N° d'ordre NNT : 2026ISAL0031

THESE de DOCTORAT DE L'INSA LYON, membre de l'Université de Lyon

Ecole Doctorale N° 160
Électronique, Électrotechnique, Automatique

Discipline : Traitement du signal et de l'image

Soutenue publiquement le 21/05/2026, par :

Morgane des Ligneris

Pulmonary embolism risk stratification based on vascular network modeling

Devant le jury composé de :

CONZE Pierre-Henri	PU, IMT Atlantique	Rapporteur
DELINGETTE Hervé	Directeur de recherche, Inria Sophia Antipolis	Rapporteur
BERTOLETTI Laurent	PUPH, Université Jean Monnet Saint-Étienne	Examinateur
JACOB Joseph	PU, University College London	Examinateur
ZULUAGA Maria A.	PU, EURECOM	Présidente du jury
FRINDEL Carole	Maître de conférence, HDR, INSA Lyon, IUF	Directrice de thèse
MERVEILLE Odyssée	Maître de conférence, INSA Lyon	Co-directrice de thèse

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

INSTITUT NATIONAL DES SCIENCES APPLIQUÉES LYON

Département de la Formation par la Recherche et des Études Doctorales (**FEDORA**)

Bâtiment Charlotte Perriand 3e étage
17, av. des Arts
69621 Villeurbanne Cédex
fedora@insa-lyon.fr

Référence : TH1354_DES LIGNERIS Morgane

L'INSA Lyon a mis en place une procédure de contrôle systématique via un outil de détection de similitudes (logiciel Compilatio). Après le dépôt du manuscrit de thèse, celui-ci est analysé par l'outil. Pour tout taux de similarité supérieur à 10%, le manuscrit est vérifié par l'équipe de FEDORA. Il s'agit notamment d'exclure les auto-citations, à condition qu'elles soient correctement référencées avec citation expresse dans le manuscrit.

Par ce document, il est attesté que ce manuscrit, dans la forme communiquée par la personne doctorante à l'INSA Lyon, satisfait aux exigences de l'Etablissement concernant le taux maximal de similitude admissible.

INSA LYON
Campus LyonTech La Doua
20, avenue Albert Einstein - 69621 Villeurbanne cedex - France
Tél. +33 (0)4 72 43 83 83 - Fax +33 (0)4 72 43 85 00
www.insa-lyon.fr

membre de UNIVERSITÉ DE LYON

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Département FEDORA – INSA Lyon - Ecoles Doctorales

