. *et al.* in *Deep learning for pulmonary embolism detection on computed tomography pulmonary angiogram: a systematic review and meta-analysis* [48]. Licensed under a Creative Commons Attribution 4.0 (CC BY) License (<https://creativecommons.org/licenses/by/4.0/>).

2.4.2 DETECTION

Detection is a technique that involves locating and identifying objects of interest in an image. In the context of pulmonary embolism, detection can locate the embolus in the image/scan by positioning a point of interest or a bounding box as shown in fig. 2.6.

2.4.3 SEGMENTATION

Segmentation involves dividing an image into multiple regions or segments based on certain features or characteristics. In the context of pulmonary embolism, segmentation can identify the boundaries of the embolus and separate it from the surrounding tissue and liquid blood. Two subclasses of the segmentation task are instances and semantic segmentation. Semantic segmentation divides the image into different regions and assigns a label or class to each region based on its semantic meaning. Semantic segmentation can segment the image into different anatomical structures such as lungs, heart, or pulmonary arteries. On the other hand, instance segmentation

Table 2.1: Pulmonary embolism labels used and evaluated in the RSNA STR challenge [12] accompanied with a short definition.

Label	Definition
Negative for PE	No evidence of pulmonary embolism is present in the scan.
Indeterminate	The imaging test results are inconclusive and do not clearly indicate the presence or absence of pulmonary embolism.
Chronic	Refers to a pulmonary embolism that has been present for a longer period and may have resulted in permanent lung damage.
Acute & Chronic	Refers to a pulmonary embolism with both acute and chronic features.
Central PE	Refers to a pulmonary embolus located in the main pulmonary arteries or their branches.
Left PE	Refers to a pulmonary embolus located in the left pulmonary artery or its branches.
Right PE	Refers to a pulmonary embolus located in the right pulmonary artery or its branches.
RV/LV Ratio ≥ 1	Ratio of the right ventricular (RV) size to the left ventricular (LV) size is greater than or equal to 1, which is suggestive of acute pulmonary embolism and increased right heart strain.
RV/LV Ratio < 1	Ratio of the right ventricular (RV) size to the left ventricular (LV) size is less than 1, which is generally indicative of a normal ratio and absence of significant pulmonary embolism.

identifies and segments individual objects of interest in the image. Instance segmentation can identify and segment each embolus present in the image. This difference is illustrated in fig. 2.7.

In summary, classification and multi-class classification are used to classify the images as normal or abnormal or differentiate between different pulmonary embolisms. Detection is used to locate the embolus in the image, while segmentation is used to separate the embolus from the surrounding tissue. Depending on the requirements of the problem, one or more of these techniques can be used in combination to provide a more accurate and comprehensive diagnosis of pulmonary embolism. It is worth noting that classification can be performed at the slice or scan levels, where each slice or scan is assigned a label.

2.5 CLASSICAL SEGMENTATION ALGORITHMS

Classical or traditional image segmentation algorithms refer to methods based on classical computer vision techniques such as thresholding, edge detection, clustering, graph cuts, and region growing, among others. These algorithms have been widely used in various applications, such as object recognition, medical imaging, and video analysis. This thesis uses the "classical segmenta-

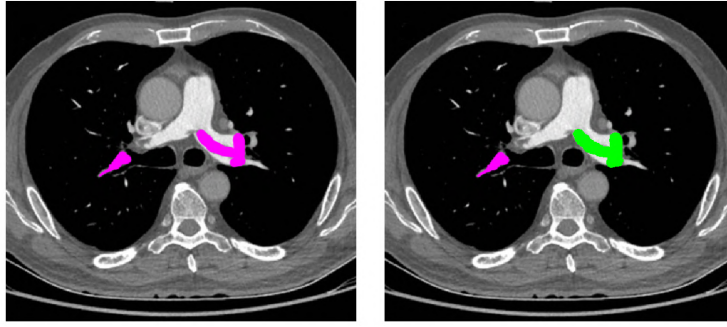


Figure 2.7: Semantic segmentation on the left image, as blood clots have the same labels. Instance segmentation on the right image, as each instance of a blood clot, has its own label.

tion algorithms" term to distinguish these methods from deep learning-based image segmentation algorithms.

For the task of detecting PE in CTA scans, Ozkan *et al.* proposed a multiple-step pipeline in [36] by segmenting multiple intermediate objects to obtain the PE. This approach requires the use of reference points indicated by the radiologist.

Masutani *et al.* introduced in [32] a method for PE detection that focuses on lung vessel segmentation and distinguishing PE based on specific features such as intensity, length, and volume. This work serves as a foundation for subsequent studies in the field. Pichon *et al.* in [38] and Zhou *et al.* in [64] also proposed a feature-based method.

However, Bouma *et al.* identified in [6] that intensity and shape alone might not be sufficient for PE detection. To address this limitation, they proposed a novel approach incorporating the shape of both the vessel and the lumen within the vessels as new features for classifying PE. This advancement in feature selection demonstrates the significance of considering additional properties beyond intensity and shape to enhance the accuracy of emboli detection.

2.5.1 PULMONARY ARTERIES SEGMENTATION

Certain CAD approaches, as outlined in [44], utilize a two-stage methodology whereby the pulmonary arteries are first segmented, followed by identifying PE candidates within the resulting mask. The former task is a specific instance of the widely researched challenge of segmenting vessels in 3D CT scans in the lungs, which has been the subject of much research and was evaluated in the Vessel12-challenge [41], where various methods and their performances are compared. Strategies, including Hessian, region-growing thresholding, and machine learning-based

techniques, have been proposed and offer promising results. However, these methods do not focus on the task of PE detection. They fail to segment the central vasculature, which is problematic since PE can also be located in these regions, as shown in fig. 2.8. Coming back to the task of PE

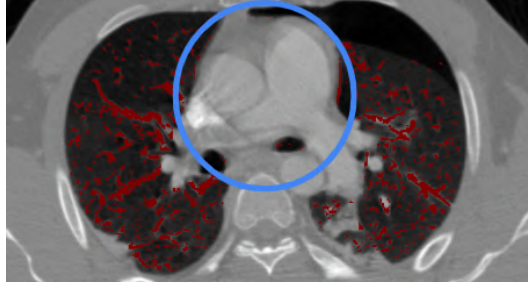


Figure 2.8: Results obtained by using a vessel segmentation algorithm to illustrate that the central vasculature (blue circle) is not included in the mask (in red).

segmentation and to address this limitation, Sebbe *et al.* [44] included the segmentation of the central vasculature in their approach, but encountered the difficult problem of touching vessels (i.e., the aorta, vena cava, and pulmonary artery have low-contrast boundaries between them), making it challenging to prevent the segmentation of the aorta and vena cava into the pulmonary artery (PA) class. The authors addressed this issue by requiring human interaction for locating and marking seven control points on each patient scan, making the process not fully automatic anymore.

Sedghi Gamechi *et al.* developed a fully automatic segmentation method for the pulmonary artery, including the central vasculature, by using optimal surface cuts [45]. This approach, however, is complex to implement, time-consuming, and requires a multi-atlas registration composed of 25 scans. Pu *et al.* also solved this problem more straightforwardly by using a U-Net deep learning network [39]. However, this approach requires labeled data for training the network. Zhang *et al.* proposed a simple system based on Harris corner detection and fuzzy connectedness [62]. Nevertheless, the method was trained on a private dataset that seemed to contain ideal cases.

To explore alternative methods that may prove effective in medical image analysis, we can draw parallels between specific challenges encountered in different domains. For instance, when addressing the issue of touching vessels, insights can be gained from segmenting cells, as demonstrated in the work of Browet *et al.* [8]. Based on graph cuts, their approach offers a promising avenue for potential adaptation and application to our problem. In the following section, we delve into the details of this method.

2.5.2 GRAPH CUTS WITH ALPHA-EXPANSIONS

The underlying method in [8] used for cell segmentation is energy minimization with graph-cuts which is proposed by Boykov *et al.* in [7]. The energy function utilized for segmentation typically comprises two key components: a data term that assesses the degree to which the segmentation corresponds to the input data and a regularization term that guarantees smoothness and consistency in the segmentation. Various metrics reflecting the semantic properties of the segmentation task can be employed to define the data term, whereas graph cuts are typically utilized in the regularization term.

The energy function is as follows:

$$E(f) = E_{smooth}(f) + E_{data}(f), \quad (2.1)$$

E_{smooth} is the energy term that measures the extent that connected pixels are piece-wise smooth [7]. E_{data} measures the disagreement between f and the observed data. Regarding the E_{smooth} term, we aim to achieve good segmentation for boundaries. Therefore, instead of using functions that make f smooth everywhere, we use a discontinuity-preserving function [7].

2.6 DEEP LEARNING METHODS

In recent years, deep-learning-based machine learning methods have been highly effective in various tasks, particularly in supervised learning tasks such as segmentation, classification, and regression. These models are trained on large datasets consisting of numerous training examples, where each sample comprises a feature that describes the instance and a label indicating the ground-truth output. However, obtaining labeled data can be costly in terms of time and money, especially in medical imaging, where data is of high resolution, and marking every pixel requires the expertise of highly paid healthcare professionals.

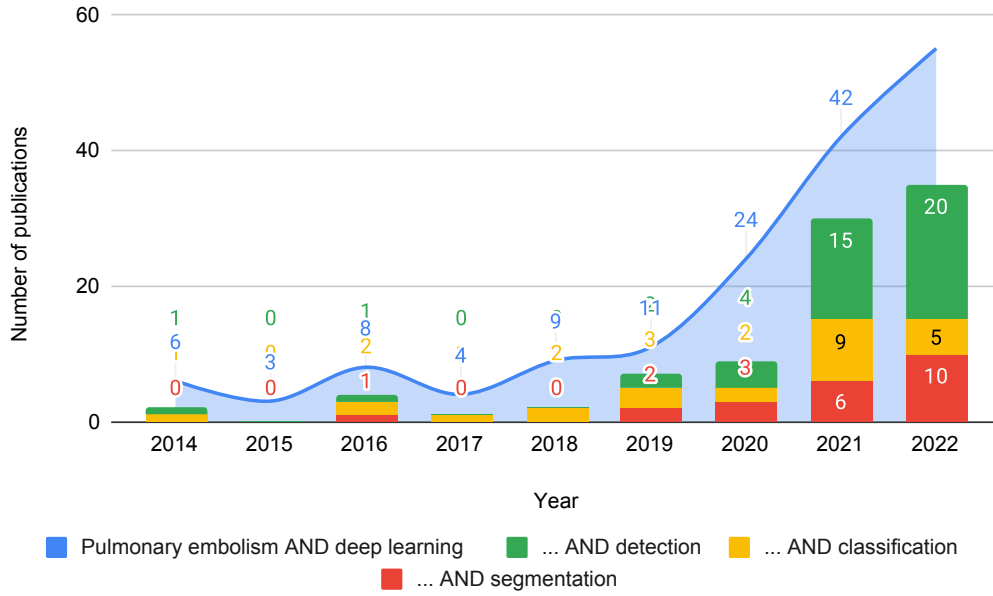


Figure 2.9: This graph was made using data obtained on 10/05/2023, from Digital Science’s Dimensions platform, available at <https://app.dimensions.ai>. The keywords were queried in the title and abstract of papers.

As shown in fig. 2.9, there has been an increasing trend in using deep learning networks for pulmonary embolism CAD since 2020. Over the past three years, there has been a shift in research focus from classification to detection and segmentation. This may be attributed to the challenge of explainability and interpretability of CAD software (detailed in section 2.3.2). In the case of detection and segmentation, radiologists can directly observe the embolus in the scan, whereas, in classification, radiologists need to search the entire scan to validate the CAD result. As a result, research in detection and segmentation is more valuable and interesting to solve. It is pertinent to note that the distinction between detection and segmentation limits the interpretation of this graph. It can sometimes be unclear. Furthermore, it is essential to emphasize that not all articles included in the graph aimed at PE segmentation; some focused solely on segmenting the arteries as a prior step and were thus included as segmentation algorithms. To be more specific, to our knowledge, only a limited number of papers directly address PE segmentation using deep learning techniques. For instance, Liu *et al.* explored this topic in [28].

The main challenge for deep learning methods in many applications, including medical imaging, is labeled data scarcity. However, several methods have been proposed to handle this problem, which is discussed in the following section.

2.6.1 WEAKLY SUPERVISED MODELS

To address this issue for the segmentation task, weakly supervised segmentation methods have been developed to reduce the requirement for labeled data. These methods will be discussed in detail in the following sections. These methods allow training a model by requiring fewer annotations from a radiologist (decreasing the cost) or by taking advantage of low-cost labels from system-biased segmentation algorithms or non-professional crowds (increasing the quality). These two improvements are illustrated in the following figure:

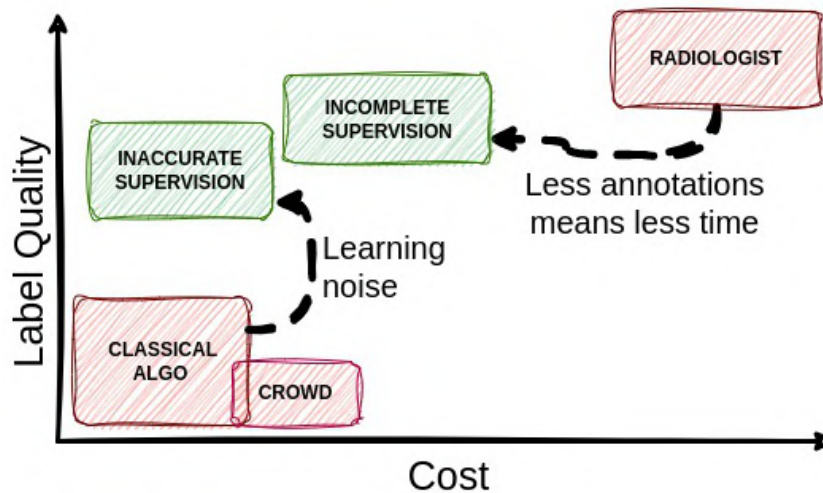


Figure 2.10: Weakly supervised models solve the label’s cost-quality trade-off.

INCOMPLETE SUPERVISION

These models leverage a combination of labeled and unlabeled data to improve the accuracy and generality of predictions. In this scenario, only a subset of the training images is annotated. This subset of images is fully pixel-wise labeled. The other images are unlabeled and form the majority of the training set. This decreases the overall cost of training the model as only a small subset of the data comes with a high-labeling cost. Here are some strategies called semi-supervised, which assume no human intervention during the training process:

- **Transfer learning:** Instead of training a new model from scratch, transfer learning leverages the knowledge and learned patterns from a pre-trained model to improve the performance and reduce the training time for a new task [66]. The pre-trained task is chosen such that the labeling cost is very low.
- **Prior knowledge** can take many forms, including expert knowledge, domain-specific rules, and constraints, and can be used to guide the learning process. An example of this strategy

is the artery-aware 3D fully convolutional network introduced by Guo et al. [17] in September 2022, where the task is to locate the PEs in the arteries mask. In this case, the arteries mask is considered as prior knowledge.

- **Data augmentation** consists of artificially increasing the size and diversity of the training set by reusing and performing operations on already labeled data. For example, in [47], they use Generative Adversarial Neural Networks to augment the scan to a healthy version and then subtract the produced scan to detect unhealthy parts. For the task of segmenting female pelvic organs on Cone Beam CT scans, J. Leger, E. Brion, *et al.* from *UCLouvain* improved performance in [26] by augmenting the data with pelvic scans from CT scanners because the dataset contained more labeled scans. In this specific example, the method is called Cross-Domain Data Augmentation.
- **Pseudo-labeling** (also known as self-training) was introduced by [25]. This technique works by repeating four steps: train the model on a batch of labeled data; predict unlabeled data using the model; use the predicted labels to calculate the loss; combine labeled and unlabeled loss and backpropagate.
- **self-supervised learning** allows discovering data-inherent patterns by defining a pretext task for which no labeled data is necessary. The supervisory signals learned in the pretext class are still usable and interesting for the task of interest. Contrastive learning, a method of self-supervision, was explored for pulmonary embolism detection by Florine Thiry in her master thesis at EPL [52]. However, note that this was not a segmentation task but a classification.

The description provided above assumes that samples are labeled or unlabeled. However, some other partially supervised methods assume that the separation between labeled and unlabeled can be defined within a subset of the sample. A toy example illustrating these differences for a CT scan dataset is shown in Figure 2.11. Like semi-supervised models, partially supervised approaches require the radiologist to sparsely annotate a subset of the scans, allowing for training a model at a lower cost. In the *3D U-Net* model as presented in [10], the sparsity is applied per-slice basis, meaning that only a subset of the slices in a sample scan is labeled.

Other methods, called **active learning**, take advantage of human interaction during the training process by querying in an optimized way ground truth for the most representative, informative, and uncertain regions or samples [37]. This process is schematically represented in fig. 2.12. This increases performance faster with a reduction of labeling cost (i.e., by minimizing the number of queries). This technique was utilized by Wang *et al.* in [56, 57] for detecting regions in a 3D CT scan containing pulmonary nodules. In [56], they achieve similar performances to a model

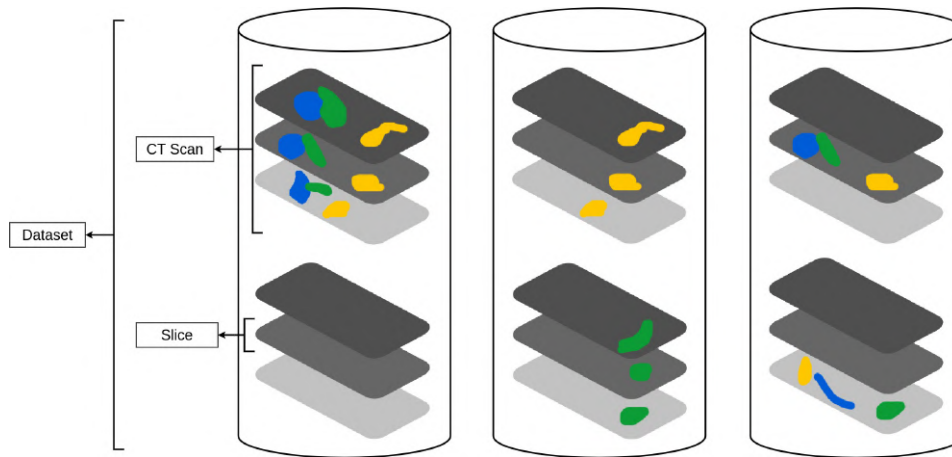


Figure 2.11: Three toy examples of limited-labeled datasets commonly encountered in medical imaging research. Semantic classes are shown in blue, yellow, and green. The annotations are incomplete in the first dataset, shown on the left, and only a subset of the scans are labeled. This is known as the incomplete-supervision case. The two datasets on the right are used for partial supervision, as defined by Zhou in [65]. In the middle dataset, annotations are available for all the scans, but only one class is labeled. All the scans are annotated in the third dataset, but only a subset of the slices are fully labeled. This can occur when radiologists focus on specific regions of the image that are of greater diagnostic relevance.

relying on the full annotated dataset, but By using only 16% of annotated samples and querying 26% of unlabeled samples.

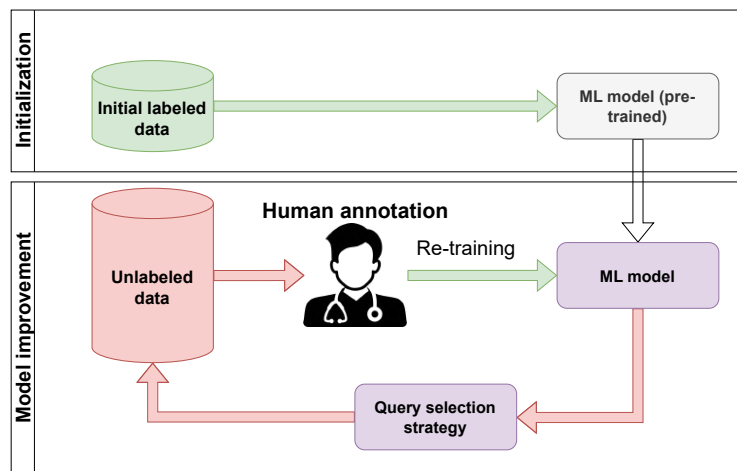


Figure 2.12: Active learning

INACCURATELY SUPERVISED SEGMENTATION

Inaccurate supervision refers to strategies where the supervision information is not always ground truth. This means that the data typically used as ground truth contains errors. These errors can arise from noisy labeling processes, mislabeling, imprecise boundaries, or when using bounding boxes as a tool. For instance, in the case of PE detection, one could imagine, as illustrated in Figure 2.10, a low-cost increase in labeled data by utilizing inaccurate computer detection algorithms that are system-biased. Another approach could involve employing a crowd of non-expert human annotators (e.g., medical students, quick-formed volunteers, etc.) who can be acquired at a lower cost but may introduce more errors. Karimi *et al.* reviewed various state-of-the-art deep learning techniques and remedies for noisy labels in medical image analysis in [21]. This review highlights that methods rely not only on the network architecture (e.g., they rely on preprocessing label cleaning, training procedures, loss functions, etc.). However, they are more studied in detection and classification tasks than in segmentation. When exploring the task of segmenting medical images, we found that Han Le *et al.* [24] addressed the problem of automatically segmenting regions of pancreatic cancer using a small amount of high-quality labeled data along with a substantial amount of noisy-labeled data. They followed the idea proposed by Ren *et al.* [40], which involves assigning a lower weight to noisy labels than clean labels during the gradient update process for loss propagation.