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
ED 206 CHIMIE	**CHIMIE DE LYON** **https ://www.edchimie-lyon.fr** Sec. : Renée EL MELHEM Bât. Blaise PASCAL 7, Avenue Jean Capelle 69621 Villeurbanne CEDEX **secretariat@edchimie-lyon.fr**	**M. Stéphane DANIELE** C2P2-CPE LYON-UMR 5265 Bâtiment F308, BP 2077 43 Boulevard du 11 novembre 1918 69616 Villeurbanne CEDEX **directeur@edchimie-lyon.fr**
ED 341 E2M2	**ÉVOLUTION, ÉCOSYSTÈME, MICROBIOLOGIE, MODÉLISATION** **http ://e2m2.universite-lyon.fr** Sec. : Bénédicte LANZA Bâtiment Atrium, 43, Boulevard du 11 Novembre 1918 69622 Villeurbanne CEDEX Tél : 04.72.44.83.62 **secretariat.e2m2@univ-lyon1.fr**	**Mme Sandrine CHARLES** Université Claude Bernard Lyon 1 UFR Biosciences, Bâtiment Mendel 43, boulevard du 11 Novembre 1918 69622 Villeurbanne CEDEX **e2m2.codir@listes.univ-lyon1.fr**
ED 205 EDISS	**INTERDISCIPLINAIRE SCIENCES-SANTÉ** **http ://ediss.universite-lyon.fr** Sec. : Bénédicte LANZA Bâtiment Atrium, 43, Boulevard du 11 Novembre 1918 69622 Villeurbanne CEDEX Tél : 04.72.44.83.62 **secretariat.ediss@univ-lyon1.fr**	**M. Jérôme LAMARTINE** Université Claude Bernard Lyon 1 Equipe Intégrité Fonctionnelle du Tissu Cutané LBTI - UMR5305 CNRS IBCP - 7 passage du Vercors, 69367 Lyon CEDEX 07 Tel : 04.72.72.26.66 **jerome.lamartine@univ-lyon1.fr**
ED 34 EDML	**MATÉRIAUX DE LYON** **http ://ed34.universite-lyon.fr** Sec. : Yann DE ORDENANA Tél : 04.72.18.62.44 **yann.de-ordenana@ec-lyon.fr**	**M. Stéphane BENAYOUN** Ecole Centrale de Lyon Laboratoire LTDS 36 avenue Guy de Collongue 69134 Ecully CEDEX Tél : 04.72.18.64.37 **stephane.benayoun@ec-lyon.fr**
ED 160 EEA	**ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE** **https ://edeea.universite-lyon.fr** Sec. : Philomène TRECOURT Bâtiment Charlotte PERRIAND 69621 Villeurbanne CEDEX Tél : 04.72.43.71.70 **secretariat.edeea@insa-lyon.fr**	**M. Philippe DELACHARTRE** INSA LYON Laboratoire CREATIS Bâtiment Blaise Pascal, 7 avenue Jean Capelle 69621 Villeurbanne CEDEX Tél : 04.72.43.88.63 **philippe.delachartre@insa-lyon.fr**
ED 512 INFOMATHS	**INFORMATIQUE ET MATHÉMATIQUES** **http ://edinfomaths.universite-lyon.fr** Sec. : Renée EL MELHEM Bât. Blaise PASCAL 7, Avenue Jean Capelle 69621 Villeurbanne CEDEX **infomaths@univ-lyon1.fr**	**M. Dragos IFTIMIE** Université Claude Bernard Lyon 1 Bâtiment Braconnier 21 Avenue Claude BERNARD 69 622 Villeurbanne CEDEX Tél : 04.72.44.79.58 **direction.infomaths@listes.univ-lyon1.fr**
ED 162 MEGA	**MÉCANIQUE, ÉNERGÉTIQUE, GÉNIE CIVIL, ACOUSTIQUE** **http ://edmega.universite-lyon.fr** Sec. : Philomène TRECOURT Bâtiment Charlotte PERRIAND 69621 Villeurbanne CEDEX Tél : 04.72.43.71.70 **mega@insa-lyon.fr**	**M. Etienne PARIZET** INSA Lyon Laboratoire LVA Bâtiment St. Exupéry 25 bis av. Jean Capelle 69621 Villeurbanne CEDEX **etienne.parizet@insa-lyon.fr**
ED 483 ScSo	**ScSo**[1] **https ://edsciencessociales.universite-lyon.fr** Sec. : Mélina FAVETON Tél : 04.78.69.72.76 Université Lyon 2 Campus Berges du Rhône, Bureau BEL 309 18, Quai Claude BERNARD 69365 LYON CEDEX 07 **melina.faveton@univ-lyon2.fr**	**M. Bruno MILLY** Univ. Lyon 2 Campus Berges du Rhône 18, quai Claude Bernard 69365 LYON CEDEX 07 Bureau BEL 319 **bruno.milly@univ-lyon2.fr**

1. ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Acknowledgments

For the acknowledgments, I will start in English to properly thank the members of my jury. I would like to thank Pierre-Henri Conze and Hervé Delingette for accepting to be my referees and for taking the time to evaluate my manuscript. And I would like to thank Laurent Bertoletti, Joseph Jacob, and Maria A. Zuluaga for participating as my examiners.

Ensuite, j'aimerais remercier chaleureusement Carole Frindel et Odyssée Merveille, mes deux encadrantes de thèse, de m'avoir fait confiance pour mener à bien cette thèse. Merci Carole de m'avoir accueilli lors de mon tout premier stage à CREATIS lors de ma troisième année de licence en 2020 (il y a donc 6ans déjà!). Venant d'une formation de bioinformatique, il me manquait une partie concrète, je m'étais posé la question de m'orienter plutôt vers la santé et l'imagerie médicale, et ce premier stage avec toi et Timothée Jacquesson a complètement confirmé mon intérêt pour le domaine. Donc je vous remercie de votre accompagnement régulier pendant cette période un peu particulière liée à la COVID et aux conditions du distanciel. Puis en 2021, j'ai été ravie que tu sois, Carole, de nouveau partante pour m'accueillir en tant qu'alternante lors de ma dernière année de master. Ensuite avec Odyssée vous avez toutes les deux accepté ma candidature pour le projet de thèse PERSEVERE. Merci, Odyssée, de m'avoir aussi mise en confiance depuis le tout début de ce projet de thèse. Merci à vous deux pour tous vos conseils et votre accompagnement chaleureux. J'ai le sentiment que vous avez su voir autant mes forces que mes faiblesses et me guider tout du long vers plus d'assurance. J'ai bien conscience du chemin parcouru et j'en suis très reconnaissante, je n'aurais pas pu espérer mieux pour me sentir accompagnée et soutenue lors de cette thèse.

Du côté clinique, je tiens à remercier Pierre Croisille, Laurent Bertoletti et Allan Serva pour tous les riches échanges qui auront guidé les axes de recherche de cette thèse. Merci Laurent de nous avoir donné l'accès à la base de données de patients. Et merci Allan pour toutes les discussions, les corrections de mes segmentations, et surtout ton implication et ta disponibilité qui ont énormément participé au bon déroulé de cette thèse. Je n'aurais pas pu la faire sans toi, et je suis ravie d'avoir pu en retour contribuer (un peu) à ta thèse de médecine. J'espère que la suite de ta carrière sera tout aussi flamboyante (mais n'oublie pas de te reposer aussi!).

J'aimerais aussi remercier Sorina Pop, Axel Bonnet, Mickaël Srida, Frederic Cervenansky, Hélène Ratiney et Gaël Vila, avec qui j'ai pu réaliser des travaux à leurs côtés au cours de mon alternance. Merci, Sorina, de m'avoir encouragé et permis de présenter ces travaux lors de la conférence ISBI 2023 en Colombie au début de ma thèse. Et j'en profite pour te remercier ici, Mickaël, pour le prêt de la tablette graphique qui m'a permis de ne pas prendre des semaines en plus à essayer de segmenter avec la souris. Concernant la Colombie, je suis ravie d'avoir profité de cette expérience aux côtés de Juliette, Raoul, Louise, Gaël, Medhi, Samaneh, Alexandre et Luc. Je garde en souvenir la sortie "Rumba en Chiva", l'initiation au crochet et de bons cocktails partagés.

J'aimerais remercier également toutes les personnes des différents services du laboratoire qui ont pu m'aider et m'accompagner à chaque fois que j'avais une question ou besoin d'aide. Merci Beatrice, Sophie, Mira, Fabienne, Anthony, Fabrice et Pierre Ferrier.

Au cours de la thèse j'ai rejoint l'équipe LIFE, motivée par ses valeurs et ambitions, aux côtés de Raoul, Pilar,

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Arthur, Laura, Axel A. et Axel B., Odyssée, Antoine, Juliette, Maylis, Ludmilla, Nils, Basile, Mathilde et Thomas. C'était chouette d'être impliqué dans le labo autour des enjeux écologiques et d'avoir essayé de sensibiliser autour des différentes fresques, de la bibliothèque et de la démarche du quota carbone. Merci à vous, je suis contente d'avoir pu y participer à vos côtés.

Si l'on reprend la chronologie de la thèse, en devenant la doctorante d'Odyssée, j'ai rejoint la fameuse "team O" aux côtés de Jonas, Charles, Sophie, Pierre R., et qui accueillent plus récemment Nathan P. et Oscar (la relève est assurée!). Je m'y suis sentie tout de suite très bien accueillie et je tiens à vous remercier toustes. Merci tout particulièrement à Sophie, Pierre et Nathan, pour votre gentillesse, pour votre bonne humeur communicative et toute votre aide (et tes superbes pâtisseries Sophie, toujours réconfortantes!).

J'ai intégré le bureau 19, qui certes est surnommé l'usine avec son air d'open space d'une dizaine de doctorants, mais au sein duquel j'ai pu me sentir bien entourée durant mes 3 ans (et demi) de thèse. Donc merci à vous toutes, Méghane Decrooq, Raphaël Dumas, Sushong, Hang Jung, Gwendal, Baptiste, Marta, Célia G., Thierry, Nathan P., Pierre-Elliott, Robin, Basile, Oscar, Maylis. Faire la thèse à vos côtés était très réconfortant. Je vous remercie pour toutes les discussions et tous les moments conviviaux.

J'ai aussi envie de remercier toutes les personnes du bureau 13, dans lequel j'ai l'impression d'avoir trouvé un second bureau, encore plus à la fin de ma thèse. Merci de m'avoir autant laissé bouturer vos plantes à foison, squatter les puzzles, et débarquer avec ma chicorée. Donc merci Raoul, Julia, Emile, Romain, Valentin M., Thibaut et plus récemment Jeanne et Léo.

Côté repas, j'ai envie de remercier toute la team Pied du saule avec qui j'ai passé des super moments de midi avec toutes sortes de débats et beaucoup de tier liste de nourriture. Un merci en particulier à Maylis, Serena, Thomas, Joseph, Maël, Nicolas M., Mathilde et Jeanne. Je ne sais pas combien de fois vous avez vu popper ma tête à l'approche du repas. En parlant de nourriture, je dois remercier les "anciens", du temps où les pauses de 16h étaient incontournables avec un accompagnement de petits princes. Merci à Suzanne, Franck, Benoit, Antoine, Medhi, Samaneh, JB, Matthis, Thierry, Célestine. J'aimerais aussi remercier Clara, Maria, Sarra, Pierre F., Mashid, avec qui j'ai pu discuter et/ou partager les repas lorsque j'avais des tupps (qu'Arnaud avait gentiment préparés). Et concernant le temps de pause, merci, Romain, Valentin M., Valentin G. et PE, j'ai apprécié avoir de nombreuses discussions autour des animés, mangas et jeux vidéo ! Mention spéciale pour Romain, j'ai adoré pouvoir débattre [avec intensité] avec toi. Maria et Slim, j'ai envie de vous remercier dans une catégorie à part, la catégorie des personnes avec un enthousiasme extrêmement communicatif. Donc merci à vous d'être rayonnants de bonne humeur.