COMBINING STRATEGIES

The task of vessel segmentation on CT images was tackled by Zhang *et al.* in [63] using a combination of two strategies: incomplete supervision and inaccurate supervision. In their approach, the authors trained their model on computer-generated labels obtained from automatic vessel enhancement using classical algorithms. However, these labels were noisy and imperfect. The authors employed an active learning strategy to mitigate the limitations of noisy labels. This strategy estimates the model's uncertainty using the method proposed by Gal *et al.* in [14] and queries a radiologist to provide ground truth at locations where the uncertainty is high. Combining incomplete and inaccurate supervision allows the authors to leverage both strategies' strengths, leading to a more robust and accurate vessel segmentation model.

3

RESEARCH QUESTIONS AND HYPOTHESES

In this chapter, the research questions and hypotheses of this thesis will be presented, which aim to investigate the creation of a new framework for segmenting pulmonary embolus on CTA scans using weakly supervised techniques and prior labels-generation with a classical segmentation algorithm. The research questions and hypotheses will be organized under one main objective and two sub-objectives:

Objective 1: Design a deep learning-based framework for segmenting (i.e., localizing precisely) pulmonary embolus on CT scans under limited annotation. Under the assumption that a radiologist can be queried by the model during the training process.

Research Question 1: Which existing weakly supervised deep learning techniques can be used to develop an accurate and robust framework for segmenting pulmonary embolus on CT scans ?

Hypothesis 1: As existing frameworks exist for tasks dealing with other medical imaging questions, we could adapt these methods to segmenting pulmonary arteries.

To answer this research question, two sub-questions and hypotheses are formulated:

Objective 1.1: Explore the implementation of a prior classical segmentation algorithm for (PE) and use it as training labels (potentially mislabeled) in the proposed deep learning framework.

Research Question 1.1: Can classical segmentation algorithms be used to localize PE in CT scans and generate a pool of PE candidate labels for training the deep-learning method in the proposed framework?

Hypothesis 1.1: We hypothesize that classical segmentation algorithms can detect a high number of candidate PE regions with high sensitivity; however, these regions may also contain false positives leading to low specificity.

Objective 1.2: Incorporate prior segmentation of anatomical parts in the chest for improved PE segmentation

Research Question 1.2: How does the use of prior segmentation of anatomical parts (i.e., pulmonary arteries, aorta, vena cava, etc...) in the chest impact the performance of the generated PE candidates ?

Hypothesis 1.2: The incorporation of prior segmentation of anatomical parts in the chest will improve the specificity as candidates outside the pulmonary arteries will no longer be selected.

In order to address these research questions, a possible framework will be proposed using a combination of weakly supervised techniques and prior classical segmentation of PE. The efficacy of this prior segmentation, which is utilized as training labels, will be evaluated by examining its effectiveness in detecting PE candidates. The evaluation will be conducted on a publicly available dataset with ground truth labels. To quantify the improvement due to the use of prior anatomical segmentation, statistical analysis will be performed. While the complete implementation of the proposed framework was not feasible within the scope of this study, it is left as a potential avenue for future research.

Overall, this thesis aims to contribute to medical imaging and healthcare by developing a new framework for segmenting pulmonary embolus on CT scans that is accurate, robust, and cost-effective. By answering the research questions and hypotheses outlined in this chapter, the proposed framework can potentially improve the detection and diagnosis of pulmonary embolism and ultimately improve patient outcomes.

4 MATERIALS AND METHODOLOGY

4.1 MATERIALS

4.1.1 DATASET

The dataset used in this thesis is called FUMPE [31] and consists of CTA scans. It is publicly available and was downloaded from Kaggle¹. The dataset comprises 35 patients, 17 males and 18 females, aged between 24 and 82 years old. Each patient scan contains, on average, 251 slices (i.e., axial images), with a minimum of 139 and a maximum of 475 slices per scan. The dataset contains 8792 slices in total, out of which only patients 24 and 32 do not have PE clots.

Each slice in the dataset comprises a 512x512 pixel array, with a slice thickness of 1mm (except for Patient24 and Patient32 with slice-thickness=2mm). Eleven of the patients had high mortality risk. The dataset was obtained after the authors obtained ethical approval from the ethics committee, and informed consent was obtained from each of the patients included in the dataset.

The CT scanning was performed using the same scanner model in two medical centers. Contrast material was injected into the patient's vein (see section 2.2.2), and the images were acquired using a standardized protocol.

4.1.2 TOOLS

In this thesis, the primary programming language used was Python, a popular choice for scientific research due to its extensive collection of libraries and tools for data analysis, machine learning, and image processing. Specifically, two image processing libraries, scikit-image, and SimpleITK, were utilized. Scikit-image offers a comprehensive collection of image processing algorithms for image filtering, segmentation, and feature extraction. On the other hand, SimpleITK provides a powerful image-processing library that offers a wide range of functions for image registration, segmentation, and analysis. Graph processing was accomplished using NetworkX.

¹FUMPE dataset: <https://www.kaggle.com/datasets/andrewmvd/pulmonary-embolism-in-ct-images>

To handle the DICOM data format of the CT scans, the PyDicom library was used. PyDicom is a Python library that offers a simple interface for reading, writing, and modifying DICOM files. Two libraries were employed for 2D and 3D visualization of the scans and intermediate results, namely Napari and Matplotlib. Napari is an open-source, interactive, fast Python library for multi-dimensional image visualization. It provides an intuitive graphical user interface allowing easy exploration and manipulation of CT scans and intermediate results. The stack of libraries employed in this thesis is shown in the following figure.

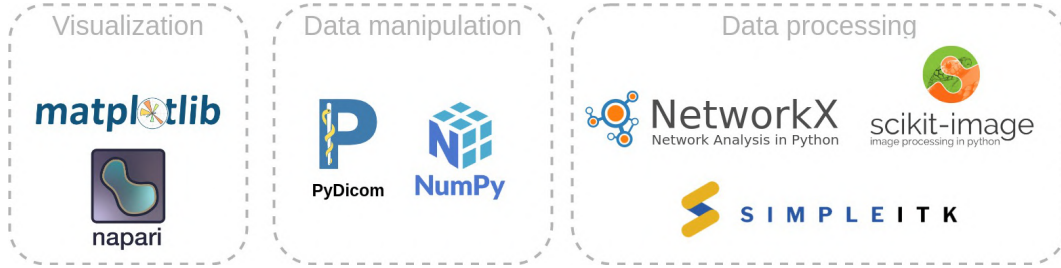


Figure 4.1: Tools Used for this Thesis

4.1.3 HARDWARE

Cloud computing was used to expedite the exploration and comparison of parameters for image processing algorithms. Paperspace Gradient was utilized for this purpose, equipped with a high-performance RTX4000 8 GB GPU, 30 GB Memory, and 8 vCPUs.

A standard portable computer was used to analyze the results, equipped with an Intel® Core™ i7-8550U CPU @ 1.80GHz × 8 with 8.0 GiB memory. This setup provided sufficient computational power for the data analysis and visualization tasks, such as generating statistical summaries, creating graphs, and preparing figures for the thesis.

4.2 ENLARGED OVERVIEW OF THE FRAMEWORK

This section provides an overview of the framework chosen for this thesis, which was inspired by Zhang *et al.*'s work on weakly-supervised training for coronary artery (blood vessels that supply oxygenated blood to the heart muscle) segmentation [63]. However, the current study focuses on a different task with semantically different annotations. The framework comprises two main stages: the first involves generating pseudo-labels using classical segmentation algorithms. In contrast, the second stage involves training a deep learning neural network and implementing a loss function that includes suggestive annotation (i.e., active learning).

In both stages, the input and output consist of the following:

- **Input:** 3D volumetric CTA scan with Hounsfield units for each pixel (with preprocessing).
- **Output:** 3D binary-volumetric scan where

$$P[x, y, z] = \begin{cases} l & \text{if pixel } x, y, z \text{ predicted as part of PE with instance } l. \\ 0 & \text{otherwise.} \end{cases} \quad (4.1)$$

, where x, y, z are, respectively, the row, column, and slice indexes. The value of $l \in \mathcal{P}$ identifies the instances of the PE (e.g., if we have three blood clots in the scan, we have $|\mathcal{P}| = 3$).

The necessity to develop a unique segmentation algorithm for the initial stage emerged due to the frequent commercial or proprietary nature of existing classical techniques, their lack of interoperability, and difficulty in adaptability. Given the complexities in implementing such segmentation software, this study's scope was confined to the development of the first stage, keeping in mind the larger two-stage framework discussed herein. Each decision in the methodology was made considering the broader perspective of the solution.

The second stage of the framework involves training a deep-learning neural network using the pseudo-labels generated in the first stage and the suggestive annotations provided by radiologists. The neural network is trained using an active learning strategy, which involves iteratively selecting the most useful regions for manual annotation and incorporating these new annotations into the training process.

The first subsection aims to introduce both stages briefly. After this, we will delve deeper into the methodology and intermediate results of the first stage, as it is the main focus of this thesis. Regrettably, we could not implement the second stage due to time constraints. Nonetheless, it serves as a foundation for future work.

4.2.1 TWO-STAGE FRAMEWORK AND ALGORITHM OVERVIEW

The two-stage framework is depicted schematically in fig. 4.2. The first stage encompasses all tasks related to data preprocessing, including noise reduction, removal of unnecessary information, and label generation using classical algorithms. The second stage focuses on the deep learning approach, involving query selection for suggestive annotation by the radiologist. Then, the framework integrates the radiologist's feedback into the loss function for model improvement.

The deep learning model is trained to refine its predictions iteratively based on the radiologist's annotations and inaccurately-generated labels. It is important to note that a single slice of the scan is illustrated to facilitate comprehension of each step within the framework. However, in practice, the dataset consists of multiple slices. Representing the entire 3D volume in this schematic was not feasible. The preprocessing step comprises a pipeline incorporating intensity-based opera-

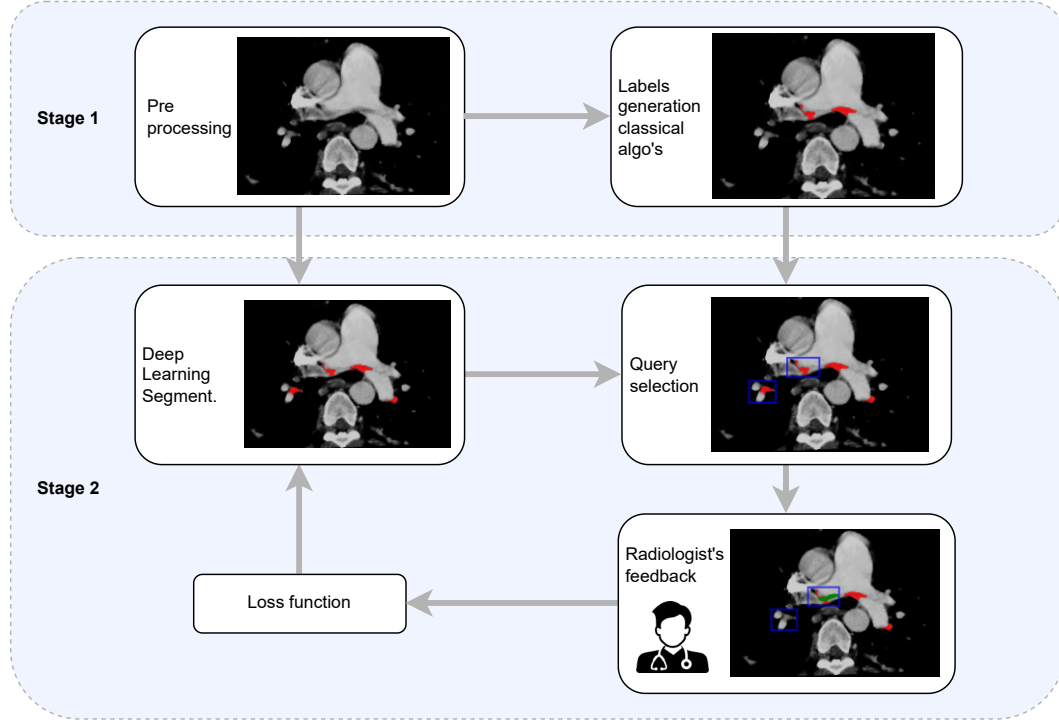


Figure 4.2: Two-Stages Framework

tions and masking/cropping operations. This allows for normalizing pixel ranges between patients and isolating the anatomical region of interest within the scan before applying segmentation algorithms. The label generation process, which utilizes classical algorithms, is underpinned by a superpixel generation process, followed by a graph-cut algorithm to semantically segment anatomical sections of the chest and then PE candidate identification. The details of these steps will be discussed in the following section.

As previously mentioned, implementing the second step was not feasible within the scope of this thesis. However, we will outline the design that would have been explored given sufficient time. The deep learning model would have been based on 3D-UNet, as proposed by Çiçek et al. in [10], which is capable of segmenting a full 3D CT volume by training the U-Net model with annotated 2D slices from the same volume. Another potential architecture is the V-Net, a direct

variant of the U-Net that internally uses 3D convolutional neural networks [34]. The query selection component, crucial for suggesting the most informative annotations to the radiologist, remains an area for further investigation. The objective would be to design an uncertainty function to evaluate how much ground truth annotations could contribute to model improvement. The loss function and subsequent backpropagation (i.e., for integrating feedback into the model) would follow the approaches described in [10, 34], depending on the specific deep learning model chosen for the framework.

The subsequent sections will delve into the details of the algorithm’s first stage, comprehensively examining its methodology and intermediate results.

4.3 STAGE 1: METHODOLOGY AND INTERMEDIATE RESULTS

The complete pipeline, presented schematically in fig. 4.3, provides an overview of the sequential steps involved in our methodology. Firstly, we employ a preprocessing step to enhance edge details and improve the overall image quality of the raw CT scans. Next, we generate superpixels using the Quickshift method and employ post-processing techniques to address the issue of disconnected pixels within a superpixel region. To estimate the class scores for each superpixel, we compute aggregated data by considering the intensities and characteristics within the region. This data aggregation allows us to infer the likelihood (i.e., scores) of a superpixel belonging to specific classes. Our approach initially estimates scores for the background and vena cava regions, followed by the aorta and pulmonary arteries. The corresponding parts and their arrangement are depicted in fig. 4.4, visually representing the circuitry. We construct a graph linking superpixels once the class scores have been computed. Finally, we apply post-processing to detect candidate emboli within the pulmonary arteries. These candidates serve as the starting point for the second stage of the framework.

4.3.1 PREPROCESSING

The first crucial step is to preprocess the data to improve the quality of medical images. This involves focusing on specific regions of interest (i.e., emboli are located in pulmonary arteries, themselves located in lungs and the central vasculature) and pixel intensities within a certain range of Hounsfield units (i.e., 0 to 350, see table 4.1). This approach enhances robustness when dealing with scans from different patients captured by various CT scanners.

The preprocessing step includes two masking and cropping steps to exclude information outside the lungs and three intensity-based operations to improve image quality and reduce noise. These steps are schematized in fig. 4.5. The body was cropped slice-wise using the Otsu threshold

4 Materials and methodology

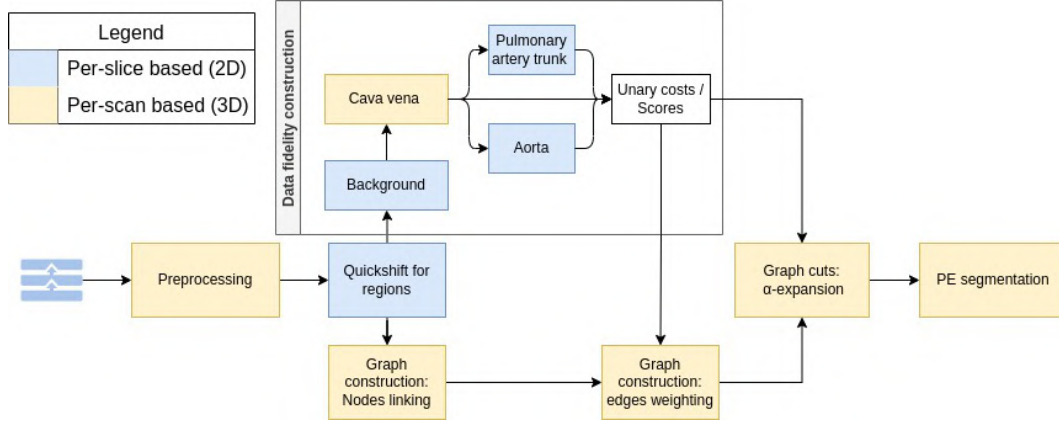


Figure 4.3: Schema of the pipeline. Colors depend on the type of computation performed: per slice on 2D images or superpixels in blue; for the whole scan in a 3D context in yellow. Arrows highlight the inputs used for each stage.

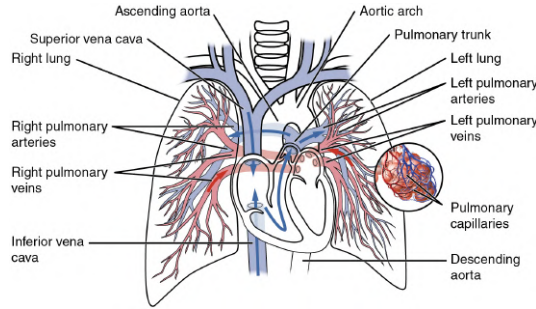


Figure 4.4: Illustration from Anatomy & Physiology, Connexions Web site. <http://cnx.org/content/col11496/1.6/>, Jun 19, 2013. Under CC BY 3.0 license <https://creativecommons.org/licenses/by/3.0/>

method [35, 42], followed by a sequence of morphological opening, closing, and finally, a contour detection algorithm. Only the contour with the biggest inner area was kept [2]. The average value of the slice-wise Otsu threshold was then used as a reference for the lung lobes extraction. In this case, only the largest volume of the thresholded slices was kept as lung lobes. A new stack of images was cropped from this mask.

Applying intensity-windowing, guided by the specific medical question at hand, is useful in retaining only valuable data [27]. In our case, focusing on PE, we are interested in intensities ranging from acute blood to contrast-enhanced blood. Looking at corresponding Hounsfield Unit from [19, 46] in table 4.1, we selected the $[0; 350]$ – HU range. We then perform histogram equalization and smoothing with a Curvature Anisotropic Diffusion filter [58]. It was chosen as a filter because it allows smoothing for complex, noisy surfaces while preserving (and enhancing)

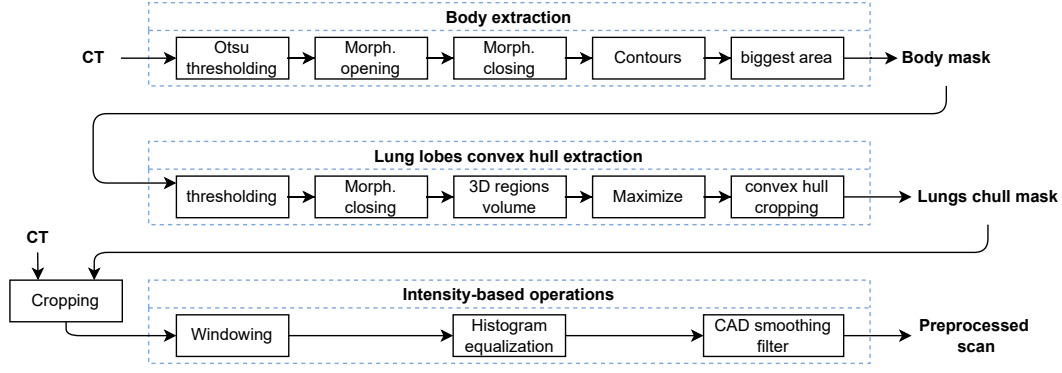


Figure 4.5: Preprocessing Steps

sharp, geometric features [51]. This method has shown to be effective for vessel enhancement in [30].

Table 4.1: Typical values in Hounsfield Units [19, 27, 46].

Substance or organ	Typical HU value
Air	- 1000
Normal lungs	-700 to -900
Fat	-100
Water	0
Acute blood	50 to 80
Muscle	50
Intravenous contrast	300
Bone	> 1000

INTERMEDIATE RESULTS

The preprocessing results, including each cropping and intensity-based operation, are illustrated in 4.6. By performing these operations, we observe that the quality of medical images can be significantly improved, leading to more accurate diagnosis results and better outcomes. The cropping offers a dual advantage, encompassing the removal of superfluous information while concurrently reducing the size of the scan.

4.3.2 SUPERPIXELS GENERATION

Given the specific characteristics of medical images, namely their large size, the intensity-based information they carry, and their low contrast and non-continuous boundaries, adopting a su-

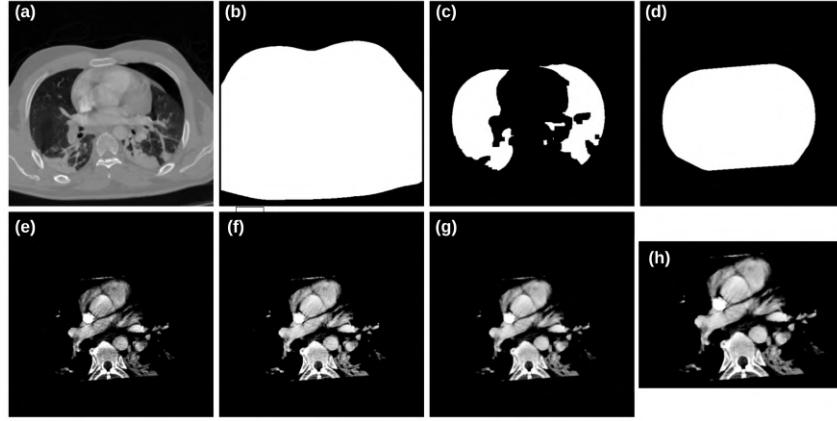


Figure 4.6: Outcomes of preprocessing steps. Mask-based operations: Original raw image in (a), body mask in (b), lungs mask in (c), and convex hull of lung mask in (d). Followed by intensity-based operations: windowing Hounsfield unit between 0 and 350 in (e); histogram equalization in (f); Curvature Anisotropic Diffusion (CAD) filter in (g); and finally, the cropped version in (h).

perpixel approach has been deemed an interesting approach to explore. Superpixels enable the grouping of pixels into regions, enhancing computational efficiency and defining boundary edges.