Côté créatif, merci Luis et Clara, c'était très agréable de pouvoir partager mon intérêt pour la peinture avec vous! Et je ne peux pas parler de créatif, sans mentionner le café Manitas, un atelier de peinture sur céramique auquel j'ai, il me semble, réussi à initier beaucoup d'entre vous. Donc merci d'avoir accepté de me suivre et de partager avec moi un moment de détente autour de craft. Merci Lorena, Julia, Marta, Maria, Sarra, Celia M., Celia G., Tatiana, Ludmilla, Maylis, Juliette, PE, Thomas, Clara, Nicolas M., Eleonore, Mathilde, Madeleine, Nathan H., Jeanne, Emile, Slim, Oscar et Séréna. J'ai envie aussi de remercier la team crochet/tricot/couture avec notamment Juliette, Maylis, Louise, Robin, Basile, Julia. C'était génial de pouvoir souffler un peu autour de ces moments créatifs avec vous toutes.

Merci aussi tout particulièrement à Maylis, Juliette, et Louise, merci d'avoir partagé de nombreuses discussions autant sur des sujets d'écologies, de féminisme, de politique, de cinéma et de livres que sur ce qui touche plus

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

à la sphère personnelle. Surtout pour ce qui touche plus à la sphère personnelle en fait. Vous êtes de chouettes amies sur qui on peut compter.

De manière générale je tiens à remercier toutes les personnes dont j'ai pu croiser le chemin et que je n'aurai pas nommées ici, avec qui tous les échanges ont rendu riches toutes ces années passées à CREATIS.

En dehors de CREATIS, j'aimerais remercier ma famille. Merci, mama, Jeje, mamaud. Merci Aliénor et Théodore. Merci tatie Alex, tonton Pascal, tonton Raph, Véro, Rachelle et Matteo. Merci à mes cousins Eymeric, Baptiste, Ludovic et Samuel. Merci à mes cousines Inès et Louise. Merci de m'avoir soutenue durant toutes mes années d'études, d'avoir été toujours derrière moi.

Enfin j'aimerais aussi remercier mes amies et amis. Merci Chloé, Louise et Emma, merci pour votre soutien, merci pour tous nos moments, vous avez grandement participé à mon équilibre durant cette thèse. Je vous admire et j'ai hâte de voir ce que vous allez réaliser (mention spéciale à Amal à qui je souhaite autant de réussite qu'à sa maman). Merci aussi, Rayan, Axel, Hortense, Beber, Solène, Romu, Waël, Adeline et Antho, de m'avoir permis de souffler autour d'un verre ou d'un bon petit plat! Aussi merci tout particulièrement à Arnaud, autant pour ton soutien durant les différents challenges rencontrés dans la thèse que pour tes bons plats maison. Merci mi amor d'avoir toujours été là dans les hauts comme dans les bas durant toutes ces années. Merci à vous toustes d'avoir été à mes côtés!

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Abstract

Pulmonary Embolism (PE) is the third most frequent acute cardiovascular syndrome, caused by the obstruction of a pulmonary artery by a blood clot. Its severity depends on the location, extent, and distribution of thrombi within the Pulmonary Arterial Tree (PAT). Current thrombotic burden scores, such as Qanadli and Mastora, rely on simplified anatomical templates and manual measurements that are time-consuming and poorly reproducible, limiting their clinical routine use. Furthermore, key cardiac biomarkers necessary for current clinical risk stratification are sometimes unavailable in emergency settings. To address those limitations, this thesis develops an automated pipeline for graph-based modeling of the PAT from Computed Tomography Pulmonary Angiography (CTPA). This graph-based modeling serves as the foundation for extracting patient-specific, imaging-derived biomarkers for PE severity assessment. This work was conducted within the Pulmonary Embolism Risk Stratification basEd on Vascular nEtwoRk modEling (PERSEVERE) project, using a cohort of 353 CTPA acquisitions from PE patients.

The first contribution is VSMOD, an open-source plugin for 3D Slicer designed for semi-automated segmentation and annotation of vascular structures in volumetric medical images. By combining interactive centerline modeling with automatic region-growing segmentation, VSMOD reduces annotation time by approximately 80% compared to conventional manual tools while preserving topological consistency. The plugin was validated on the PAT and applied successfully on the cerebrovascular network. VSMOD's annotations on the PERSEVERE dataset form the ground truth for the subsequent work presented in this thesis.

The second contribution addresses the lack of annotated data on the PERSEVERE cohort. Transfer learning experiments from a public healthy-subject dataset revealed a significant domain gap, confirming the necessity of a dataset-specific approach. We used a deep learning-assisted iterative annotation strategy that combines successive rounds of model training, automated prediction, and manual corrections. Using this process, we trained robust nnU-Net models for pulmonary artery and thrombi segmentation across all 353 patients.