We identified three methods for solving this task: KMeans (SLIC), Felzenszwalb’s, and Quickshift clustering algorithms. Felzenszwalb’s algorithm, proposed in [13], allows the clustering of pixels on 2D images and is a graph-based method. In contrast, K-Means has the advantage of working for 3D images [1]. Finally, Quickshift is a more recent method from [54] that allows fast computation of hierarchical segmentation on multiple scales.

INTERMEDIATE RESULTS

Identified methods were evaluated for the task at hand, as depicted in the accompanying fig. 4.7. Ultimately, the quick shift approach was selected due to its rapid computational speed and tendency to establish region boundaries on anatomical object boundaries. The accurate positioning of these boundaries can be explained by the multi-scale property of this method [54]. To achieve these region delimitations, we performed parameter tuning for the Quickshift algorithm, specifically adjusting the sigma and kernel size values. Sigma represents the width used for Gaussian smoothing during the internal preprocessing, with a value of zero indicating no smoothing applied. On the other hand, kernel size refers to the width of the Gaussian kernel used for the sample density, where a higher value corresponds to fewer clusters. We chose not to use the preprocessing (i.e., $\sigma = 0$) and settled on a kernel size of 5. Other combinations of values were tested, but they are not illustrated in this thesis.

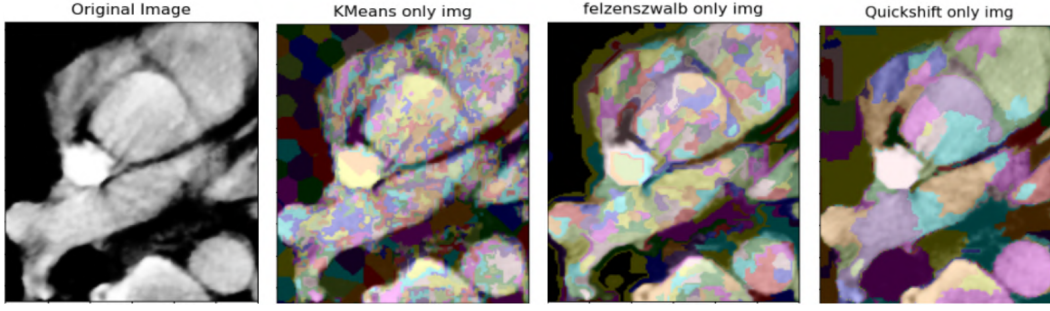


Figure 4.7: Comparison of three different superpixels generation algorithms.

Let's define the superpixels mathematically using array L as a 3-dimensional array where each element corresponds to a label l associated with the pixel at the given index:

$$L : \mathbb{N}^{n \times m \times p} \rightarrow \{1, 2, \dots, n_superpix\} \quad (4.2)$$

where n , m , and p correspond to the dimensions of the CT scan. Given that the Quickshift algorithm is designed to work for 2D images, we apply it slice by slice. Therefore, all pixels associated with the same label l are necessarily in the same slice. This constraint can be expressed mathematically as:

$$\forall l \in \{1, 2, \dots, n_superpix\}, \exists s \in \{1, 2, \dots, p\} \text{ such that } \forall (x, y, z) \in L^{-1}(l), z = s \quad (4.3)$$

This expression states that for every label l in the set of all labels (from 1 to $n_superpix$), there exists a slice s (from 1 to p , where p is the number of slices) such that for all pixels (x, y, z) mapped to the label l by L , the z -coordinate (slice index) is s .

Furthermore, we have removed superpixels containing only pixels with values close to zero. This approach allows us to decrease the problem's size while preserving the essential information. This can be mathematically expressed as follows:

$$\forall l \in \{1, 2, \dots, n_superpix\}, I_l = \frac{1}{|L^{-1}(l)|} \sum_{(x,y,z) \in L^{-1}(l)} I(x, y, z) \quad (4.4)$$

$$\forall l \in \{1, 2, \dots, n_superpix\}, I_l < \epsilon \Rightarrow L^{-1}(l) = 0 \quad (4.5)$$

The first equation calculates the mean intensity I_l for each label l . The second equation states that if I_l is less than ϵ , then the label l in array L is set to 0. Note that $L^{-1}(l)$ denotes the set of pixels associated with label l , and $|L^{-1}(l)|$ is the number of such pixels. The eliminated regions (i.e., the

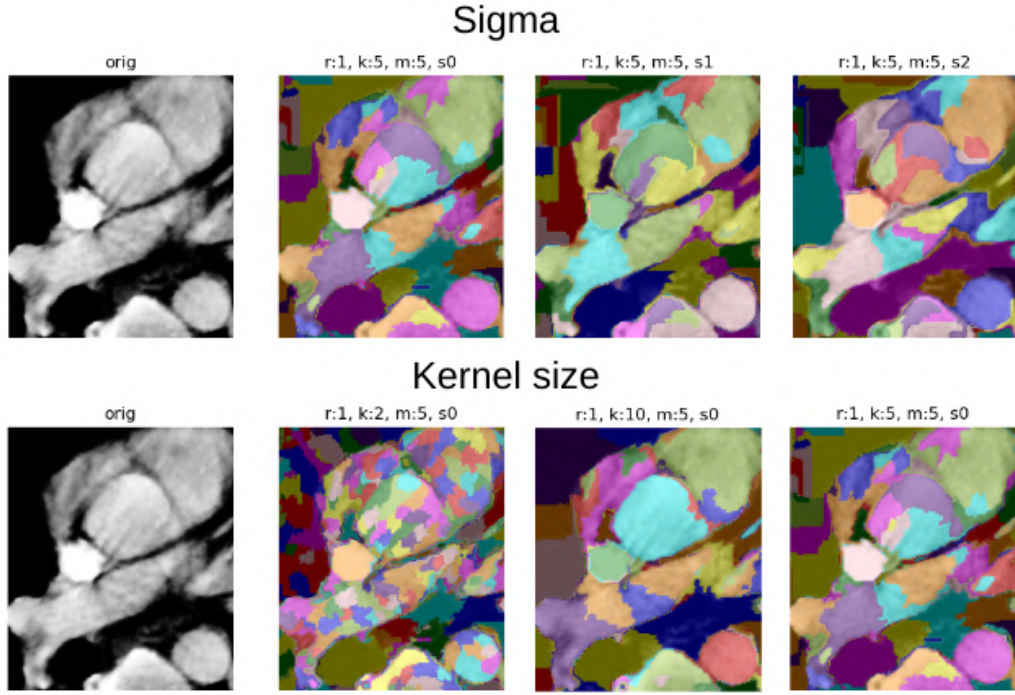


Figure 4.8: Parameter tuning for the Quickshift method.

regions set to label 0) can be classified as background (i.e., non-valuable for segmentation) after the segmentation process.

Typically, a CT scan contains about 10^7 pixels, whereas utilizing superpixels reduces this figure to approximately 10^4 superpixels. This intelligent grouping approach significantly reduces the amount of data while minimizing the loss of information.

4.3.3 ENERGY MINIMIZATION AND GRAPH CUTS

Now that we have extracted superpixels from the scan, the segmentation task can be approached using the concept of energy minimization. In this thesis, the focus is on identifying two classes, namely emboli, and others. However, due to the difficulties of the task, it may be beneficial to incorporate additional semantic classes, such as vena cava and aorta, to enhance the accuracy of the segmentation. Adding these classes allows additional information and constraints to be utilized, and prior knowledge about the human body structure can be incorporated. Moreover, detecting emboli that are present in pulmonary arteries can be accomplished using a post-processing step.

The study by Browet *et al.* inspired the main methodology we employ here [8]. To adapt this approach to our segmentation task, we modify the graph structure utilized in their study. Specifically, Browet *et al.* form their graph based on neighboring pixels; however, in our case, we define links between superpixels based on certain spatial conditions. This modification allows us to use the alpha expansion method to solve the energy minimization problem. Additionally, we aggregate data over the superpixels, such as by taking the mean value of all pixels within a superpixel.

To understand why this generalization is feasible, we examine the definition of the energy function in [7]:

$$E(f) = E_{smooth}(f) + E_{data}(f), \quad (4.6)$$

Let's clarify this function applied to our case. We define p as the superpixel id, with $p \in \{1, 2, \dots, n_superpix\}$. Each superpixel p is assigned to a class $f_p \in \{\text{background, pulmonary artery, aorta, vena cava}\}$.

The data fidelity term $E_{data}(f)$ is defined as the sum of cost function $D_p(f_p)$ across all superpixels in set $\mathcal{P}$. The function $D_p(f_p)$ represents some measure of fit between the observed data for superpixel p and the class f_p to which p is assigned. This can be represented mathematically as:

$$E_{data}(f) = \sum_{p \in \mathcal{P}} D_p(f_p), \quad (4.7)$$

The smoothing term E_{smooth} is defined as the sum of a function $V_{p,q}(f_p, f_q)$ over all pairs of neighboring superpixels in set $\mathcal{N}$. This set is defined as

$$\mathcal{N} : \mathcal{P} \rightarrow 2^{\mathcal{P}} \quad (4.8)$$

$$\mathcal{N} = \{(p, q) \mid p \in \mathcal{P}, q \in Neighbor(p)\} \quad (4.9)$$

This set $\mathcal{N}$ includes all pairs of superpixels (p, q) such that p is in $\mathcal{P}$ and q is a neighbor of p . The function $V_{p,q}(f_p, f_q)$ imposes a penalty based on the difference between the class assignments of superpixel p and its neighbor q . This can be represented mathematically as:

$$E_{smooth} = \sum_{\{p,q\} \in \mathcal{N}} V_{\{p,q\}}(f_p, f_q), \quad (4.10)$$

If $V(f_p, f_q) = |f_p - f_q|$ and the number of labels is finite, then the exact minimum of the energy function can be found in polynomial time via graph cuts. However, in general, this problem is NP-hard, and thus, we need to use Boykov's fast method.

4.3.4 DATA FIDELITY

This section describes the function $S_p(f_p)$, which represents the unary score of assigning the label f_p to the pixel p . This score is afterward used to compute $D_p(f_p)$, which is the unary cost defined in eq. (4.11). This score is computed for each superpixel and $S_p(f_p) \in [0; 1]$.

For a scan containing $n_superpix$ superpixels and 4 (i.e., background, vena cava, pulmonary arteries, and aorta) different labels, D can be seen as an array of dimension $(n_superpix, 4)$. In the following sections, we define a score function for each of the four semantic labels that outputs a floating point value between 0 and 1 and attributes a higher value if the superpixel is more likely to be labeled with this semantic class. Later, we use a cost function instead of a score when using the energy minimization. The relation between both is simply given by:

$$D_p(f_p) = 1 - S_p(f_p). \quad (4.11)$$

Before digging into the scoring functions, we must introduce the concept of a **knowledge map**. A knowledge map is a simple, approximate atlas or reference model that guides the prior anatomical segmentation process. More specifically, it is a set of 2-dimensional points on the axial plane, each linked to a semantic class: background (e.g., the part corresponding to the spine), aorta, and pulmonary artery.

As illustrated in the pipeline schema in fig. 4.3, our approach involves first computing the scores for the background, followed by the vena cava, aorta, and pulmonary arteries. To identify the aorta and pulmonary arteries, we use the known position of the vena cava and score superpixels based on their relative position to the vena cava. This technique helps improve the accuracy of the results by leveraging spatial information in the image.

Let's now mathematically define these 2D points. First, we define the full scan C , the set of all points (x, y, z) in the CT scan.

$$C = \{(x, y, z) \mid x, y, z \in \mathbb{N}\} \quad (4.12)$$

Then, we define a slice $S(z)$ as a subset of C . $S(z)$ includes all points (x, y) in the axial plane at slice number z .

$$S = \{(x, y) \mid (x, y, z) \in C\} \quad (4.13)$$

A visual overview of reference points can be found in fig. 4.9. This 2D knowledge map was developed for this thesis, but it could be extended to a 3D knowledge map to improve the results, albeit at the cost of taking more time to define generic points that work for multiple scans.

The reference points on the knowledge map are divided into two categories:

- The set **KF** of **fixed points** is a set of indexes of the CT scan array. They are 'fixed' because each point directly corresponds to a specific pixel in slice S , common to every slice in C .

$$KF = \{(x, y) \mid KF \subset S, \quad \forall z\} \quad (4.14)$$

- The set of **relative points**, $KR(z)$, for slice number z is defined as the set of points obtained by adding each translation vector in $KR_{\text{translation}}$ to $VC_{\text{center}}(z)$. Let's denote the set of all possible translation vectors as $KR_{\text{translation}}$. These are 2-dimensional vectors in the axial plane of the CT scan.

$$KR_{\text{translation}} = \{(x, y) \mid x, y \in \mathbb{N}\} \quad (4.15)$$

The center position of the vena cava (which is explained later in this section) for a given slice z is denoted by $VC_{\text{center}}(z)$. This is a 2-dimensional point in the axial plane.

$$VC_{\text{center}} : \mathbb{N} \rightarrow \mathbb{N}^2 \subset S \quad (4.16)$$

This function maps a slice number (a natural number) to a 2-dimensional point (a pair of natural numbers). The set of relative points, $KR(z)$, for slice number z is then defined as the set of points obtained by adding each translation vector in $KR_{\text{translation}}$ to $VC_{\text{center}}(z)$.

$$KR(z) = \{(x, y) \mid (x, y) = VC_{\text{center}}(z) + (x', y'), (x', y') \in KR_{\text{translation}}\} \quad (4.17)$$

This means that a point (x, y) belongs to $KR(z)$ if it can be obtained by adding a translation vector (x', y') from $KR_{\text{translation}}$ to the center position of the vena cava, $VC_{\text{center}}(z)$.

Each semantic class has its own knowledge map. The set $K_a(z)$ for a given anatomical structure and slice number z is defined as the union of the set of relative points and the set of fixed points for that structure. This can be represented mathematically as:

$$K_a(z) = KR_a(z) \cup KF_a, \quad (4.18)$$

$$\forall a \in \{\text{background, vena cava, aorta, pulmonary arteries}\}$$

The relative points add a degree of freedom to the scoring process for the aorta and pulmonary arteries, which are easy to locate relative to the vena cava. The knowledge map provides a spatial reference that improves the accuracy of the results and facilitates the scoring of each semantic class. To score a superpixel with class membership f , we compute a similarity metric for each reference point in the knowledge map of label f , using the Gaussian Kernel function defined as:

$$Ker(p, r) = e^{-\gamma \|centroid(p) - r\|_2^2}, \quad (4.19)$$

where $p \in 1, \dots, n_superpix$ is the superpixel, $centroid(p) \in \mathbb{N}^2$ its coordinates in the axial plane, $r \in K$ (see eq. (4.18)) is the reference point, and $\gamma \in \mathbb{R}$ is a kernel weight used to adjust the sensitivity. We then take the maximum kernel obtained from all the reference points as the final score for the superpixel. This function gives a higher score to superpixels near a reference point on the knowledge map.

The determination of the points in the knowledge map was achieved through an iterative process of trial and error. By observing the output of the scoring function, we could assess the effectiveness of different choices and make decisions regarding the position of these reference points.

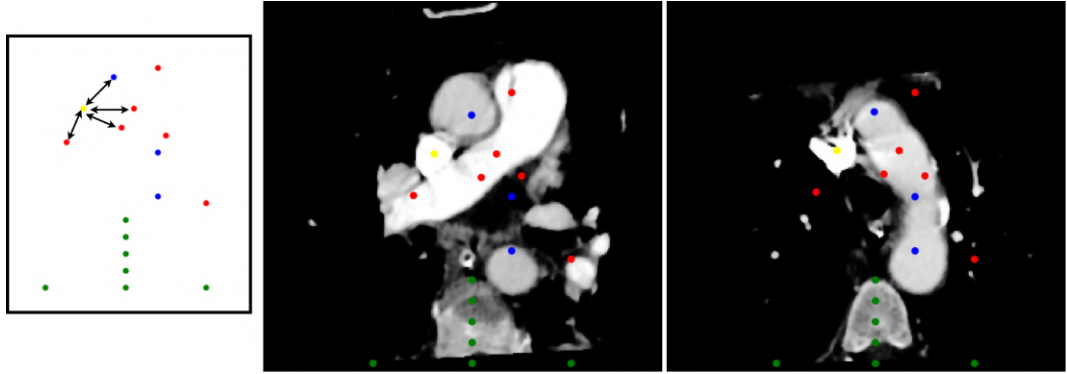


Figure 4.9: Mosaic of three images showing the application of the knowledge map to different slices. (a) A representation of the knowledge map, where yellow represents the vena cava, green represents the spine, blue represents the aorta, and red represents the pulmonary arteries. Points of the first type with a relative position to the vena cava are indicated by arrows, while other points are of the fixed type. (b) The knowledge map is applied to one slice. (c) The knowledge map is applied to another slice.

We can now use this knowledge map for the anatomical prior segmentation. Each class has its scoring function:

- **Background:** This class includes all superpixels that are not associated with the other three semantic classes, and it can contain substances or organs such as the spine, heart, bones, airways, muscles, and fat, among others. The scoring function for the Background class is defined as follows:

$$S_p(f_p = \text{background}) = \max(\sigma(I_p); K_{\text{background}}(z_p)), \quad (4.20)$$

where the first term is the alpha-beta-sigmoid function [20] of the intensity of the superpixel. The intensity is computed as the mean of all intensities within the region of the superpixel in eq. (4.4), and this term scores regions with lower intensities as background. The second term is a score obtained using the knowledge map of the spine, as described in the previous paragraph. z_p is the z index or the slice number that contains the superpixel p .

- **Vena cava:** The method used for scoring the vena cava class is based on the vena cava's known position at the top of the scan and its vertical-cylindrical shape to propagate information through slices (see fig. 4.11). To be more specific, we focus on the superior vena cava, which is located close to the pulmonary arteries in the trunk (see fig. 4.4). The compu-

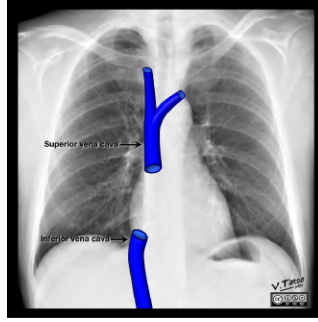


Figure 4.10: Superior and inferior vena cava in a front view of a CT scan - Source: Case courtesy of Vincent Tatco, Radiopaedia.org, rID: 46331

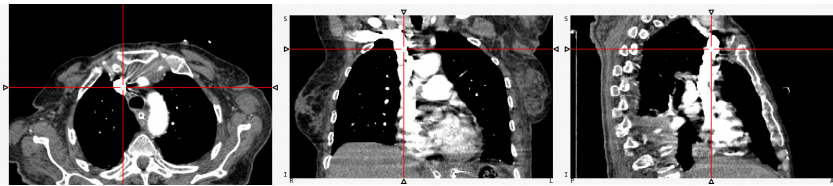


Figure 4.11: Red lines are centered inside the vena cava in the axial plane in (a). The same lines are also shown in (b) and (c), highlighting the vertical tubular structure of the vena cava.

tation of scores for each slice is performed iteratively, starting from the highest slice index,

$i = \max(Z)$, where Z is the set of all z-coordinates in the scan, and then decrementing i by 1 for each iteration. The score for a given superpixel p in slice i , assigned to the vena cava ($f_p = \text{vena cava}$), is calculated using the following function:

$$S_p(f_p = \text{vena cava}) = I_p \cdot \text{Ker}(\text{Centroid}(p), VC_{\text{center},i}) \cdot Z_{\text{range},i}, \quad \forall p \in S(i) \quad (4.21)$$

In this equation, I_p is the intensity of superpixel p , $\text{Ker}(\cdot, \cdot)$ is the Gaussian Kernel function that measures similarity, and $Z_{\text{range},i}$ is a weight used to decrease the score in vertical regions where the vena cava is unlikely to be found (e.g., the lower part of the thorax). The reference point for the next slice, $VC_{\text{center},i-1}$, is then computed using the superpixels' scores and the current slice's reference point. This is achieved with the following function:

$$VC_{\text{center},i-1} = \frac{w_1 \cdot VC_{\text{center},\text{init}} + w_2 \cdot VC_{\text{center},i} + \sum_{p \in S(i)} S_p(\text{vena cava}) \cdot \text{Centroid}(p)}{w_1 + w_2 + \sum_{p \in S(i)} S_p(\text{vena cava})}, \quad (4.22)$$

In this equation, $VC_{\text{center},\text{init}}$ is a reference point based on a predefined approximate position of the vena cava common to any patient and slice. The weights w_1 and w_2 are used to adjust the influence of this initial reference point and the result from the current slice, respectively. The term $\text{Centroid}(p)$ represents the coordinates of the centroid of superpixel p . This methodology ensures that the reference point is updated to prevent convergence to other high-intensity regions during the score propagation process. The iterative process is illustrated in fig. 4.12.

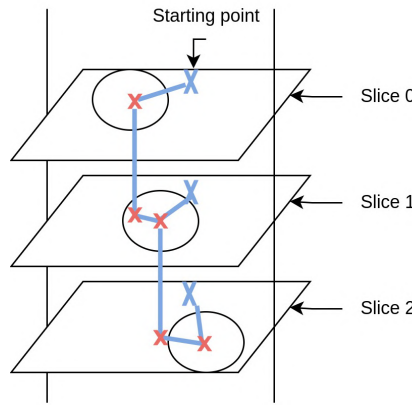


Figure 4.12: Iterative process of computing vena cava per-slice center.

- **Aorta:** The scoring function for the aorta relies on the knowledge map approach. This scoring function fully exploits two types of reference points in the knowledge map. Specif-

ically, we use reference points relative to the vena cava to score the ascending aorta. We use fixed reference points for the descending aorta, while for the aortic arch, we utilize both types of points. The ascending aorta near the vena cava allows for more precise scoring using the relatively positioned points. Refer to fig. 4.13 for localizing these anatomical parts.

$$S_p(f_p = \text{aorta}) = I_p \cdot Z_{\text{range,aorta}}(z_p) \cdot \max_{k \in K_{\text{aorta}}} (Ker(\text{Centroid}(p), k)) \quad (4.23)$$

In this equation, K_{aorta} represents the set of reference points for the aorta.

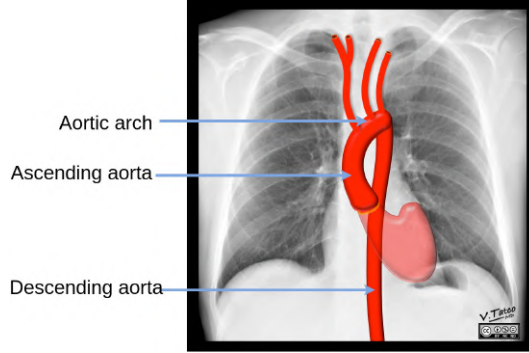


Figure 4.13: Annotated aorta in a frontal view - Source of the image without annotations: *Case courtesy of Vincent Tatco, Radiopaedia.org, rID: 46331.*

- **Pulmonary arteries:** Finally, we evaluate the left-right pulmonary arteries and pulmonary trunk, which are crucial for detecting pulmonary embolus (see fig. 4.4). To accomplish this, we rely on the knowledge map generated from the reference points and the lung mask to identify smaller pulmonary arteries distributed throughout the lungs. Scoring this class is particularly important, as it is a sub-region for detecting embolus. The scoring function for the pulmonary arteries introduces a new term, $\text{InLungs}(p)$, which provides a binary indication of whether the centroid of a given superpixel p lies within the lung mask. This lung mask is defined by the CT scan region corresponding to the lungs, as detailed in previous sections of this work. Mathematically, we can define this function as follows:

$$\text{InLungs}(p) : \mathcal{P} \rightarrow \{0, 1\} \quad (4.24)$$

where $\mathcal{P}$ is the set of all superpixels. The function is defined as:

$$\text{InLungs}(p) = \begin{cases} 1, & \text{if the centroid of } p \text{ lies within the lung mask,} \\ 0, & \text{otherwise.} \end{cases} \quad (4.25)$$

In this manner, $\text{InLungs}(p)$ effectively differentiates superpixels within the lung mask from those outside, providing a valuable discriminative feature for the scoring function. The scoring function is thus defined as:

$$S_p(f_p = \text{PA}) = I_p \cdot Z_{\text{range, pa}}(z_p) \cdot \left(\max_{k \in K_{pa}} (\text{Ker}(\text{Centroid}(p), k)) + \text{InLungs}(p) \right), \quad (4.26)$$

INTERMEDIATE RESULTS

Now that we have defined the score function for each class of the prior-knowledge segmentation task (i.e., pulmonary arteries, vena cava, aorta, and background), we can observe the results obtained for each class on a scan.

To facilitate the interpretation of the computed scores, we transform the superpixel-based scores into pixel-based scores. This transformation allows us to generate a colormap where the color corresponding to the score of a superpixel is uniformly assigned to each pixel within that superpixel.

Let's denote the set of all superpixels as $P \in 1, 2, \dots, n_superpix$, and let $S_p(f_p)$ be the score when assigning label f to superpixel p . The score of a single pixel at coordinates (x, y, z) , denoted as $s[x, y, z]$, is defined as the score of the superpixel that contains that pixel. This can be formulated mathematically as follows:

$$s_f[x, y, z] = S_p(f), \quad \forall (x, y, z) \in C \text{ such that } p = L[x, y, z] \quad (4.27)$$

In this expression, C is the set of all pixels in the scan, and $L[x, y, z]$ refers to the superpixel label containing the pixel at (x, y, z) . In this way, each pixel inherits the score of the superpixel to which it belongs, enabling a detailed visualization of the superpixel scores.

In fig. 4.14, we can observe the achievement of our objective: assigning a high score to the between-vessels tissues and the spine. The between-vessels area was identified based on the intensities, while the spine scoring was accomplished by utilizing the knowledge map. Including the knowledge map was crucial, as it enabled us to score the spine despite its high-intensity pixels, which would not have activated the function relying on low-intensity superpixels in its absence. This demonstrates the effectiveness of incorporating domain-specific knowledge to improve the segmentation pipeline.

We can further analyze the scores obtained for the superpixels where pulmonary embolus (PE) is known to be present based on the ground truth. For instance, in fig. 4.15, it can be observed that

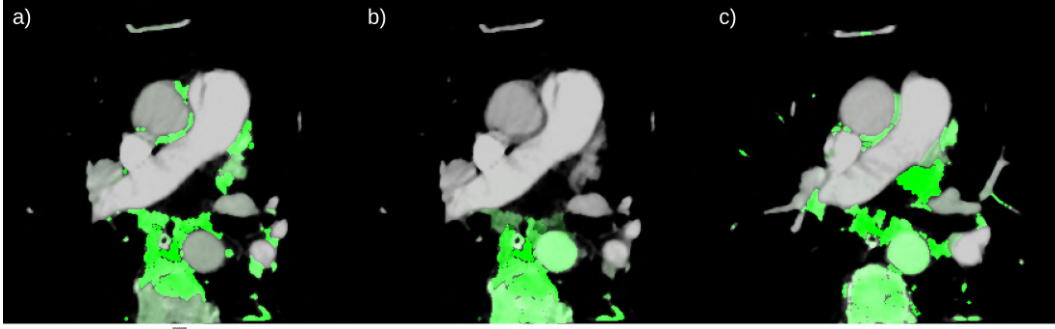


Figure 4.14: Scoring the background semantic label on superpixels by taking the maximum value (c) between an intensity-based score (a) and a spinal-knowledge map score (b). The green color means higher scores.

both PE regions in the slices obtained high scores for the background class. This result is expected, as these regions appear darker due to the presence of blood clots, thus activating the low-intensity filter in the scoring function. Note that this result is significant and will play a crucial role in the post-processing step of the pipeline when identifying PE candidates.

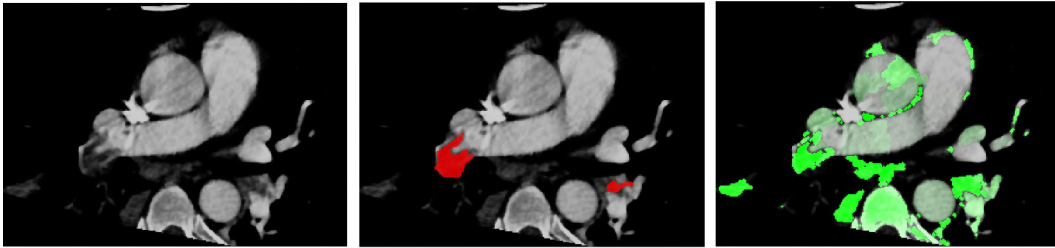


Figure 4.15: Pat01-slice109. From left to right: original slice; overlay with ground truth PE in red; background scores

An analysis of the results obtained for the aorta and pulmonary artery semantic classes is conducted together due to the challenge of separating these classes caused by low-contrast boundaries, specifically in the case of touching vessels. In fig. 4.16, the scores obtained for these two classes overlap significantly in the middle part of the pulmonary artery, precisely at the center of the image. However, in other regions, we observe that the highest score corresponds to the true semantic class. The issue encountered in the central part of the pulmonary artery could be resolved by incorporating the graph smoothing term of the alpha-expansion algorithm. By encouraging label consistency among neighboring regions, the algorithm can address the challenges posed by low-contrast boundaries in distinguishing the aorta and pulmonary artery classes.

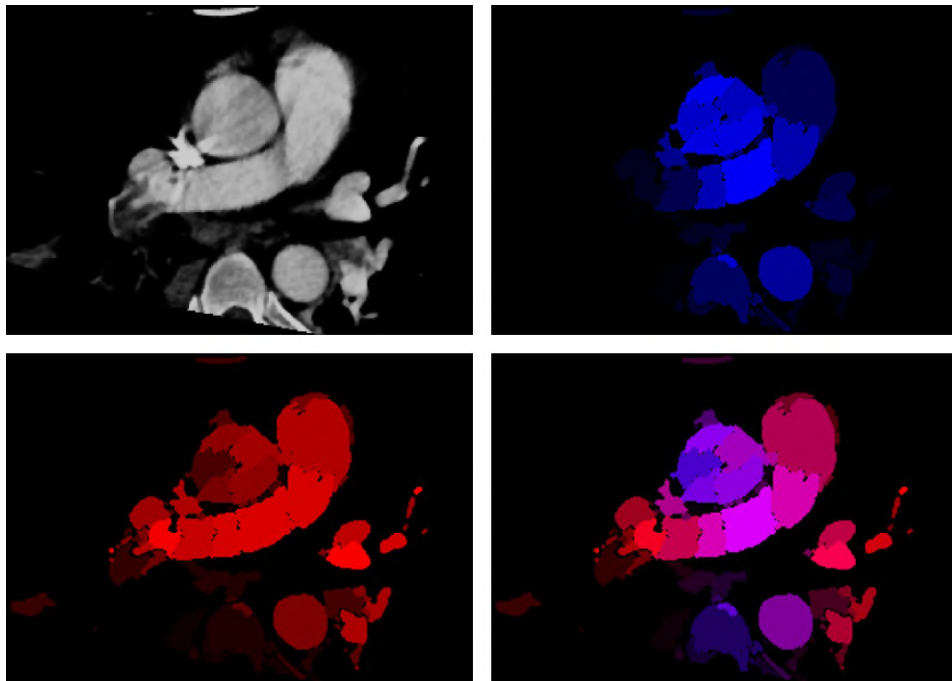


Figure 4.16: Original slice and scores obtained for the aorta (blue) and the pulmonary arteries (red). The last figure is a color-additive representation of both scores.

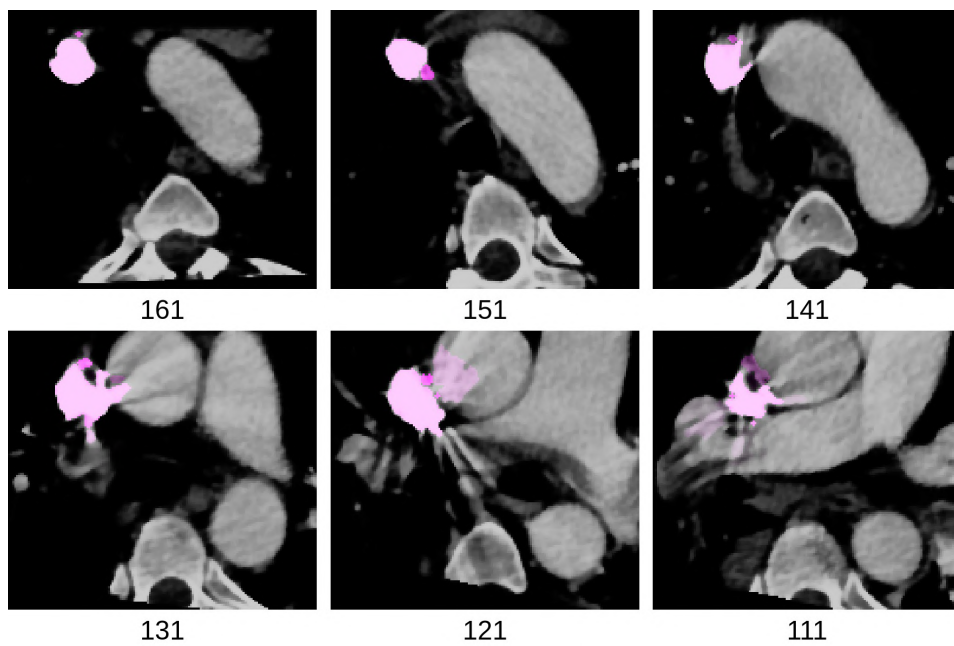


Figure 4.17: Pat01 - Vena cava scoring results for different slice numbers.

A sample result obtained through the iterative process for vena cava computation is depicted in fig. 4.17. Starting from the top slice, it is evident that the reference point $VC_{initial}$ serves as a suitable starting point, enabling successful segmentation of the vena cava in the initial slices. Subsequently, the iterative process, which incorporates the positional information of the vena cava from the previous slice, facilitates the propagation of the segmentation across the subsequent slices. It is noteworthy that despite the presence of low-contrast boundaries, the obtained delimitation of the vena cava is satisfactory.

To gain a comprehensive view of the scoring function within a broader context, we present the collective scores using an RGBA overlay, as depicted in fig. 4.18. Specifically, fig. 4.18 consists of a mosaic comprising six images, where four of them represent the scoring results of superpixels from a single slice. The images encompass the original scan, scores for the vena cava superpixels depicted in yellow, scores for the aorta superpixels depicted in blue, scores for the pulmonary arteries superpixels depicted in red, scores for the background superpixels depicted in green, and a visual representation of the scores achieved through additive blending, where the RGB values are combined. This final image provides a coherent anatomical representation; however, some uncertainty arises when superpixels exhibit similar scores for both the aorta and pulmonary arteries. These areas of ambiguity indicate regions where the scoring function struggles to differentiate between the two classes effectively.

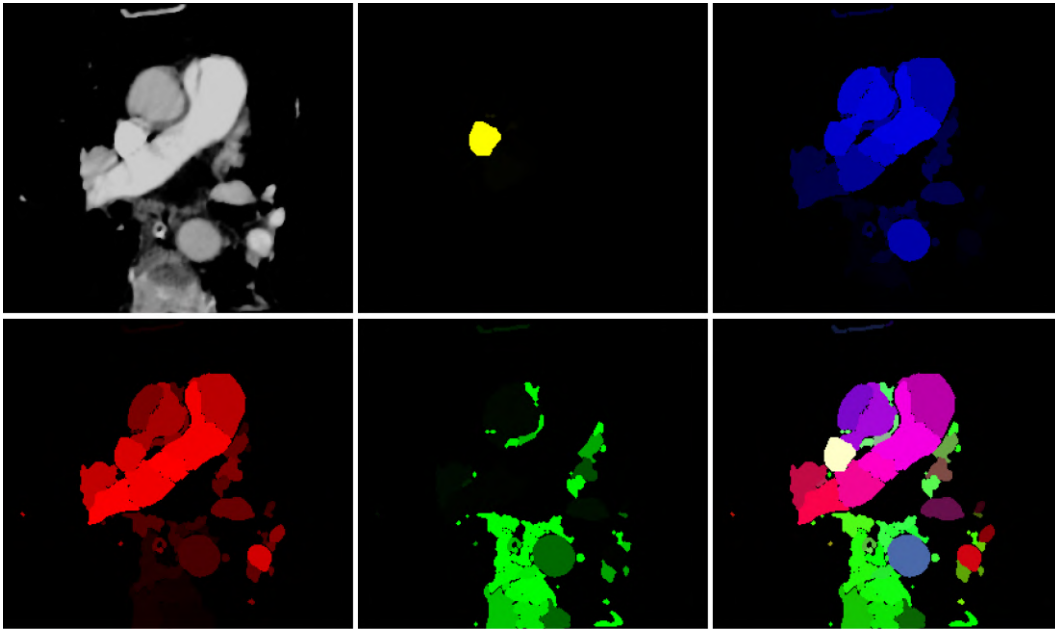


Figure 4.18: Scoring results for each semantic class

4.3.5 GRAPH CONSTRUCTION

Having obtained the score functions (and consequently, the cost functions) for each superpixel and anatomical-semantic class, we can now proceed to the second phase of energy minimization, which involves addressing the challenge of distinguishing between the aorta and the pulmonary arteries. This is achieved by implementing the smoothness term to alleviate this issue.

To achieve this, we need the inclusion of a graph structure to represent the spatial arrangement of the superpixels. The edges connecting the nodes in the graph play a crucial role in facilitating the propagation of label information. Through the graph-cut algorithm, the graph is partitioned into a fixed number of subgraphs, each corresponding to a specific label.

The construction of the graph relies on establishing adjacency connections between superpixels. These connections are of two types: horizontal and vertical.

- The **horizontal connections** are formed between superpixels that share a boundary within the same slice. To mathematically define the term "share a boundary within a slice", we need to consider adjacent pixels within the same slice that are labeled differently. If we define the function $N(x, y)$ as the set of direct neighbors of the pixel at position (x, y) in the axial plane (i.e., above, below, left, or right), then two superpixels p and q share a boundary within slice i if and only if there exist pixels (x_1, y_1) and (x_2, y_2) such that (x_1, y_1, i) is in superpixel p , (x_2, y_2, i) is in superpixel q , and (x_2, y_2) is in $N(x_1, y_1)$. If we denote the set of all superpixels in a given slice i as $S(i)$, the set of all pairs of horizontally adjacent superpixels, denoted $\mathcal{E}_H$, can be defined as:

$$\begin{aligned} \mathcal{E}_H = \{ (p, q) | p, q \in S(i) \text{ for some } i, \text{ and} \\ \exists (x_1, y_1) \in C \text{ and } \exists (x_2, y_2) \in N(x_1, y_1) \text{ such that} \\ L[x_1, y_1, i] = p \text{ and } L[x_2, y_2, i] = q \} \end{aligned} \quad (4.28)$$

- **Vertical connections**, on the other hand, are formed between superpixels on different but adjacent slices and overlap when projected onto the axial plane. We can define an overlap function $\mathcal{O}(p, q)$ that returns true if superpixels p and q overlap, and false otherwise:

$$\mathcal{O}(p, q) : \mathcal{P} \times \mathcal{P} \rightarrow \{0, 1\}$$

$$\mathcal{O}(p, q) = \begin{cases} 1, & \text{if } \exists (x, y, i), (x, y, j) \in C \text{ such that} \\ & L[x, y, i] = p, L[x, y, j] = q, \text{ and } |i - j| = 1 \\ 0, & \text{otherwise} \end{cases}$$

With this definition, the set of vertically adjacent superpixels $\mathcal{E}_V$ can be defined as:

$$\mathcal{E}_V = \{(p, q) | p \in S(i), q \in S(i + 1) \text{ for some } i, \text{ and } \mathcal{O}(p, q) = 1\} \quad (4.29)$$

These rules ensure a graph's construction accurately represents the spatial structure and relationships between superpixels in the 3D scan. Note that superpixels that are more than one slice apart are not considered adjacent and, thus, are not linked.

Adherence to these rules generates an unweighted graph, which we aim to transform into a weighted graph to allow the application of the α -expansion and enable control over the graph cuts.

To establish an edge weight between two superpixels, we introduce a function belonging to the class of semi-metrics. The weight assigned to the edge connecting two superpixels is influenced by the percentage of overlap (for vertical links) and the scores obtained (for horizontal links).

As discussed in section 4.3.4, we deduce that the scoring functions are sufficiently reliable for superpixels in the background and vena cava regions. In contrast, the scores obtained for pixels belonging to the aorta or pulmonary artery regions have high uncertainty. As a result, we could potentially reduce the edge weights for superpixels belonging to the vena cava and background to allow a graph cut with their neighbors. To achieve this, we need to define a function that takes into account the following physical properties:

- An edge linked to a high background or vena cava score should carry a low weight.
- An edge between two superpixels with high aorta or pulmonary artery scores should carry a high weight to enhance the smoothing effect of the alpha-expansion algorithm.

- Vertical edges should carry more weight than horizontal edges, as adjacent superpixels on different slices are likelier to belong to the same structure. Thus, their connection should be given more weight.

With these properties in mind, we can proceed to define both weight functions:

- **The weight of a vertical edge** $(p, q) \in E_V$, connecting superpixels p and q , is calculated using the following formula:

$$W(p, q) = \frac{|p \cap q|}{\min(|p|, |q|)}, \quad \forall (p, q) \in E_V \quad (4.30)$$

Here, $|\cdot|$ represents the cardinality (number of pixels) of a superpixel, and $p \cap q$ denotes the intersection of superpixels p and q in the axial plane, i.e., the overlapping pixels. This weight calculation aimed to assign higher weights to superpixels with larger intersections.

- For **the horizontal edge weight**, it is calculated based on the semantic scores of each superpixel. An edge connecting two superpixels with high scores for the aorta or pulmonary artery should carry a high weight, thereby enhancing the smoothing effect of the alpha-expansion algorithm. To achieve this, we define a confidence function for the background and vena cava classes:

$$\begin{aligned} \text{Conf}_{\text{bg}}(p) &= \min \left(S_p(f_p = \text{bg}) - S_p(f_p = \text{ao}), \right. \\ &\quad \left. S_p(f_p = \text{bg}) - S_p(f_p = \text{pa}) \right) \\ \text{Conf}_{\text{vc}}(p) &= \min \left(S_p(f_p = \text{vc}) - S_p(f_p = \text{ao}), \right. \\ &\quad \left. S_p(f_p = \text{vc}) - S_p(f_p = \text{pa}) \right), \end{aligned} \quad (4.31)$$

Next, for each horizontal edge connecting superpixels p and q , we define $M(p, q)$, the maximum confidence of the superpixels connected by the edge:

$$\begin{aligned} M(p, q) = M(q, p) &= \max \left(\text{Conf}_{\text{bg}}(p), \text{Conf}_{\text{bg}}(q), \right. \\ &\quad \left. \text{Conf}_{\text{vc}}(p), \text{Conf}_{\text{vc}}(q) \right), \\ &\quad \forall (p, q) \in E_H \end{aligned}$$

However, since we desire a low weight for high-confidence edges in line with physical properties, we define the weight function for horizontal edges:

$$W(p, q) = 1 - \frac{1}{1 + \alpha e^{-\beta M(p, q)}}, \quad \forall (p, q) \in E_H \quad (4.32)$$

where α and β control the function's shape, we can map high confidence to low edge weight.