The third contribution is a fully automated graph-based pipeline that transforms the artery, the thrombus, and the lungs masks into an anatomically oriented directed graph of the PAT. A graph-correction step significantly reduced non-physiological artifacts compared with existing tools. This results in a graph where each edge encodes vessel morphology, hierarchical location within the PAT, and PE obstruction, from which local and global biomarkers can be extracted across multiple hierarchical levels. The pipeline permits the automatic computation of adaptations of the Qanadli and Mastora thrombotic burden scores, achieving strong correlation with manual measurements, as well as the spatial distribution of embolic volume across lobes and hierarchical levels.

Finally, we applied the graph modeling framework to a neurovascular imaging context, vali-

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

dating the applicability of our modeling methodology on a different organ. The goal of this experiment was to quantitatively compare the vessel-segmentation capabilities on Dual-Energy CT and Spectral Photon Counting CT. This analysis was made possible by the morphological information encoded within the graph, demonstrating the broader applicability of the tools developed throughout this thesis.

As a whole, these contributions establish a complete and reproducible workflow from raw CTPA data to structured, clinically meaningful vascular biomarkers, laying the foundation for automated, patient-specific severity assessment of PE and for future prognostic studies in patients with PE.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Resumé

L'embolie pulmonaire (EP) est le troisième syndrome cardiovasculaire aigu le plus fréquent. Sa sévérité dépend non seulement de la charge thrombotique, mais également de sa distribution au sein de l'arbre artériel pulmonaire (AAP). Les scores de charge thrombotique existants, tels que ceux de Qanadli et de Mastora, reposent sur des modèles anatomiques simplifiés et des évaluations manuelles, limitant leur reproductibilité et leur utilisation en routine. Par ailleurs, les biomarqueurs cardiaques utilisés pour la stratification du risque peuvent être indisponibles en situation d'urgence. Face à ces limites, cette thèse propose un pipeline automatisé de modélisation de l'AAP par graphes à partir de l'angiographie pulmonaire, permettant l'extraction de biomarqueurs d'imagerie spécifiques au patient pour l'évaluation de la sévérité de l'EP. Les travaux ont été menés dans le cadre du projet PERSEVERE, à partir d'une cohorte de 353 patients.

La première contribution est VSMOD, un plugin open source pour 3D Slicer dédié à l'annotation semi-automatisée de structures vasculaires. En combinant une modélisation interactive de la ligne centrale avec une segmentation automatique par croissance de région, VSMOD réduit le temps d'annotation d'environ 80% tout en préservant la cohérence topologique. Le plugin a été validé sur des artères pulmonaires et appliqué avec succès à des artères cérébrales. Les annotations produites ont servi de vérité terrain pour les contributions suivantes.

La deuxième contribution porte sur l'apprentissage des modèles de segmentation en contexte d'annotations très limitées. Des expériences de transfert d'apprentissage à partir de données publiques de sujets sains ont mis en évidence un décalage de domaine par rapport aux données de sujets pathologiques, confirmant la nécessité d'une approche spécifique à la cohorte. Une stratégie d'annotation itérative assistée par apprentissage profond, combinant des cycles successifs d'entraînement de modèles, de prédiction automatique et de correction manuelle, a permis de répondre à cette limitation. Cette approche a conduit à l'entraînement de modèles nnU-Net robustes pour la segmentation des artères pulmonaires et des emboles sur l'ensemble de la cohorte.

La contribution principale de cette thèse est un pipeline de modélisation par graphes de l'AAP, transformant les masques des artères, des emboles et des poumons en un graphe orienté. Une étape de correction spécifique permet de réduire significativement les artefacts non physiologiques par rapport aux outils existants. Le graphe obtenu encode, pour chaque arête, la morphologie vasculaire, la localisation hiérarchique dans l'AAP et le degré d'obstruction, permettant l'extraction de biomarqueurs locaux et globaux à différents niveaux hiérarchiques. Ce pipeline permet de calculer automatiquement des adaptations des scores de charge thrombotique de Qanadli et de Mastora, fortement corrélées aux évaluations manuelles, ainsi que la distribution spatiale du volume embolique par lobe et par niveau hiérarchique.

Enfin, ce cadre de modélisation a été appliqué à l'imagerie neurovasculaire pour comparer

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

quantitativement des acquisitions issues de scanners à double énergie et à comptage photonique spectral. Cette analyse, fondée sur les informations morphologiques encodées dans le graphe, démontre la généricité de l'approche proposée.