4.3.6 ALPHA-EXPANSION

To minimize the energy function eq. (4.6), we will follow the graph-cuts method proposed by Boykov *et al.* in [7]. Two algorithms that use graph cuts to compute a local minimum are presented:

- α - β swap: for a pair of labels α, β , this move exchanges the labels between an arbitrary set of pixels labeled α and another arbitrary set labeled β .
- α -expansion: for a label α , this move assigns the label α to an arbitrary set of pixels.

This thesis will focus on the second method, which requires E_{smooth} to be a metric. We know that the solution is within a known factor of the global minimum. This method was used in Browet *et al.*'s cells segmentation paper [8].

Before delving into this concept, let's define the appropriate smoothing function to be incorporated in the alpha expansion based on the weight functions from the previous section. Specifically, we must ensure zero weight between adjacent superpixels with identical labels. We introduce a new function that takes into account label dependencies:

$$V_{\{p, q\}}(f_p, f_q) = W_{\{p, q\}}(p, q)(1 - \delta(f_p, f_q)). \quad (4.33)$$

This function considers the connection weight $W_{p, q}$ and the labels assigned to superpixels p and q . We use the Kronecker delta function δ to compare the labels assigned to the superpixels p and q . By incorporating label dependency, we ensure that neighboring superpixels with the same label have zero weight, smoothing the variation of labels between adjacent regions.

A function V is a metric if it satisfies the following properties for all α, β, γ in the domain of V :

$$V(\alpha, \beta) = V(\beta, \alpha) \quad (\text{Symmetry})$$

$$V(\alpha, \beta) \geq 0 \quad (\text{Non-negativity})$$

$$V(\alpha, \beta) = 0 \iff \alpha = \beta \quad (\text{Identity of indiscernibles})$$

$$V(\alpha, \beta) \leq V(\alpha, \gamma) + V(\gamma, \beta) \quad (\text{Triangle inequality})$$

To prove that $V_{p,q}(f_p, f_q)$ is a metric, we need to show that it satisfies these properties.

- Symmetry: This is immediately evident from the definition of $V_{p,q}(f_p, f_q)$, as swapping f_p and f_q does not change the value of the function.
- Non-negativity: Since $W_{p,q}(p, q)$ is a weight and thus non-negative, and $1 - \delta(f_p, f_q)$ is either 0 or 1, the product is always non-negative.
- Identity of indiscernibles: If $f_p = f_q$, then $\delta(f_p, f_q) = 1$ and thus $V_{p,q}(f_p, f_q) = 0$. Conversely, if $V_{p,q}(f_p, f_q) = 0$, this implies that $f_p = f_q$.
- Triangle inequality: This requires a more detailed proof that takes into account the specific form of $V_{p,q}(f_p, f_q)$, the internal weights $W(p, q)$ and the semantics of the labels f_p, f_q . As we use the same construction of weights as in [8] for the horizontal edges and label-independent weights for vertical edges, we assume based on [8] that this function validates the triangular inequality.

In fig. 4.19, we visually represent the graph structure before applying the alpha-expansion algorithm using a toy example. The graph is constructed with edges representing pixel-to-pixel connections, weights assigned according to the function $W(p, q)$, and pixel-to-label connections, weights determined by the data fidelity term $D_p(label)$.

In fig. 4.20, we examine the same toy example but during the execution of the α -expansion algorithm. The graph has one less pixel-to-pixel edge when both pixels are labeled with the same class. Conversely, when the pixels have different labels, the pixel-to-pixel edges are retained. This behavior is influenced by the smoothing function $V_{p,q}(f_p, f_q)$.

Both figures aim to illustrate the alpha-expansion algorithm's effects and the graph structure's dynamics. The purpose is to demonstrate how the graph's connectivity evolves during optimization.

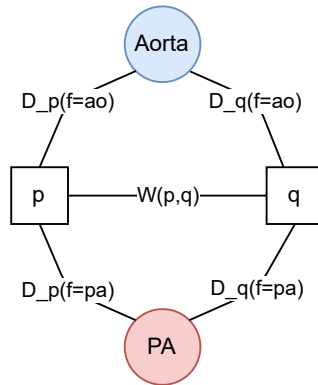


Figure 4.19: Toy Example of the Weights for Two-Pixels and Two-Classes

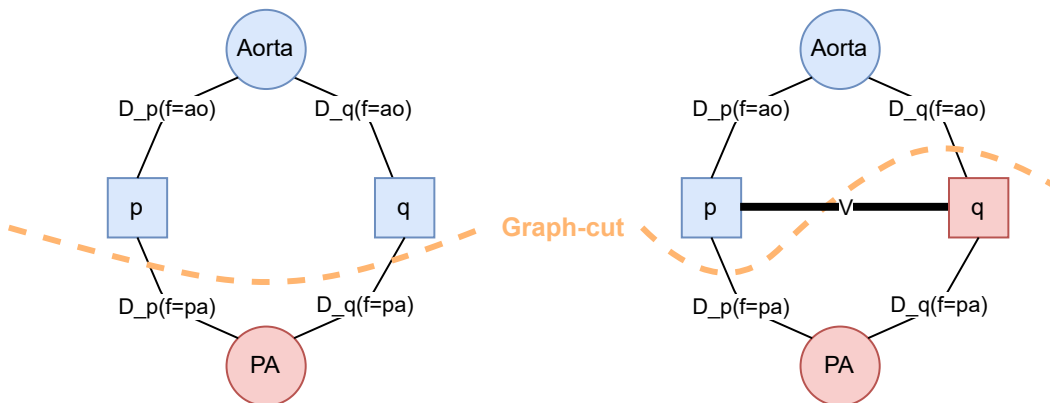


Figure 4.20: Toy Example for Graph-Cuts with Two-Pixels and Two-Classes

Here is a description of the α -expansion:

Algorithm 1 Alpha-Expansion Algorithm

```

1: Initialize arbitrary labeling  $f$ 
2: Set  $success := 0$ 
3: for each label  $\alpha$  in  $L = \{\text{vena cava, aorta, pulmonary arteries, background}\}$  do
4:   Find  $\hat{f} = \arg \min E(f')$  among  $f'$  within one  $\alpha$ -expansion of  $f$ 
5:   if  $E(\hat{f}) < E(f)$  then
6:     Set  $f := \hat{f}$  and  $success := 1$ 
7:   end if
8: end for
9: if  $success = 1$  then
10:   Go to step 2
11: end if
12: return  $f$ 

```

In the context of this algorithm, the α -expansion move (line 4) allows for the reassignment of an arbitrary set of pixels to the label α . This move expands the pixels assigned to α and can help improve the overall energy function $E(f)$. The algorithm iteratively applies these α -expansion moves for each label in L , updating the labeling f whenever a move decreases the energy. The process continues until no α -expansion move can further decrease the energy.

From this point forward, the class attributed to superpixel p is given by $A(p)$.

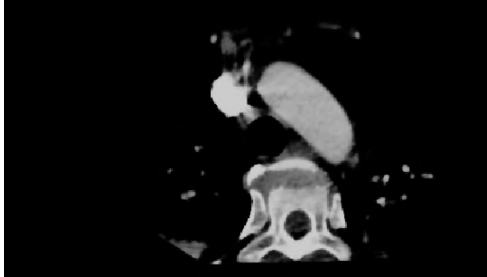
INTERMEDIATE RESULTS

In the previous sections, we computed the labeling costs and constructed a weighted graph, which serves as input for the alpha expansion algorithm. By executing this algorithm, we obtain a segmentation of the background, aorta, vena cava, and pulmonary arteries. To illustrate the effectiveness of our approach, we present sample slices obtained from two patients in fig. 4.21 and fig. 4.22.

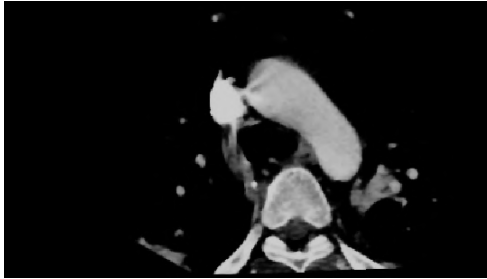
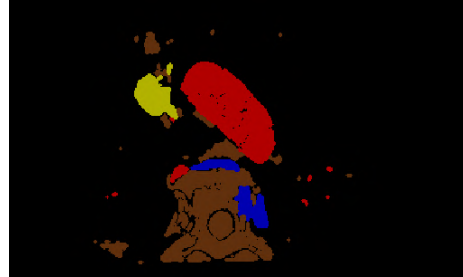
The algorithm considers physical properties and label dependencies by leveraging the weighted graph. This leads to a smooth and accurate segmentation of the four regions of interest, with clear boundaries delineating each region. This can be observed across all the images of patients 1 and 2, particularly in slice 100 of patient 2, where the segmentation is notably precise.

However, in the case of patient 1, we encountered some misclassifications where spine regions were erroneously labeled as the aorta. Additionally, there is a discrepancy in the semantical interpretation of the same object between the topmost and second topmost presented slices. The

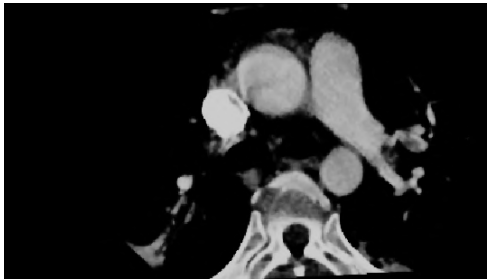
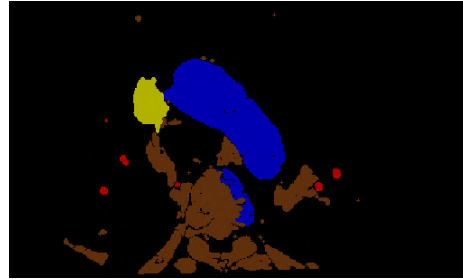
second slice accurately labels it as the aorta, while the topmost does not. To address these issues, improvements can be made to the knowledge map or by fine-tuning the weights of the vertical edges, ensuring that these two scans segment the object into the same semantic class, leading to an expected segmentation within the aorta class.



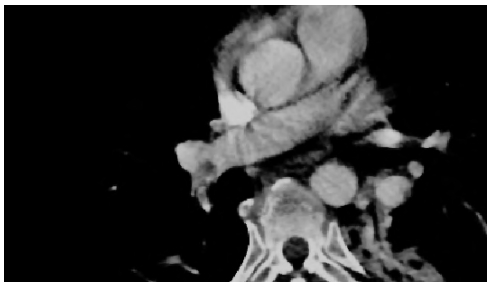
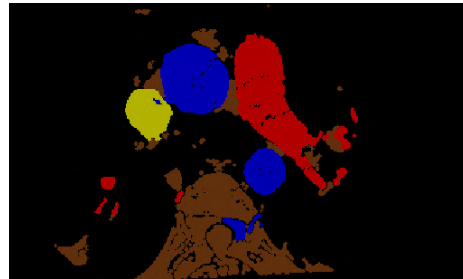
Patient 1 - Slice 156



Patient 1 - Slice 145



Patient 1 - Slice 136



Patient 1 - Slice 110

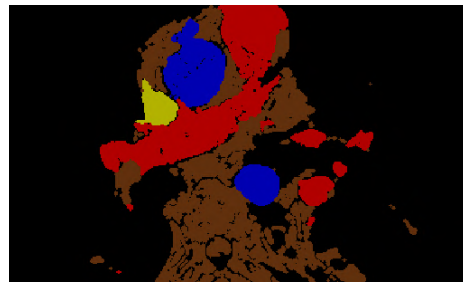
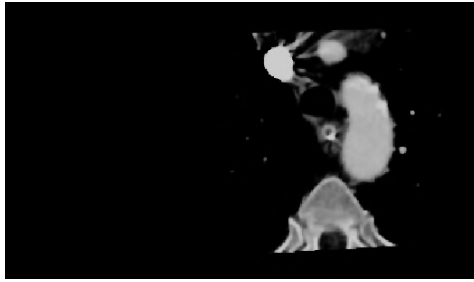
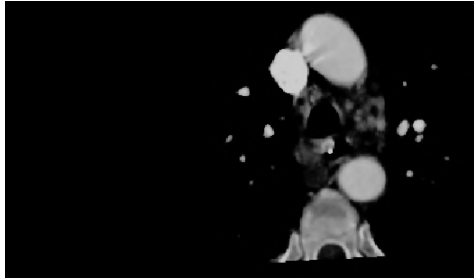
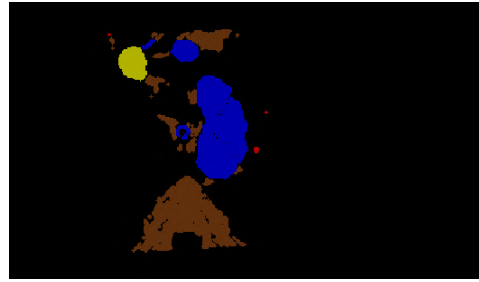


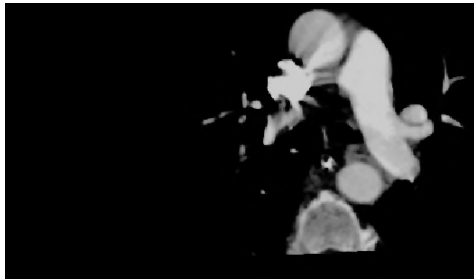
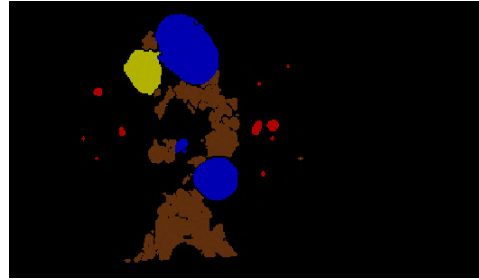
Figure 4.21: Segmentation results for scan 1 and 4 selected slices. Colormap: Background in brown and brown, pulmonary arteries in red, vena cava in yellow, aorta in blue



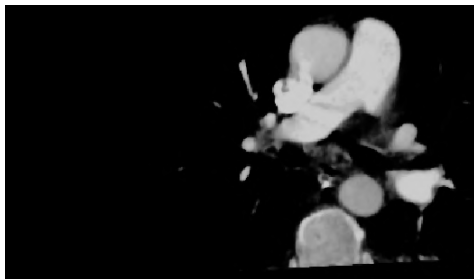
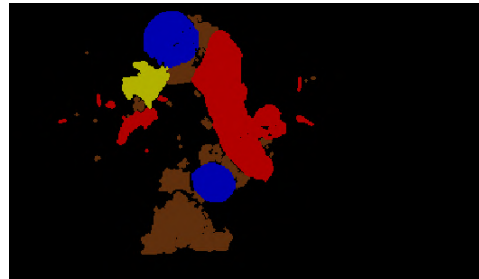
Patient 2 - Slice 145



Patient 2 - Slice 122



Patient 2 - Slice



Patient 2 - Slice 100

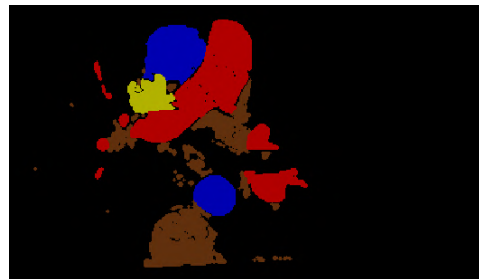


Figure 4.22: Segmentation results for scan 2 and 4 selected slices. Colormap: Background in brown and brown, pulmonary arteries in red, vena cava in yellow, aorta in blue

4.3.7 PE LOCALIZATION USING PRIOR-ANATOMICAL-SEGMENTATION

In this section, we will conclude the process of generating labels for pulmonary embolus (PE) in a CT scan. Building upon the previous sections, where we obtained superpixels, anatomical

scores, and an anatomical segmentation of the scan, we will now leverage these extracted features to segment PE candidates. These candidates will serve as generated labels for the second stage of the framework, although this second stage is not extensively studied in this thesis.

Blood clots within the CT scan exhibit lower pixel intensities compared to other pixels within the pulmonary arteries. Consequently, we can anticipate identifying them by examining the intensities of superpixels within the pulmonary arteries mask derived from the anatomical-prior segmentation. However, it is important to note that these low-intensity pixels may also be erroneously labeled as part of the background class. As a result, relying solely on the pulmonary arteries mask may not be sufficient for detecting PE candidates.

To address this challenge, we will explore two distinct approaches for generating PE candidates: by examining the darker regions within the pulmonary arteries mask and by considering mislabeled background regions. Each approach will be discussed in the following subsections.

SEARCHING IN PULMONARY ARTERIES MASK

This approach focuses on regions in the mask with an average intensity lower than a set threshold, as emboli are expected to manifest as darker regions in the CT scan.

Mathematically, this corresponds to identifying a set of nodes, $P_{\text{suspicious}}$, within the graph G , such that the mean intensity of each node $p \in P_{\text{suspicious}}$ is below the defined threshold. Formally, we define:

$$P_{\text{suspicious}} = \{p | I_p < \text{threshold}, A(p) = \text{pa}\} \quad (4.34)$$

where I_p denotes the mean intensity of node p and $A(p)$ corresponds to the assigned label of node p (in this case, we are interested in nodes labeled as pulmonary artery 'pa'). A subgraph G_{sus} of the original graph G is then constructed, consisting only of the nodes in $N_{\text{suspicious}}$.

The connected components of G_{sus} are then examined, representing potential clusters of nodes that form emboli. Small connected components with fewer than 3 nodes are ignored to avoid false positives due to noise.

The output is a list N_{out} , which contains the node IDs of the remaining nodes in the connected components of G_{sus} , representing the potential locations of emboli within the pulmonary artery mask.

MISLABELED BACKGROUND REGIONS

Unlike the previous approach, this approach focuses on regions that are initially classified as background but may contain emboli. The identification process involves several validity tests to narrow down the suspicious regions.

Mathematically, the process starts by identifying a set of nodes N_{filter1} in the graph G , such that each node p in N_{filter1} satisfies three conditions:

- The label of the node p is 'background', formally, $A(p) = \text{'background'}$
- There exists at least one neighboring node of p that is labeled as PA
- None of the neighboring nodes of p are labeled as CV or 'aorta' (AO)

This set of nodes is formally defined as:

$$\begin{aligned} N_{\text{filter1}} = \{ & p \mid A(p) = \text{'background'}, \\ & \exists q \in N(p) : A(q) = \text{'pa'}, \\ & \nexists q \in N(p) : A(q) \in \{\text{'cv'}, \text{'ao'}\} \} \end{aligned}$$

A subgraph G_{filter1} of G is then constructed, consisting only of the nodes in N_{filter1} . The connected components of G_{filter1} are then examined. Small connected components with fewer than 3 nodes are ignored, as well as components that span only one slice in the volumetric data, as these are unlikely to represent emboli.

An additional filtering step is applied, examining the geometric properties of the remaining connected components. Only components with a total area between 2000 and 15000 pixels and an extent greater than 0.15 are considered potential emboli. These values were chosen by exploring typical values of PE in the ground truth.

The final output of this function is a list N_{filter2} , containing the node IDs of the remaining nodes in the filtered connected components of G_{filter1} , representing the potential locations of emboli within the background mask.

4.4 EVALUATION METRICS

We have chosen two types of evaluation metrics to evaluate and analyze the results obtained from the pipeline. The first set of metrics is based on [16] and enables us to assess the algorithm's performance detecting pulmonary embolus (PE) within the extensive and large CT scan. These metrics

also allow us to evaluate the PE candidates incorrectly proposed by the segmentation pipeline. The metrics are defined as follows:

- *True Positive (TP)*: True positives are defined as PE candidates provided by the algorithm that fall within the ground truth emboli. Only one is considered if several PE candidates fall within the same embolus.
- *False Positive (FP)*: False positives are PE candidates outside the true emboli.
- *False Negative (FN)*: False negatives represent true embolus without any corresponding PE candidates.
- *Sensitivity or Recall*: Sensitivity is defined as the number of embolus detected divided by the total number of embolus in the test scans. Mathematically, it can be represented as:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (4.35)$$

- *Positive Predictive Value (PPV) or Precision*: The probability that a positive PE candidate result is actually indicative of a PE.

$$\text{PPV} = \frac{TP}{TP + FP} \quad (4.36)$$

The second set of metrics is used to evaluate the segmentation itself pixel-wise. We employ Intersection-Over-Union (IoU) and Dice coefficient (DICE), which are commonly used metrics for segmentation tasks on images with unbalanced data (as is the case with PE, as it constitutes a small portion of the scan). These metrics can be defined as follows:

- *Intersection Over Union (IoU)*: The IoU of a predicted segmentation relative to ground truth is defined as:

$$\text{IoU} = \frac{PRED \cap GT}{PRED \cup GT}, \quad PRED, GT \subset C \quad (4.37)$$

where $PRED$ represents the predicted segmentation, GT represents the ground truth segmentation, and both are subsets of $\mathcal{C}$, representing the set of all pixels in the CT scan.

- *DICE*: The DICE coefficient is defined as:

$$\text{DICE} = \frac{2 \cdot |PRED \cap GT|}{|PRED| + |GT|}, \quad PRED, GT \subset C \quad (4.38)$$

5 RESULTS

We only obtained results for the label-generation stage of the framework, and thus only this classical segmentation algorithm will be analyzed. The following chapter did not analyze patients without PE in the ground truth (i.e., Patient 24 and Patient 32 [31]).

5.1 QUANTITATIVE RESULTS

Based on the different metrics depicted in section 4.4, we evaluated the performance of the label generation algorithm. The algorithm aimed to detect and segment PE regions in CT scans and provide a pool of candidate labels for the subsequent deep-learning method.

5.1.1 DETECTION

Evaluation is first made in terms of detection, meaning that we observe how the algorithm can segment PE pixels in regions where PE is present in the ground truth of the scan. To measure its effectiveness, several metrics are employed, including sensitivity, true positives, false positives, and false negatives. Since the algorithm searches for PE superpixel candidates in both the pulmonary arteries mask (pathological case) and the background mask, the metrics are computed based on the use of each search mask individually and when both masks are utilized together (overall case).