L'ensemble de ces contributions établit un workflow complet et reproductible, allant de l'angiographie pulmonaire jusqu'à des biomarqueurs structurés et cliniquement interprétables, posant les bases d'une évaluation automatisée et personnalisée de la sévérité de l'embolie pulmonaire et d'études pronostiques futures chez les patients atteints d'EP.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Table of contents

Acknowledgments v

Acronyms xxv

Introduction 1

1 Medical context 5
- I Anatomical context . . . 6
  - I.1 Pulmonary anatomy . . . 6
  - I.2 Cerebral anatomy . . . 10
- II Pulmonary Embolism . . . 12
  - II.1 Physiological mechanisms . . . 12
  - II.2 Clinical signs . . . 13
  - II.3 Comorbidities . . . 13
- III Assessing PE severity . . . 14
  - III.1 Risk stratification . . . 14
  - III.2 Clinical Scores . . . 15
  - III.3 Thrombotic burden . . . 15
- IV PERSEVERE Project . . . 22
- V Conclusion . . . 23

2 Datasets and segmentation metrics 25
- I Datasets . . . 26
  - I.1 In-house dataset: the PERSEVERE dataset . . . 26
  - I.2 Pulmonary Artery Segmentation Challenge 2022 dataset . . . 33
- II Segmentation evaluation metrics . . . 36
  - II.1 Overlap metrics . . . 36
  - II.2 Topological metric . . . 38
- III Conclusion . . . 39

3 State of the art 41
- I Segmentation methods . . . 42
  - I.1 Deep learning architectures for medical image segmentation . . . 42
  - I.2 Lung and lobe segmentation . . . 47
  - I.3 PAT segmentation . . . 48
  - I.4 Thrombi segmentation . . . 50
- II From segmentation to graph-based modeling . . . 52
  - II.1 Centerline-based vascular modeling . . . 52
  - II.2 Graph construction and representation . . . 53
  - II.3 Applications of vascular graph models . . . 54

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

III Conclusion . . . 55

4 Development of 3D Slicer plugin **57**
I Limits of existing tools . . . 58
I.1 Preliminary annotation strategies . . . 58
I.2 Statement of need . . . 64
II VSMOD plugin . . . 64
II.1 Centerline module . . . 64
II.2 Segmentation module . . . 67
III Evaluation . . . 68
III.1 Reduction of annotation time . . . 68
III.2 Plugin usage on different vascular networks . . . 69
IV Conclusion . . . 69

5 Segmentation model training **71**
I Leveraging the PARSE public dataset . . . 72
I.1 Datasets and architectures . . . 72
I.2 Training configuration . . . 73
I.3 Results . . . 73
II PERSEVERE database segmentation . . . 77
II.1 Deep learning-assisted annotation strategy . . . 77
II.2 First iteration . . . 79
II.3 Second iteration . . . 82
II.4 Deployment on the full PERSEVERE database . . . 83
III Conclusion . . . 84

6 Graph modeling and Biomarker extraction **85**
I Methodology . . . 86
I.1 PAT graph modeling . . . 86
I.2 Hierarchical labeling . . . 91
I.3 PE feature extraction . . . 95
I.4 Patient level feature extraction . . . 97
II Results . . . 99
II.1 From voxel-graph to branch graph . . . 99
II.2 Graph error correction and comparison with VesselGraph pipeline . . . 101
II.3 Graph-based representation of the PERSEVERE cohort . . . 103
II.4 Hierarchical annotation of the PAT graph . . . 103
II.5 Thrombotic burden characterization through graph-derived biomarkers . . . 105
II.6 Computational time of the pipeline . . . 106
III Conclusion . . . 108

7 Experiment on neurovascular images **111**
I Study context . . . 112
I.1 Acquisition and reconstruction protocol . . . 113
I.2 Preprocessing . . . 113

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

II Methodology . . . 115
II.1 Vessel segmentation and ROI corrections . . . 115
II.2 Graph-based metric extraction . . . 115
III Results . . . 116
III.1 Subjective image quality analysis . . . 116
III.2 The circle of Willis . . . 116
III.3 Lenticulostriate arteries . . . 118
III.4 Coil surroundings . . . 118
IV Conclusion . . . 121

Conclusion **123**

Publication list **127**

Softwares and libraries **129**

Bibliography **131**

A Synthèse en français **139**
Introduction . . . 139
Chapitre 1 : Contexte médical . . . 141
Chapitre 2 : Jeu de données et métriques . . . 142
Chapitre 3 : État de l'art . . . 144
Chapitre 4: Développement d'un plugin semi-automatique d'annotation des structures vasculaires . . . 145
Chapitre 5 : Entraînement des modèles de segmentation vasculaire et des thrombus . 147
Chapitre 6 : Modélisation sous forme de graphe et extraction des biomarqueurs . . . 149
Chapitre 7 : Quantification de l'apport du SPCCT pour la visualisation des vaisseaux cérébraux . . . 151
Conclusion . . . 154