In the following tables, we provide a detailed analysis of these specific searching cases: the tables (Table 5.1, Table 5.2, and Table 5.3) showcases the best and worst performance outcomes for each scenario.

In the context of overall PE detection (Table 5.1), the average sensitivity across all cases was 0.67, with an average of 64 false positives per scan, leading to a positive predictive value (PPV) of 0.04. Notable cases include PAT007, which demonstrated the best sensitivity with a perfect score of 1.00, and PAT017, which had the highest number of false positives (111), constituting the worst case in this category.

When examining pathological region PE detection (Table 5.2), the sensitivity significantly decreased to 0.17 on average, while the average number of false positives per scan decreased to 7.

Table 5.1: Overall PE detection best and worst cases

Case	Patient ID	Sens	FP/S	PPV
All Cases	/	0.67	64	0.04
Best Sensitivity Scan	PAT007	1.00	49	0.04
Best FP Scan	PAT028	0.00	28	0.00
Worst Sensitivity Scan	PAT002	0.00	65	0.00
Worst FP Scan	PAT017	0.67	111	0.02

This resulted in a slightly improved PPV of 0.10. PAT005 had the best sensitivity (0.50) in this category, while PAT012 showed the highest number of false positives (17).

Table 5.2: Pathological region PE detection best and worst cases

Case	Patient ID	Sens	FP/S	PPV
All Cases	/	0.17	7	0.10
Best Sensitivity Scan	PAT005	0.50	8	0.09
Best FP Scan	PAT031	0.13	1	0.50
Worst Sensitivity Scan	PAT021	0.00	2	0.00
Worst FP Scan	PAT012	0.00	17	0.00

For the background region PE detection (Table 5.3), the sensitivity reached the overall case at 0.66, and the false positives per scan were 62 on average, leading to a PPV of 0.04. The case with the best sensitivity was PAT026, while the case with the highest false positives was again PAT017.

Table 5.3: Background region PE detection best and worst cases

Case	Patient ID	Sens	FP/S	PPV
All Cases	/	0.66	62	0.04
Best Sensitivity Scan	PAT026	1.00	50	0.02
Best FP Scan	PAT012	0.33	24	0.04
Worst Sensitivity Scan	PAT028	0.00	28	0.00
Worst FP Scan	PAT017	0.67	109	0.02

The fig. 5.1 presents four heatmaps, each representing a specific performance metric: Sensitivity, Number of False Positives, Number of False Negatives, and Number of True Positives. Each heatmap visualizes these metrics across three categories: Overall, Pathological Area, and Background.

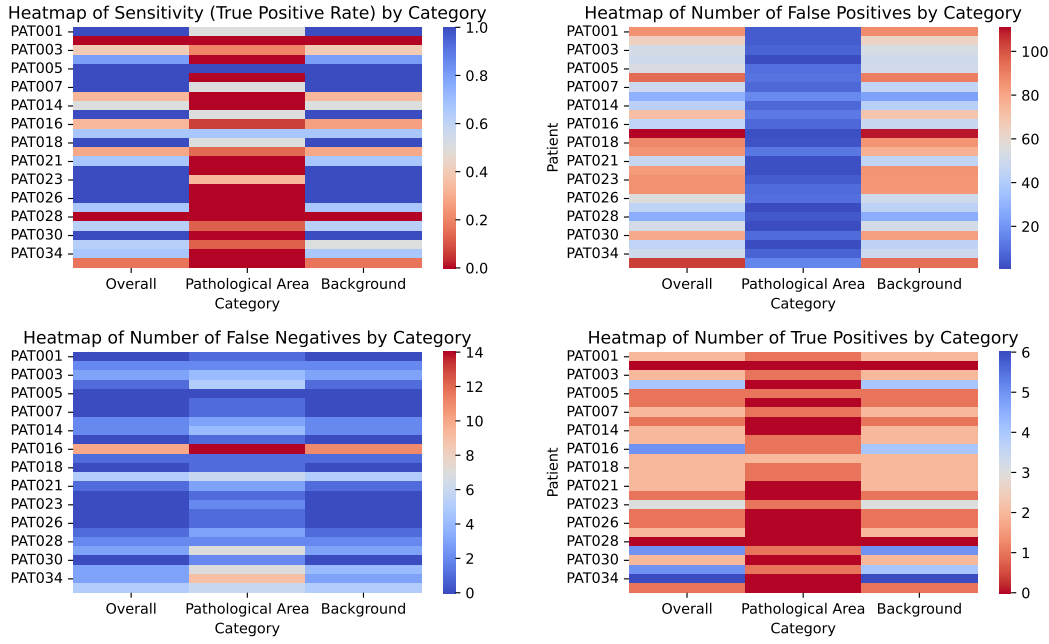


Figure 5.1: Comparative analysis of performance metrics across categories and patients. The diverging color scales in each heatmap allow for a clear differentiation between high and low values, indicating better performance with cooler colors.

We have also incorporated boxplots in our analysis to further illustrate the distribution and variability of our detection metrics across different patients and regions. These graphical representations (see fig. 5.2 and fig. 5.3) provide a comprehensive depiction of our dataset. The line inside the box marks the median value. The box's bounds represent the first (Q1) and third (Q3) quartiles, offering a sense of where the bulk of the data points lie[59].

5.1.2 SEGMENTATION

Having analyzed the detection performance of the algorithm, attention is now turned toward evaluating its segmentation capability. The Dice coefficient, presented in fig. 5.5 as boxplots, serves as a metric to assess the overlap between the algorithm's segmentation and the ground truth. Two scenarios are considered: one where all candidates (true positives and false positives) are employed to compute the segmentation pixel-wise metric and another where only the true positive candidates are utilized. Similarly, the Intersection over Union metric is visualized using boxplots in fig. 5.4.

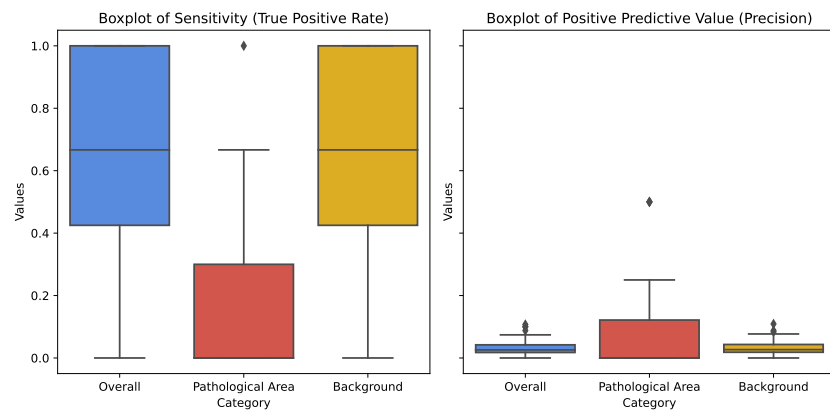


Figure 5.2

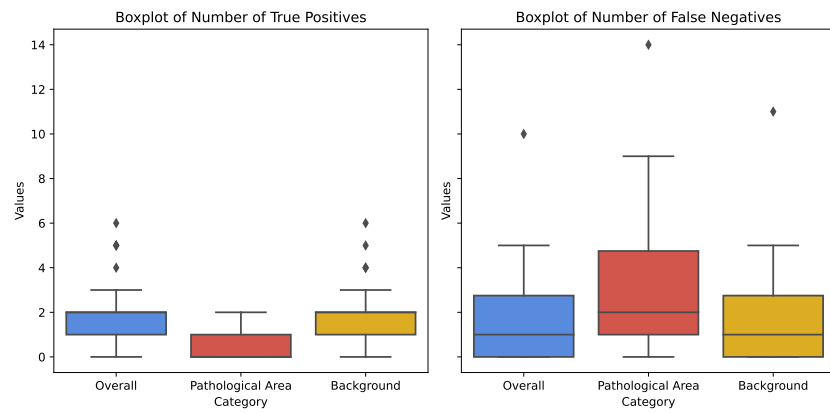


Figure 5.3

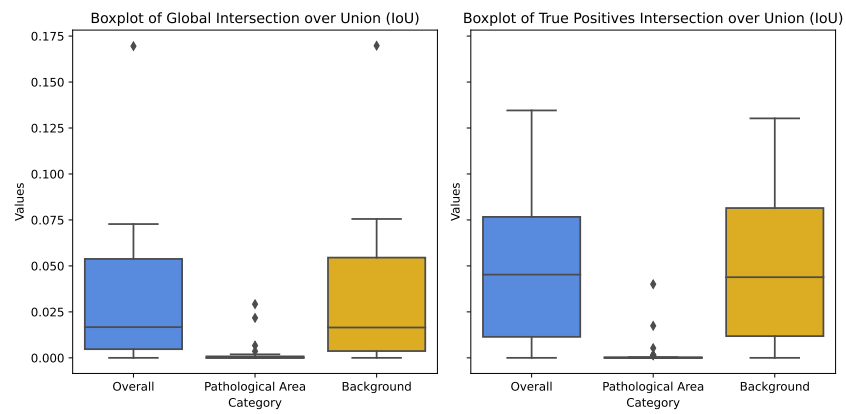


Figure 5.4

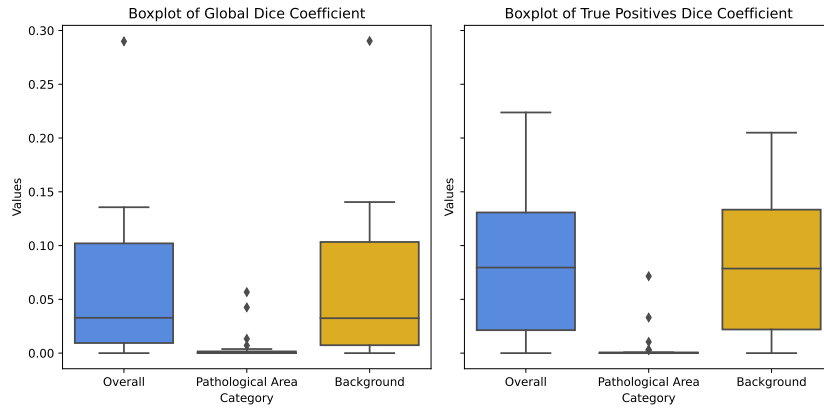


Figure 5.5

5.2 VISUAL ANALYSIS OF PULMONARY EMBOLUS DELINEATION

In this section, we provide a visual analysis of the segmentation results for PE candidates and their corresponding ground truth on a subset of slices from our dataset. An interpretation in the caption accompanies each figure to provide insights into the performance of the segmentation algorithm and highlight notable observations. Each illustration includes an overlay of the pulmonary artery (PA) mask, the segmented PE regions using contour lines, and their ground truths. Additionally, a zoomed-in view of the region of interest for each case is included to provide a closer examination of the segmentation results. It is important to note that this chapter includes interpretations and judgments based on the analysis of specific slices. However, the subsequent discussion chapter will present a comprehensive evaluation and discussion of the results, their significance, and their implications.

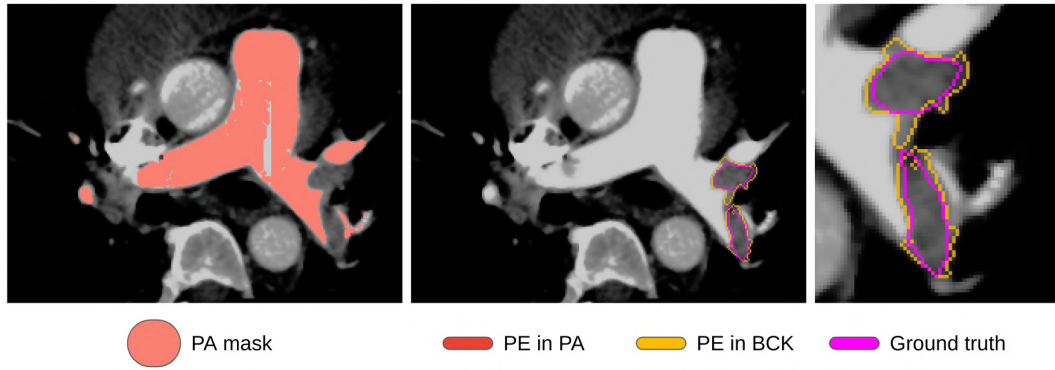


Figure 5.6: *Patient 29*. The algorithm successfully located and segmented the two regions in this slice that contained ground truth PE. The boundaries of the segmented areas closely align with those of the ground truth. This slice showcases the most promising visual result among all the analyzed scans.

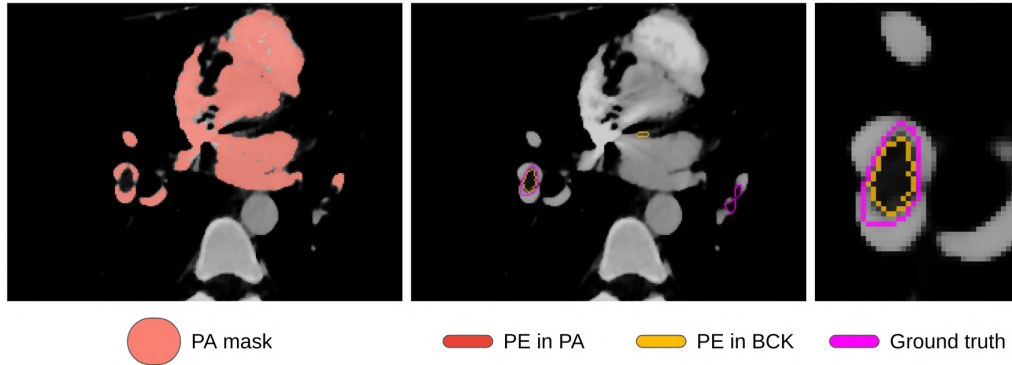


Figure 5.7: *Patient 27*. For this slice, the algorithm produced one false positive, one false negative, and one true positive. Once again, the segmentation of the true positive demonstrates boundaries that closely match the ground truth.

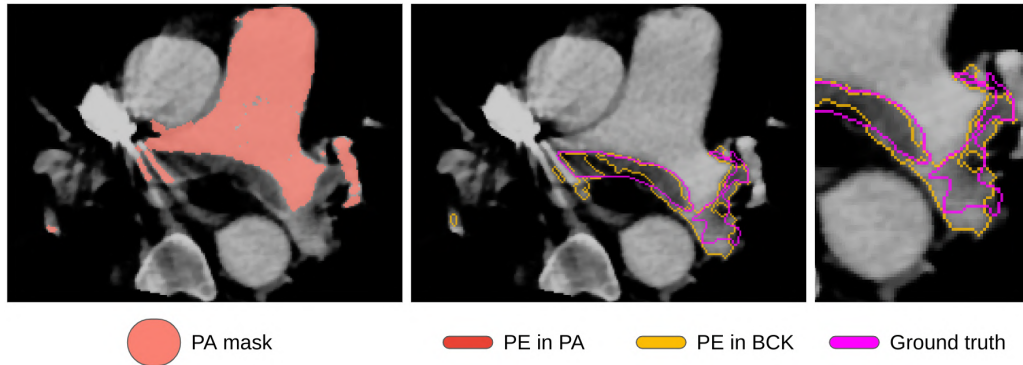


Figure 5.8: *Patient 3*. This slice also yields satisfying results, as a significant portion contains ground truth PE. The algorithm successfully located and segmented the PE, despite the complex and irregular shape of the embolus.

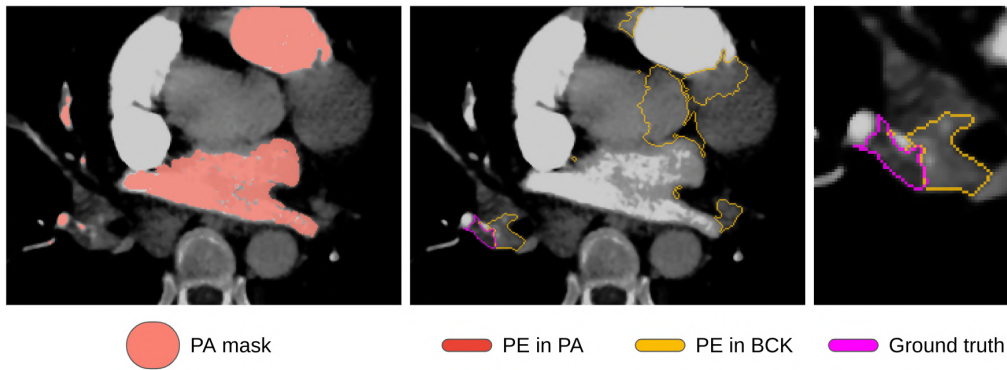


Figure 5.9: *Patient 4*. Multiple false positives were detected in different background areas of this slice. Although we detected a superpixel adjacent to a ground truth embolus, there was no overlap between the two regions. Notably, the intensity of the ground truth PE was relatively darker than the detected PE, indicating a deviation from the searching conditions of the superpixels.

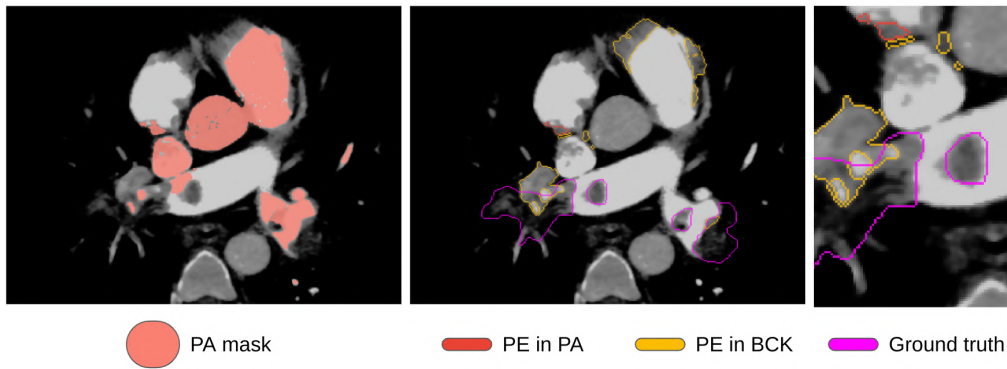


Figure 5.10: *Patient 7*. In this case, the algorithm did not detect the PE in the center of the image, which is easily detectable by a human observer. Moreover, the segmentation of the pulmonary arteries was incorrect, with both the vena cava and aorta being erroneously classified as pulmonary arteries. The majority of the segmented regions in this slice are false positives.

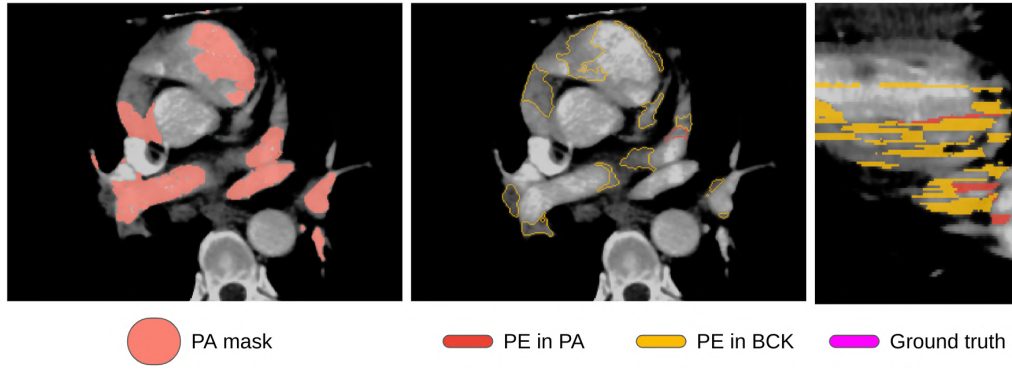


Figure 5.11: *Patient 12*. In this case, the algorithm encountered numerous segmentation errors, failing to accurately locate the PE regions. Additionally, a visual inspection of the sagittal plane on the left reveals that the PE candidates are connected across multiple slices, indicating a failure in the segmentation algorithm.

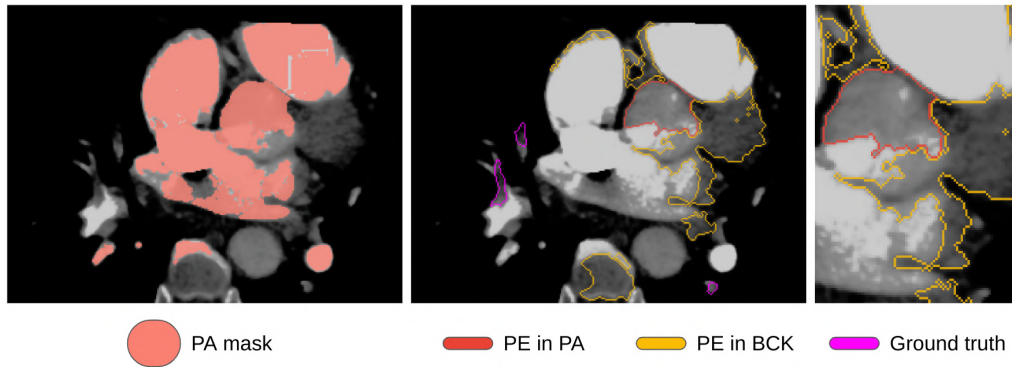


Figure 5.12: *Patient 16*. Like most PE in the ground truth, our algorithm did not detect it. Furthermore, in this case, the segmentation of the pulmonary arteries was unsuccessful, resulting in the appearance of multiple PE candidates. One particular region, outlined in red, was erroneously segmented as a PE from searching in the PA mask.

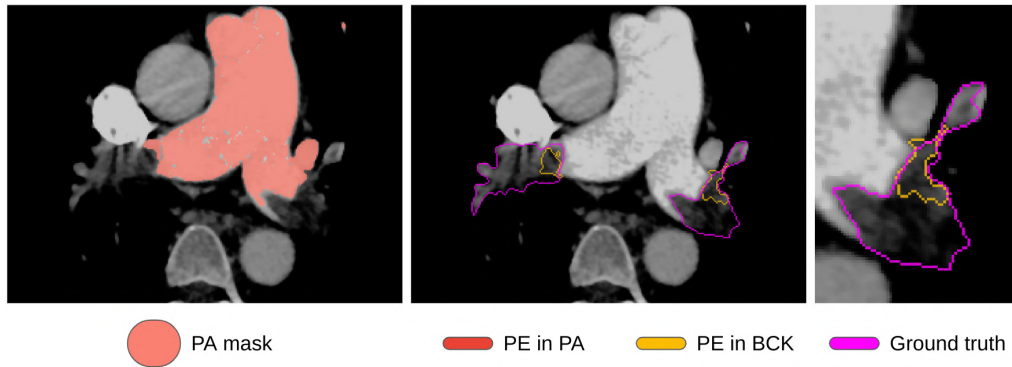


Figure 5.13: *Patient 17*. The segmentation algorithm successfully located the two regions where the ground truth PE is present. However, it was under-segmented, as only a small subset of the scan was accurately segmented.

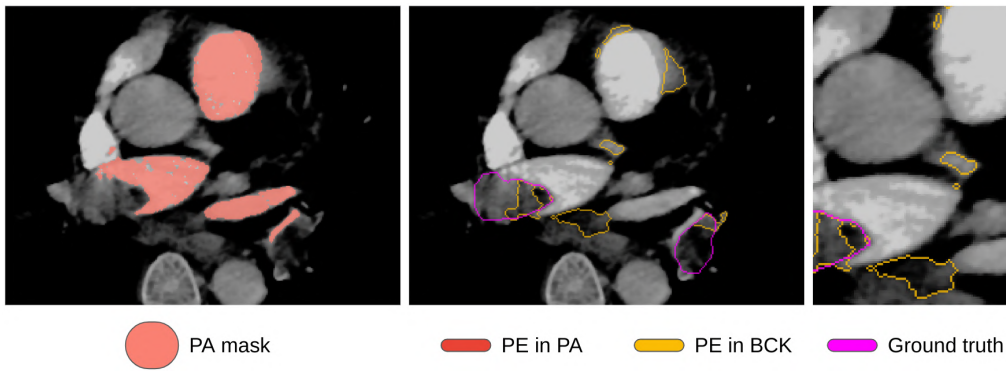


Figure 5.14: *Patient 17 other slice*. This is another slice from patient 17 (fig. 5.13) where additional instances of false positive are observed.

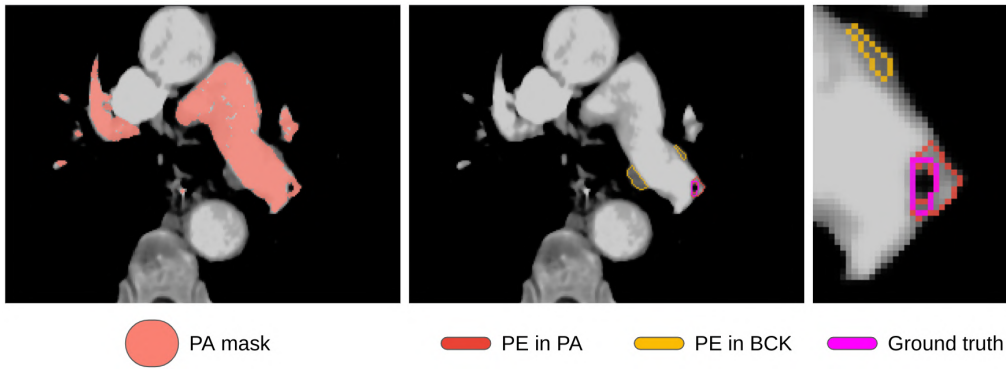


Figure 5.15: *Patient 18*. The algorithm detected a PE by searching within the pulmonary artery mask. However, it is an additional region of the true embolus. Furthermore, the dark region indicated by the ground truth was not preserved during the post-processing of the Quickshift algorithm.

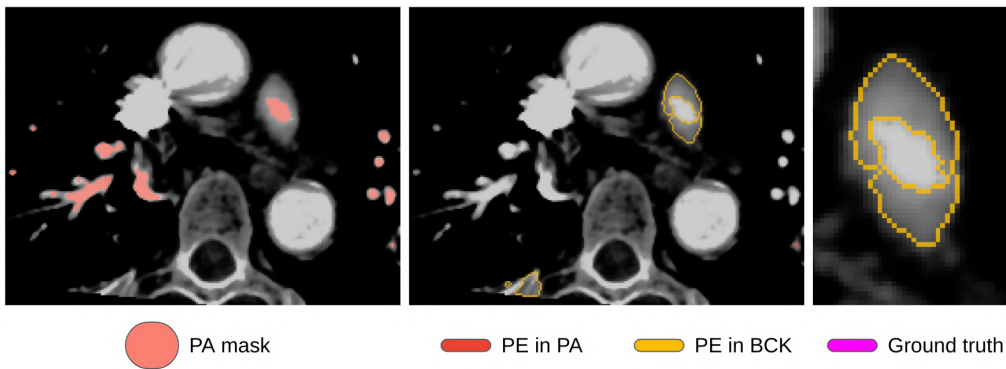


Figure 5.16: *Patient 19*. the algorithm incorrectly segmented a darker region surrounding the pulmonary artery in this slice.

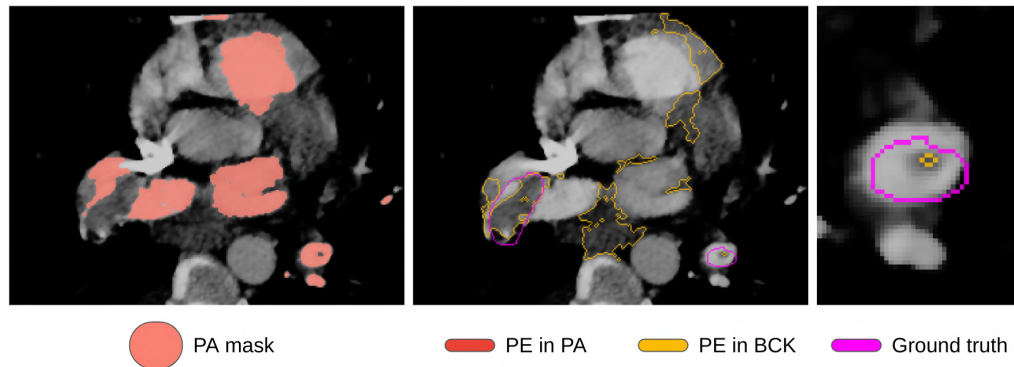


Figure 5.17: *Patient 23*. This slice exhibits a significant number of false positives. However, the algorithm successfully detected both PE regions present in the slice. Nevertheless, the PE on the right side is under-segmented, as indicated in the zoomed-in window.

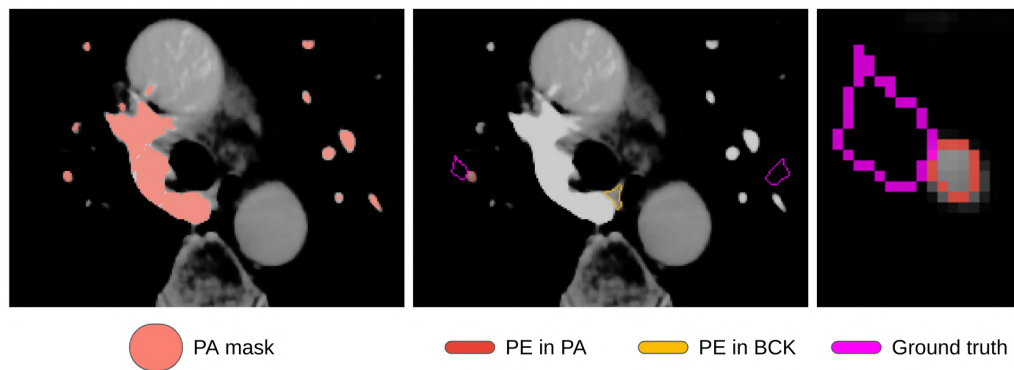


Figure 5.18: *Patient 31*. In this case, the ground truth PE was not visible due to the pre-processing steps included in our pipeline. However, the algorithm successfully located a candidate region that is close to, but not overlapping with, the ground truth PE.

6 DISCUSSIONS

This discussion chapter aims to provide an in-depth analysis of the key findings derived from this thesis, situating them within the broader context of the existing literature and exploring their implications. Additionally, recommendations for future research will be discussed.

The central objective of our research was to explore the development of a weakly supervised framework to address the challenge of limited annotated data in the field of pulmonary embolus (PE) detection in CT scans. We investigated using classical segmentation algorithms to replace high-cost ground truth labels. Moreover, we explored the integration of prior anatomical segmentation, specifically leveraging the pulmonary arteries mask to focus the search on relevant anatomical regions. While the scope of this thesis primarily focused on implementing the classical segmentation algorithm, the broader objective of the framework was kept in mind.

In response to Research Question 1, we discovered that the framework developed by Zhang *et al.* in [63] for segmenting coronary arteries in CT scans could be adapted for our task of segmenting PE. This adaptation did not present any major compatibility issues, suggesting the potential for leveraging existing frameworks for similar medical imaging tasks. However, certain framework components, such as query selection for suggested annotation and the deep learning model, were left for future work. We acknowledge the complexity of implementing this framework, as it requires multiple components, each of which needs to be robust to ensure the success of the overall process. Furthermore, we encountered difficulties reusing existing classical segmentation algorithms due to limited accessibility and compatibility issues. Consequently, we undertook the challenging task of developing our own algorithm for the first stage of the framework while other components were not implemented. Due to the complexity, we provided a brief overview of our findings for this research question, focusing more in-depth on the other two research questions.

For Research Question 1.1, we employed a classical segmentation algorithm based on superpixels generation. This approach offered several advantages, including reducing dataset size by aggregating pixels and preserving important information and the ability to smooth boundaries between vessels and embolus. It facilitated the grouping of pixels into coherent regions. We successfully obtained a pool of PE candidate labels by applying graph processing and filtering techniques. How-

ever, we observed many false positives (see fig. 5.3), which posed a challenge for the subsequent stages of the weakly supervised framework. These candidates' low sensitivity (see fig. 5.2) limited their usability, highlighting the need for further refinement to enhance the overall performance. Furthermore, even for the true positive cases, the algorithm struggled to precisely segment the PE regions, as evidenced by fig. 5.9, fig. 5.13, fig. 5.16, fig. 5.17 and fig. 5.18. However, we did encounter some instances where the algorithm successfully segmented emboli with high intersection over union, as shown in Figure fig. 5.6, fig. 5.7 and fig. 5.8 (with the complex shape). Unfortunately, these positive results did not substantially impact the overall DICE and IoU values, as they were computed on a per-scan basis and averaged (see fig. 5.4 and fig. 5.5).

Regarding Research Question 1.2, we incorporated prior segmentation of anatomical parts (pulmonary arteries, vena cava, and aorta) using the alpha-expansion for graph-cut optimization, inspired by the work of Browet *et al.* in [8]. We adapted this technique from a grid pixelized image format to our superpixels format. Although we achieved segmentation of anatomical parts, we found that its robustness varied across different scans. PE regions were then sought in both the pulmonary arteries mask and the background mask based on simple thresholds. By comparing the PE detection performance in the pulmonary arteries mask and the background mask, we found utilizing the anatomical prior segmentation beneficial for better localization of PE regions. In cases where the anatomical segmentation was well-performed, it improved detecting and segmenting PE candidates by reducing the number of false positives. However, the inconsistency in performance across scans reduced overall PE segmentation performance, as the algorithm could be misled by inaccurate prior anatomical segmentation.

In order to provide a comprehensive discussion, we compared our results to a review conducted by Gonzalez *et al.* in [16], which evaluated the performance of 8 different algorithms in detecting pulmonary embolus (PE). The PE candidates from the pulmonary arteries mask were used for this comparison. However, it is important to acknowledge the limitations of this comparison.

Firstly, the comparison was limited by the differences in the datasets used. Unfortunately, we could not access the same dataset as in the review, as the ground truth annotations were not publicly available. The variations in the datasets might introduce differences in complexity, image quality, and the distribution of PE instances, affecting the performance comparison.

Furthermore, Gonzalez's *et al.* review extended the ground truth by 5mm in all directions, resulting in the ground truth regions growing by 5 pixels. However, we did not apply this extension in our case. As a result, our method may underestimate the number of true positives, as more pos-

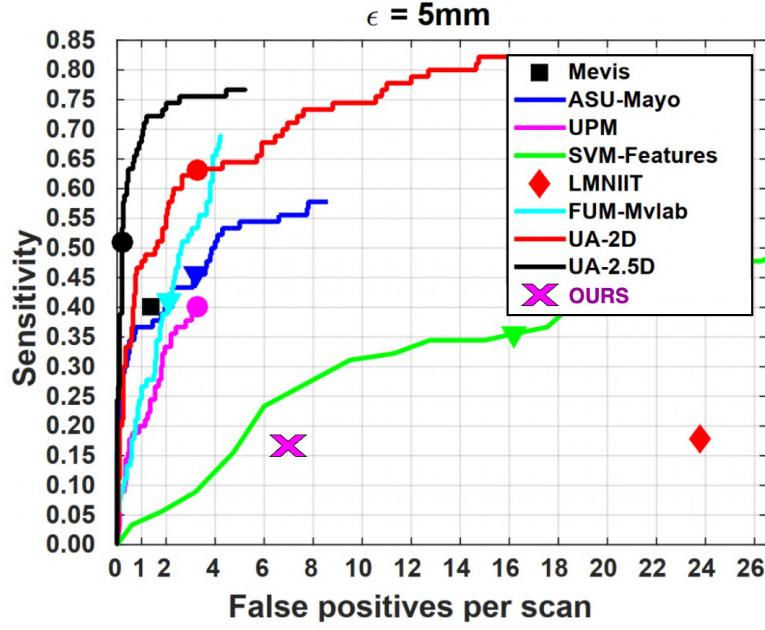


Figure 6.1: Our method with the "x" marker compared to other methods. Adapted from Fig 3. in Computer-aided detection for pulmonary embolism challenge (CAD-PE) [16].

itive instances could have been detected if we had considered the expanded ground truth (notable examples in fig. 5.18 and fig. 5.9).

Another factor is the methodology used to capture and report evaluation metrics. Gonzalez’s *et al.* review likely tuned the threshold of each algorithm to optimize sensitivity and false positive rate, which provide a more accurate representation of their performance. In our case, we could not perform a thorough threshold tuning due to the computationally intensive nature of obtaining these metrics. As a result, our method exhibits a dot representation of the results similar to method *Mevis* and method *LMNIIT*, which may not accurately reflect the true performance.

Lastly, it is important to note that the methods compared in [16] review focused solely on detecting pulmonary embolus. In contrast, our framework goes beyond detection and includes the segmentation of PE. This additional task adds complexity to our framework, as accurately localizing and delineating the PE regions is challenging. Therefore, comparing our method with those in the review must consider this added complexity.

Considering these limitations, comparing our results and Gonzalez’s *et al.* review provides some insights, but it should be interpreted cautiously. Furthermore, we evaluate the intermedi-

ate inaccurate label generation obtained from our first stage, developed in the broader two-stage framework.

Several key observations emerge in evaluating the suitability of these results as a starting point for the second stage of the framework. Firstly, it is evident that a considerable number of False Positives were obtained, coupled with relatively low sensitivity. The second stage intends to mitigate false positives by integrating radiologists' feedback annotations; however, the lower sensitivity raises concerns regarding the efficacy of incorporating such labeled data if the deep learning model is not sufficiently proficient in detecting False Negatives. It is worth noting that for our specific application, the medical context places greater importance on minimizing false negatives than positives. This raises an important consideration for further exploration and refinement of the framework to enhance its ability to identify true positive cases while minimizing false negatives.

6.1 LIMITATIONS AND FUTURE IMPROVEMENT

Our study has several limitations that warrant discussion and suggestions for future work to address these limitations and improve the proposed framework.

Firstly, due to the complexity of implementing each pipeline step, we could not focus extensively on parameter optimization. As a result, weaknesses were present in certain pipeline steps, which impacted the overall results. A potential future work could involve using stronger filters during the label generation step to avoid the deletion of superpixels that may contain valuable information. Additionally, developing a more robust 3D knowledge map for graph cuts, as opposed to the approximate 2D knowledge map used in this study, could enhance the performance of the segmentation of anatomical parts (see fig. 5.10). Furthermore, a more thorough exploration of selection criteria for PE candidates is needed, as time constraints limited the testing of various decision criteria. Future studies could explore better and more features-based decision criteria to improve this crucial part of the pipeline.

Another limitation of our study is the absence of validation by a radiologist. As we recognized that our results were not satisfactory, we did not validate them with the expertise of a professional radiologist. If improved metrics and visual interpretations are obtained, it is essential to validate the algorithm's outputs with radiologists to ensure the clinical relevance and accuracy of the findings.

Moreover, we did not thoroughly assess the accuracy of the ground truth annotations. At times, we observed inconsistencies where neighboring regions exhibited similar intensities, raising uncertainty about potential errors in our interpretations or the ground truth itself. Future work could involve meticulously examining the ground truth annotations to ensure their precision and address any ambiguities.

In conclusion, our study is limited by weaknesses in different pipeline steps, which were left unaddressed due to time constraints. However, future work should aim to improve intermediate results and address these limitations. Despite these limitations, the proposed framework, coupled with refinements, holds promise for further exploration and advancement in the field of segmenting pulmonary embolus on CT scans.

6.2 ADDITIONAL RESEARCH QUESTION

In addition to the limitations and future work mentioned above, we propose an additional research question that arose from the complexity and challenges encountered during this thesis. Considering the difficulties associated with segmenting PE in CT scans, we question whether the added value of segmenting PE instances, instead of solely detecting them, justifies the significant challenges and complexities accompanying the segmentation task.

Exploring this research question can illuminate the clinical relevance and practicality of investing resources in segmenting PE instances. Additionally, understanding the perspectives of radiologists and clinicians regarding the value of segmenting PE can provide valuable insights into the potential benefits and challenges faced in a real-world clinical setting.

Ultimately, the goal is to ensure that the efforts invested in segmentation are justified by the clinical benefits derived from the detailed analysis of PE instances.

7 CONCLUSION

This thesis set out to explore two main objectives: the implementation of a label generation algorithm using classical computer vision methods to act as inaccurate labels for a weakly supervised model and the incorporation of a prior anatomical segmentation to investigate its effect on performance when searching exclusively for PEs within the pulmonary arteries.

Our findings indicated that graph cuts enhanced the visibility of anatomical boundaries, and superpixels effectively reduced the problem size. However, this approach resulted in many false positive PE candidates and not all ground truth PEs were successfully detected. Moreover, even when a PE was identified, its segmentation often displayed a low intersection over the union or Dice score.

Introducing prior anatomical segmentation reduced the number of false positives, although at the expense of lower sensitivity. Furthermore, the robustness of this anatomical segmentation varied significantly between patients, leading to inconsistent performance results. Nevertheless, in instances where the prior segmentation performed well, we obtained some appropriate segmentations of PE regions.

Notably, our current results require further exploration and refinement. The level of sensitivity and specificity presented is insufficient for the segmentations to be used as labels, and the variability in PE candidate segmentation is considerable. Moreover, time constraints limited our ability to fine-tune each step of the pipeline effectively, and our comparisons were drawn with detection-focused methods rather than segmentation-focused ones.

Looking toward the future, there is a wealth of potential avenues for further research. The parameters from each step of our pipeline could benefit significantly from more dedicated fine-tuning. Implementing the second part of the weakly supervised framework also presents an opportunity for future work. Above all, the ultimate future research question remains to ascertain whether the time and effort invested in segmentation instead of detection can be justified by the resultant clinical benefits derived from a more detailed analysis of PE instances.

7 Conclusion

In conclusion, our research represents an exploration of a more nuanced understanding of PE detection and segmentation. It studies the promise of combining classical computer vision methods with deep learning techniques to potentially streamline and enhance the PE diagnostic process, even though the current outcomes necessitate further refinement. This thesis forms a basis for future studies, and with continued exploration, we aim to contribute significantly to improving PE diagnosis and treatment.

ACRONYMS

CTA	Computed Tomography Angiography
DICOM	Digital Imaging and Communications in Medicine
DL	Deep Learning
EPL	Ecole Polytechnique de Louvain
GPU	Graphics Processing Unit
PE	Pulmonary Embolus
RGBA	RGBA stands for Red, Green, Blue, and Alpha, and it is a color model that represents colors using combinations of red, green, and blue channels, along with an alpha channel for transparency.

GLOSSARY

$\mathcal{P}$	A set of labels/instances of pulmonary embolism in a scan.
alpha expansion	A graph-based optimization technique used to minimize the energy function in graph-cut algorithms.
anatomical prior knowledge	Prior knowledge about the human body structure that can be incorporated into the segmentation process.
Aorta	The main artery that carries oxygenated blood from the heart to the rest of the body.
background	The part of the image corresponding to non-relevant or non-anatomical regions, such as the spine.
CAD filter	Curvature Anisotropic Diffusion filter, a technique used for smoothing images while preserving sharp features.
connected components	Clusters or groups of connected nodes in a graph.
embolies	Blood clots or other foreign bodies that obstruct blood vessels.
graph-cuts	An algorithmic technique for solving energy minimization problems on graphs.
histogram equalization	A method to adjust an image's contrast by redistributing its histogram's intensities.
horizontal connections	Connections formed between superpixels that share a boundary within the same slice.
Hounsfield units	A unit of measurement used in CT scans to quantify radiodensity.
K	The knowledge map is the union of the set of relative points and fixed points for a specific semantic class and slice.
K_a	The knowledge map for a specific anatomical structure and slice represents the set of relative and fixed points.
Ker	The Gaussian Kernel function used to compute the similarity metric between a superpixel and a reference point on the knowledge map.
KF	The set of fixed points on the knowledge map, representing the CT scan array indices.