B Appendix **157**
I Excluded PERSEVERE patients list by criteria . . . 157
I.1 Different manufacturer . . . 157
I.2 Incomplete clinical blood biomarkers . . . 157
I.3 Image quality . . . 157
I.4 Lung and lobe segmentation . . . 158
II PERSEVERE dataset reconstruction kernels . . . 159
III PERSEVERE statistics . . . 160
IV PARSE statistics . . . 161
V Cerebral Artery Segmentation 2023 challenge . . . 162
V.1 CAS dataset . . . 162
V.2 CAS statistics . . . 163
V.3 CAS2023 Challenge . . . 164
V.4 Training efficiency . . . 168
VI Advanced parameters for RANSAC: the threshold . . . 169

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

VII Details of the quality check during the hierarchical labeling of the PAT 170

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

List of Figures

1.1 Illustration of the pulmonary anatomy. Encyclopædia Britannica, Inc. 6
1.2 Illustration of the pulmonary arterial tree hierarchical anatomy. 7
1.3 Variation in the common trunks of the RUL segmental pulmonary arteries. Figure extracted from Fourdrain et al. 2018. 8
1.4 Most common anatomic variations in the origins of the RUL segmental pulmonary arteries (A1, A2, A3). TA, truncus arteriosus. Figure extracted from Michaud et al. 2023. 9
1.5 Illustration of pulmonary anatomical structures on an axial view on a contrast-enhanced Computed Tomography Pulmonary Angiography from the patient PA000005 of the PARSE 2022 dataset (see Chapter 2 I.2). 9
1.6 Illustration of the Circle of Willis from MedlinePlus Medical Encyclopedia Image. . 10
1.7 Illustration of cerebral vascular structures on sagittal view on an MRA from patient 001 of the Brava dataset. 11
1.8 Illustration of pulmonary embolism mechanism. From Servier Medical Art by Servier, licensed under a Creative Commons Attribution 3.0 Unported License. 12
1.9 Table of risk stratification criteria to categorize patients. Figure extracted from Konstantinides et al. 2020. 14
1.10 Illustration of a proximal clot and a distal clot. From Servier Medical Art by Servier, licensed under a Creative Commons Attribution 3.0 Unported License. 16
1.11 Illustration of the location considered in the calculation of the Qanadli score. Modification of a Figure extracted from Alis and Haramati 2023 18
1.12 Illustration of the location considered in the calculation of the Mastora score. Modification of a Figure extracted from Alis and Haramati 2023 19
1.13 PERSEVERE project overview of the three main steps. 23

2.1 Summary of the number of excluded PERSEVERE patients by criteria. 27
2.2 Distribution of intensities of the SOMATOM Definition AS and SOMATOM Definition scanner. 28
2.3 Representative examples of images excluded from the cohort due to insufficient quality: (A) truncated field of view in patient 0041, (B) motion artifact in patient 0195, and (C) damaged lung parenchyma in patient 0252, displayed in 3D Slicer. . . 30
2.4 Images sizes distribution on the final 353 patients of the SOMATOM Definition AS and SOMATOM Definition scanner. 31
2.5 Illustration of the 10 sample patient case with an overview of the location of the thrombi. Each lobe is color-coded, and the interlobar portion corresponds to the part where an artery crosses from one lobe to another. The peripheral arteries correspond to the segmental level. 33
2.6 Images size distribution on the 90 PARSE patients. 34
2.7 Representative example of a CTPA volume with its corresponding pulmonary artery annotations for patient PA000005 from the PARSE training dataset, displayed in 3D Slicer. 35

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

2.8 Illustration of Betti numbers for 2D and 3D shapes. Figure by Pierre Rougé inspired by Stenseke 2021. . . . 38

3.1 Illustration of the 2D U-Net architecture. Figure extracted from Ronneberger, Fischer, and Brox 2015. . . . 42
3.2 Illustration of the Mean teacher architecture. Figure extracted from Tarvainen and Valpola 2018. The student model weights are θ and noise η, and the teacher model weights are θ' and noise η'. . . . 44
3.3 Illustration of the cascaded architecture implemented for cerebrovascular segmentation. Figure extracted from Rougé, Passat, and Merveille 2024. . . . 45
3.4 Illustration of the nnU-Net automatic configuration. Figure extracted from Isensee et al. 2021. . . . 46
3.5 Illustration of how a segmentation error, in this case, overestimating a vessel boundary, which will have little impact on the Dice score, while introducing a spurious branch in the reconstructed vascular tree. . . . 50
3.6 Illustration of two centerline-based graph representations from a vessel segmentation. The voxel graph, where each centerline voxel forms a node and edges connect neighboring voxels. The branch graph, where nodes correspond to vessel bifurcations and endpoints, and edges represent the vessel segments connecting them. . . . 53
3.7 Illustration of the graph construction pipeline applied to a PAT segmentation. Left: the 3D binary mask of the pulmonary arterial tree. Right: the corresponding graph representation obtained through the `VesselGraph` pipeline of the Voreen framework, where nodes mark vessel bifurcations and endpoints, and edges encode individual vessel branches. 54