Glossary

knowledge map	An approximate and simple atlas or reference model that guides the anatomical segmentation process.
knowledge points	A set of points on the axial plane of the scan, each linked to a semantic class, used in the knowledge map.
KR	The set of relative points on the knowledge map for a specific slice, obtained by adding translation vectors to the vena cava center position.
KRtranslation	The set of all possible translation vectors in the axial plane of the CT scan.
millimeter	millimeter
PE candidates	Potential locations or regions that may contain pulmonary embolus.
Pulmonary Arteries	blood vessels that carry oxygen-poor blood from the right side of your heart to your lungs.
pulmonary arteries	The blood vessels that carry deoxygenated blood from the heart to the lungs.
Quickshift	An algorithm for fast hierarchical image segmentation.
S	The set of 2-dimensional points on the axial plane of the CT scan.
slice number	The index of a specific slice in the CT scan.
Spine	Vertebrae extending from the skull to the small of the back.
super-pixels	Regions obtained by grouping similar pixels together.
unary cost	The cost associated with assigning a specific label to a superpixel, typically derived from the unary score.
unary score	A measure of the likelihood or confidence of assigning a specific label to a superpixel.
VCcenter	The center position of the vena cava in the axial plane for a given slice.
vCPUs	Virtual Central Processing Units.
vertical connections	Connections formed between superpixels that are on different but adjacent slices and overlap when projected onto the axial plane.
windowing	A technique used to modify the display of image intensities within a specific range to enhance visualization.

BIBLIOGRAPHY

1. R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Ssstrunk. “SLIC superpixels compared to state-of-the-art superpixel methods”. *IEEE transactions on pattern analysis and machine intelligence* 34:11, 2012, pp. 2274–2282.
2. N. Adaloglou. *Introduction to medical image processing with Python: CT lung and vessel segmentation without labels*. 2022. URL: <https://theaisummer.com/medical-image-python/>.
3. Q. An, B. Jiang, and J. Yuan. “The Overview of Medical Image Processing Based on Deep Learning”. In: *Proceedings of 2021 International Conference on Medical Imaging and Computer-Aided Diagnosis (MICAD 2021)*. Ed. by R. Su, Y.-D. Zhang, and H. Liu. Springer Singapore, Singapore, 2022, pp. 411–417. ISBN: 978-981-16-3880-0.
4. *Blood Clots - Hematology.org*. URL: <https://www.hematology.org/education/patients/blood-clots>.
5. N. Bouchier. *L’UCL Saint-Luc de Bruxelles à la pointe en imagerie médicale avec le scanner spectral Philips IQon - Informations Presse | Philips*. 2016. URL: <https://www.philips.be/fr/a-w/about/news/archive/standard/about/news/communiqués/Healthcare/2016/20160713-ucl-saint-luc-de-bruxelles-a-la-pointe-en-imagerie-medicale-avec-le-scanner-spectral-philips-iqon.html> (visited on 05/05/2023).
6. H. Bouma, J. J. Sonnemans, A. Vilanova, and F. A. Gerritsen. “Automatic detection of pulmonary embolism in CTA images”. *IEEE transactions on medical imaging* 28:8, 2009, pp. 1223–1230.
7. Y. Boykov, O. Veksler, and R. Zabih. “Fast approximate energy minimization via graph cuts”. *IEEE Transactions on pattern analysis and machine intelligence* 23, 11 2001, pp. 1222–1239.
8. A. Browet, C. D. Vleeschouwer, L. Jacques, N. Mathiah, B. Saykali, and I. Migeotte. “Cell segmentation with random ferns and graph-cuts”. In: 2016, pp. 4145–4149.
9. A. B. Cheikh, G. Gorincour, H. Nivet, J. May, M. Seux, P. Calame, V. Thomson, E. Delabrousse, and A. Cromb. “How artificial intelligence improves radiological interpretation in suspected pulmonary embolism”. *European Radiology* 32:9, 2022, pp. 5831–5842.

10. Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger. “3D U-Net: learning dense volumetric segmentation from sparse annotation”. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Athens, Greece, October 17–21, 2016, Proceedings, Part II* 19. Springer. 2016, pp. 424–432.
11. E. Colak, F. C. Kitamura, S. B. Hobbs, C. C. Wu, M. P. Lungren, L. M. Prevedello, J. Kalpathy-Cramer, R. L. Ball, G. Shih, A. Stein, et al. “The RSNA pulmonary embolism CT dataset”. *Radiology: Artificial Intelligence* 3:2, 2021, e200254. URL: <https://kaggle.com/competitions/rsna-str-pulmonary-embolism-detection>.
12. E. Colak, F. C. Kitamura, S. B. Hobbs, C. C. Wu, M. P. Lungren, L. M. Prevedello, J. Kalpathy-Cramer, R. L. Ball, G. Shih, A. Stein, et al. “The RSNA pulmonary embolism CT dataset”. *Radiology: Artificial Intelligence* 3:2, 2021, e200254.
13. P. F. Felzenszwalb and D. P. Huttenlocher. “Efficient graph-based image segmentation”. *International journal of computer vision* 59, 2004, pp. 167–181.
14. Y. Gal and Z. Ghahramani. “Dropout as a bayesian approximation: Representing model uncertainty in deep learning”. In: *international conference on machine learning*. PMLR. 2016, pp. 1050–1059.
15. S. Z. Goldhaber and C. G. Elliott. “Acute Pulmonary Embolism: Part I: Epidemiology, Pathophysiology, and Diagnosis”. *Circulation* 108, 22 2003, pp. 2726–2729. ISSN: 00097322. DOI: [10.1161/01.CIR.0000097829.89204.0C](https://doi.org/10.1161/01.CIR.0000097829.89204.0C). URL: <http://www.circulationaha.org>.
16. G. González, D. Jimenez-Carretero, S. Rodriguez-López, C. Cano-Espinosa, M. Cazorla, T. Agarwal, V. Agarwal, N. Tajbakhsh, M. B. Gotway, J. Liang, et al. “Computer aided detection for pulmonary embolism challenge (CAD-PE)”. *arXiv preprint arXiv:2003.13440*, 2020.
17. J. Guo, X. Liu, Y. Chen, S. Zhang, G. Tao, H. Yu, H. Zhu, W. Lei, H. Li, and N. Wang. “AANet: Artery-Aware Network for Pulmonary Embolism Detection in CTPA Images”. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*. Ed. by L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li. Springer Nature Switzerland, Cham, 2022, pp. 473–483. ISBN: 978-3-031-16431-6.
18. K. T. Horlander, D. M. Mannino, and K. V. Leeper. “Pulmonary embolism mortality in the United States, 1979-1998: an analysis using multiple-cause mortality data”. *Archives of internal medicine* 163, 14 2003, pp. 1711–1717.
19. W. Huda and R. Slone. “Review of Radiologic Physics. Baltimor”. *Md: Williams & Wilkins*, 1995.

20. Q. Huo and B. Ma. "Online adaptive learning of continuous-density hidden Markov models based on multiple-stream prior evolution and posterior pooling". *Speech and Audio Processing, IEEE Transactions on* 9, 2001, pp. 388–398. DOI: [10.1109/89.917684](https://doi.org/10.1109/89.917684).
21. D. Karimi, H. Dou, S. K. Warfield, and A. Gholipour. "Deep learning with noisy labels: Exploring techniques and remedies in medical image analysis". *Medical Image Analysis* 65, 2020, p. 101759. ISSN: 1361-8415. DOI: <https://doi.org/10.1016/j.media.2020.101759>. URL: <https://www.sciencedirect.com/science/article/pii/S1361841520301237>.
22. C. Kearon. "Diagnosis of pulmonary embolism". *Cmaj* 168:2, 2003, pp. 183–194.
23. C. S. Kwok, C. W. Wong, S. Lovatt, P. K. Myint, and Y. K. Loke. "Misdiagnosis of pulmonary embolism and missed pulmonary embolism: A systematic review of the literature". *Health Sciences Review* 3, 2022, p. 100022. ISSN: 2772-6320. DOI: <https://doi.org/10.1016/j.hsr.2022.100022>. URL: <https://www.sciencedirect.com/science/article/pii/S2772632022000113>.
24. H. Le, D. Samaras, T. Kurc, R. Gupta, K. Shroyer, and J. Saltz. "Pancreatic Cancer Detection in Whole Slide Images Using Noisy Label Annotations". In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*. Ed. by D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, and A. Khan. Springer International Publishing, Cham, 2019, pp. 541–549. ISBN: 978-3-030-32239-7.
25. D.-H. Lee et al. "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks". In: *Workshop on challenges in representation learning, ICML*. Vol. 3. 2. 2013, p. 896.
26. J. Léger, E. Brion, P. Desbordes, C. De Vleeschouwer, J. A. Lee, and B. Macq. "Cross-Domain Data Augmentation for Deep-Learning-Based Male Pelvic Organ Segmentation in Cone Beam CT". *Applied Sciences* 10:3, 2020. ISSN: 2076-3417. DOI: [10.3390/app10031154](https://doi.org/10.3390/app10031154). URL: <https://www.mdpi.com/2076-3417/10/3/1154>.
27. B. P. Little. "Approach to Chest Computed Tomography". *Clinics in Chest Medicine* 36, 2 2015, pp. 127–145. ISSN: 02725231. DOI: [10.1016/j.ccm.2015.02.001](https://doi.org/10.1016/j.ccm.2015.02.001).
28. W. Liu, M. Liu, X. Guo, P. Zhang, L. Zhang, R. Zhang, H. Kang, Z. Zhai, X. Tao, J. Wan, and S. Xie. "Evaluation of acute pulmonary embolism and clot burden on CTPA with deep learning". *European Radiology* 30:6, 2020, pp. 3567–3575. ISSN: 1432-1084. DOI: [10.1007/s00330-020-06699-8](https://doi.org/10.1007/s00330-020-06699-8). URL: <https://doi.org/10.1007/s00330-020-06699-8>.

29. X. Ma, E. C. Ferguson, X. Jiang, S. I. Savitz, and S. Shams. “A multitask deep learning approach for pulmonary embolism detection and identification”. *Scientific Reports* | 12, 123, p. 13087. DOI: [10.1038/s41598-022-16976-9](https://doi.org/10.1038/s41598-022-16976-9). URL: <https://doi.org/10.1038/s41598-022-16976-9>.
30. R. Manniesing, M. A. Viergever, and W. J. Niessen. “Vessel enhancing diffusion: A scale space representation of vessel structures”. *Medical Image Analysis* 10:6, 2006, pp. 815–825. ISSN: 1361-8415. DOI: <https://doi.org/10.1016/j.media.2006.06.003>. URL: <https://www.sciencedirect.com/science/article/pii/S1361841506000442>.
31. M. Masoudi, H.-R. Pourreza, M. Saadatmand-Tarzjan, N. Eftekhari, F. S. Zargar, and M. P. Rad. “A new dataset of computed-tomography angiography images for computer-aided detection of pulmonary embolism”. *Scientific data* 5:1, 2018, pp. 1–9.
32. Y. Masutani, H. MacMahon, and K. Doi. “Computerized detection of pulmonary embolism in spiral CT angiography based on volumetric image analysis”. *IEEE Transactions on Medical Imaging* 21:12, 2002, pp. 1517–1523.
33. R. F. MD. *Pulmonary Embolism Diagnosis: The 3-Step Process*. 2022. URL: <https://www.verywellhealth.com/pulmonary-embolism-diagnosis-4163944>.
34. F. Milletari, N. Navab, and S.-A. Ahmadi. “V-net: Fully convolutional neural networks for volumetric medical image segmentation”. In: *2016 fourth international conference on 3D vision (3DV)*. Ieee. 2016, pp. 565–571.
35. N. Otsu. “A Threshold Selection Method from Gray-Level Histograms”. *IEEE Transactions on Systems, Man, and Cybernetics* 9:1, 1979, pp. 62–66. DOI: [10.1109/TSMC.1979.4310076](https://doi.org/10.1109/TSMC.1979.4310076).
36. H. Özkan, O. Osman, S. Şahin, and A. F. Boz. “A novel method for pulmonary embolism detection in CTA images”. *Computer Methods and Programs in Biomedicine* 113:3, 2014, pp. 757–766. ISSN: 0169-2607. DOI: <https://doi.org/10.1016/j.cmpb.2013.12.014>. URL: <https://www.sciencedirect.com/science/article/pii/S0169260713004045>.
37. J. Peng and Y. Wang. “Medical image segmentation with limited supervision: A review of deep network models”. *IEEE Access* 9, 2021, pp. 36827–36851.
38. E. Pichon, C. L. Novak, A. P. Kiraly, and D. P. Naidich. “A novel method for pulmonary emboli visualization from high-resolution CT images”. In: *Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display*. Vol. 5367. SPIE. 2004, pp. 161–170.
39. J. Pu, J. K. Leader, J. Sechrist, C. A. Beeche, J. P. Singh, I. K. Ocak, and M. G. Risbano. “Automated identification of pulmonary arteries and veins depicted in non-contrast chest CT scans”. *Medical Image Analysis* 77, 2022, p. 102367.

40. M. Ren, W. Zeng, B. Yang, and R. Urtasun. “Learning to reweight examples for robust deep learning”. In: *International conference on machine learning*. PMLR. 2018, pp. 4334–4343.
41. R. D. Rudyanto, S. Kerkstra, E. M. van Rikxoort, C. Fetita, P.-Y. Brillet, C. Lefevre, W. Xue, X. Zhu, J. Liang, İ. Öksüz, D. Ünay, K. Kadipaşaoğlu, R. S. J. Estépar, J. C. Ross, G. R. Washko, J.-C. Prieto, M. H. Hoyos, M. Orkisz, H. Meine, M. Hüllebrand, C. Stöcker, F. L. Mir, V. Naranjo, E. Villanueva, M. Staring, C. Xiao, B. C. Stoel, A. Fabijanska, E. Smistad, A. C. Elster, F. Lindseth, A. H. Foruzan, R. Kiros, K. Popuri, D. Cobzas, D. Jimenez-Carretero, A. Santos, M. J. Ledesma-Carbayo, M. Helmberger, M. Urschler, M. Pienn, D. G. Bosboom, A. Campo, M. Prokop, P. A. de Jong, C. Ortiz-de-Solorzano, A. Muñoz-Barrutia, and B. van Ginneken. “Comparing algorithms for automated vessel segmentation in computed tomography scans of the lung: the VESSEL12 study”. *Medical Image Analysis* 18:7, 2014, pp. 1217–1232. ISSN: 1361-8415. DOI: <https://doi.org/10.1016/j.media.2014.07.003>. URL: <https://www.sciencedirect.com/science/article/pii/S136184151400111X>.
42. R. Samet and Z. Yildirim. “A New Methodology for Blood Vessel Segmentation on Lung CT Images”. In: 2016, pp. 1–7. DOI: [10.1109/NicoInt.2016.1](https://doi.org/10.1109/NicoInt.2016.1).
43. A. P. Sauter, N. Shapira, F. K. Kopp, J. Aichele, J. Bodden, A. Knipfer, E. J. Rummeny, and P. B. Noël. “CTPA with a conventional CT at 100 kVp vs. a spectral-detector CT at 120 kVp: Comparison of radiation exposure, diagnostic performance and image quality”. *European Journal of Radiology Open* 7, 2020, p. 100234. ISSN: 2352-0477. DOI: <https://doi.org/10.1016/j.ejro.2020.100234>. URL: <https://www.sciencedirect.com/science/article/pii/S235204772030023X>.
44. R. Sebbe, B. Gosselin, R. Lédée, C. Léger, E. Coche, and B. Macq. “Framework for computer-aided detection of pulmonary embolism”. *International Journal* 1:1, 2015.
45. Z. Sedghi Gamechi, A. M. Arias-Lorza, Z. Saghir, D. Bos, and M. de Bruijne. “Assessment of fully automatic segmentation of pulmonary artery and aorta on noncontrast CT with optimal surface graph cuts”. *Medical Physics* 48:12, 2021, pp. 7837–7849.
46. E. Seeram. *Computed Tomography: Physical Principles, Clinical Applications, and Quality Control 1994 I Philadelphia*.
47. M. M. R. Siddiquee, Z. Zhou, N. Tajbakhsh, R. Feng, M. B. Gotway, Y. Bengio, and J. Liang. “Learning fixed points in generative adversarial networks: From image-to-image translation to disease detection and localization”. In: *Proceedings of the IEEE/CVF international conference on computer vision*. 2019, pp. 191–200.

48. S. Soffer, E. Klang, O. Shimon, Y. Barash, N. Cahan, H. Greenspan, and E. Konen. “Deep learning for pulmonary embolism detection on computed tomography pulmonary angiogram: a systematic review and meta-analysis”. *Scientific reports* 11:1, 2021, p. 15814.
49. P. D. Stein, M. L. Terrin, A. Gottschalk, A. Alavi, and J. W. Henry. “Value of ventilation/perfusion scans versus perfusion scans alone in acute pulmonary embolism”. *The American journal of cardiology* 69:14, 1992, pp. 1239–1241.
50. V. F. Tapson. “Acute Pulmonary Embolism”. *New England Journal of Medicine* 358, 10 2008, pp. 1037–1052. ISSN: 0028-4793. DOI: [10.1056/NEJMr072753](https://doi.org/10.1056/NEJMr072753).
51. T. Tasdizen, R. Whitaker, P. Burchard, and S. Osher. “Geometric surface smoothing via anisotropic diffusion of normals”. In: *IEEE Visualization, 2002. VIS 2002*. 2002, pp. 125–132. DOI: [10.1109/VISUAL.2002.1183766](https://doi.org/10.1109/VISUAL.2002.1183766).
52. F. Thiry. “Contrastive learning for annotation enrichment and automatic classification of pulmonary embolisms”. Anglais, 2021. URL: <http://hdl.handle.net/2078.1/thesis:33092>.
53. T. Thompson and C. Kabrhel. “Overview of acute pulmonary embolism in adults”. *Post TW, ed. UpToDate. Waltham*, 2023. URL: <https://www.uptodate.com/contents/overview-of-acute-pulmonary-embolism-in-adults>.
54. A. Vedaldi and S. Soatto. “Quick shift and kernel methods for mode seeking”. In: *Computer Vision—ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12–18, 2008, Proceedings, Part IV 10*. Springer. 2008, pp. 705–718.
55. F. Verschuren, G. Liistro, R. Coffeng, F. Thys, J. Roeseler, F. Zech, and M. Reynaert. “Volumetric Capnography as a Screening Test for Pulmonary Embolism in the Emergency Department”. *Chest* 125, 3 2004, pp. 841–850. ISSN: 00123692. DOI: [10.1378/chest.125.3.841](https://doi.org/10.1378/chest.125.3.841).
56. W. Wang, R. Feng, J. Chen, Y. Lu, T. Chen, H. Yu, D. Z. Chen, and J. Wu. “Nodule-plus R-CNN and deep self-paced active learning for 3D instance segmentation of pulmonary nodules”. *Ieee Access* 7, 2019, pp. 128796–128805.
57. W. Wang, Y. Lu, B. Wu, T. Chen, D. Z. Chen, and J. Wu. “Deep active self-paced learning for accurate pulmonary nodule segmentation”. In: *Medical Image Computing and Computer Assisted Intervention—MICCAI 2018: 21st International Conference, Granada, Spain, September 16–20, 2018, Proceedings, Part II 11*. Springer. 2018, pp. 723–731.
58. J. Weickert et al. *Anisotropic diffusion in image processing*. Vol. 1. Teubner Stuttgart, 1998.
59. Wikipedia. *Box plot* — *Wikipedia, The Free Encyclopedia*. <http://en.wikipedia.org/w/index.php?title=Box%20plot&oldid=1150807106>. [Online; accessed 18-May-2023]. 2023.

60. Q. Yang, Y. Liu, T. Chen, and Y. Tong. "Federated machine learning: Concept and applications". *ACM Transactions on Intelligent Systems and Technology (TIST)* 10:2, 2019, pp. 1–19.
61. G. Zantonelli, D. Cozzi, A. Bindi, E. Cavigli, C. Moroni, S. Luvarà, G. Grazzini, G. Danti, V. Granata, and V. Miele. *Acute Pulmonary Embolism: Prognostic Role of Computed Tomography Pulmonary Angiography (CTPA)*. 2022. DOI: [10.3390/tomography8010042](https://doi.org/10.3390/tomography8010042).
62. C. Zhang, M. Sun, Y. Wei, H. Zhang, S. Xie, and T. Liu. "Automatic segmentation of arterial tree from 3D computed tomographic pulmonary angiography (CTPA) scans". *Computer Assisted Surgery* 24:sup2, 2019, pp. 79–86.
63. J. Zhang, G. Wang, H. Xie, S. Zhang, N. Huang, S. Zhang, and L. Gu. "Weakly supervised vessel segmentation in X-ray angiograms by self-paced learning from noisy labels with suggestive annotation". *Neurocomputing* 417, 2020, pp. 114–127.
64. C. Zhou, H.-P. Chan, S. Patel, P. N. Cascade, B. Sahiner, L. M. Hadjiiski, and E. A. Kazerooni. "Preliminary investigation of computer-aided detection of pulmonary embolism in threedimensional computed tomography pulmonary angiography images". *Academic radiology* 12:6, 2005, p. 782.
65. Z.-H. Zhou. "A brief introduction to weakly supervised learning". *National science review* 5:1, 2018, pp. 44–53.
66. F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He. "A Comprehensive Survey on Transfer Learning". *Proceedings of the IEEE* 109:1, 2021, pp. 43–76. DOI: [10.1109/JPROC.2020.3004555](https://doi.org/10.1109/JPROC.2020.3004555).

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/epl