4.1 Example of the lungmask models' segmentation results on patient PA000005 from the PARSE dataset . . . 59
4.2 Example of the TotalSegmentator models' segmentation results on patient PA000005 from the PARSE dataset . . . 59
4.3 Illustration of the region growing initialisation using `Draw tube` and operations on `lungmask` automatic segmentation to pulmonary vessel masks. . . . 61
4.4 Labeling arteries until the segmental level on patient PA000005 from the PARSE dataset . . . 62
4.5 Illustration of the test of PAT segmentation using RVesselX . . . 63
4.6 Pulmonary artery annotation workflow with VSMOD. Left-to-right: (A) RANSAC centerlines (B) automatic region-growing seeds (C) final connected segmentation. . 64
4.7 Centerline module . . . 65
4.8 Advanced RANSAC parameter available . . . 66
4.9 Segmentation module . . . 67
4.10 Brain vessel annotation workflow (MRA). Left-to-right: (a) RANSAC centerlines (b) automatic region-growing seeds (c) final connected segmentation demonstrating cross-organ applicability. . . . 69

5.1 Qualitative result of the 4 models on PARSE PA000078, PA000046, PA000285 patients. Dice values in red and clDice values in black. . . . 74
5.2 Qualitative result of the 4 models on PERSEVERE 0060 and 0128 patients. Dice values in red and clDice values in black. . . . 75

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

5.3 Active-learning-inspired iterative annotation protocol used to obtain reliable pulmonary artery and thrombi segmentations across the PERSEEVERE database. . . 77
5.4 Illustration of the four input configurations used for nnU-Net training across both iterations. Models differ in the anatomical context provided as segmentation targets. 78
5.5 The 3 nnU-Net configurations automatically generated and their pairwise ensemble evaluated for each model on our PERSEVERE dataset. 79
5.6 Thrombi segmentation results on the two first-iteration PERSEVERE test patients: patient 0031 (top) and patient 0216 (bottom). Predictions are shown in red and the ground truth in green. 80

6.1 General workflow of the proposed pipeline. A) Segmentation of arteries, lungs, lobes and thrombi described in Chapter 5. B) Graph modeling and feature extraction pipeline. Steps (1-3) generate the graph and produce local biomarkers, which are subsequently aggregated in step (4) to produce patient-level biomarkers. 86
6.2 Preprocessing of the artery mask, to extract the centerline and the distance map. . 87
6.3 Graph construction steps.The (A) arteries mask is skeletonized, from which we obtain the centerline of the arteries (B). The centerlines are transformed into a voxel-level graph (C) that is then converted into the branch-level graph (D). 88
6.4 Illustration of the cleaning process, detecting a spurious vessel caused by an irregularity on the surface of the mask, and a cycle caused by contact between two close arteries. 89
6.5 Illustration of an artifact or a spurious vessel caused by an irregularity on the surface of the mask. 90
6.6 Illustration of a cycle caused by the contact of two vessels by a few voxels, and the resulting constriction pattern. 90
6.7 Hierarchical levels of pulmonary arteries, each level is presented with a different color. 92
6.8 Laterality labeling, obtained by superposition from the right and left lungs masks . 93
6.9 Lobes labeling, obtained by superposition of the lobes masks with the terminal nodes and propagating the labels upwards until level 2. RUL: Right Upper Lobe; RML: Right Middle lobe; RLL: Right Lower Lobe; LUL: Left Upper Lobe; LLL: Left Lower Lobe. 93
6.10 Illustration of the morphological criteria evaluated to determine if an artery is a trunk. 94
6.11 Example of the Total Embolic Volume (TEV) computed on patient 0055. 95
6.12 Thrombus intersection with each edge it traverses. (A) a representation of the Right Upper Lobe view with the artery mask (green) and the modelized graph. (B) The volume of the artery, with each artery volume labeled with the vessel ID. (C) A clot inside the artery. (D) The same clot volume intersected with each corresponding artery it traverses, allowing us to associate the values with each corresponding edge. 96
6.13 Representation of all the local features aggregated (vessels, location, and PE) at the edge level at the end of the 3 main steps of the pipeline (1) PAT modeling, 2) Hierarchical labeling, and 3) PE characterization. 97
6.14 Illustration of the anatomical positions of the arteries considered by both Qanadli and Mastora scores. Modification of a Figure from Alis and Haramati 2023 98

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

6.15 Thrombus intersection with each hierarchical level it traverses in the RUL. (A) RUL view with the artery mask (green) and the modelized graph. (B) The volume of the artery, with each artery volume labeled with the associated hierarchical level. (C) A clot inside the artery. (D) The same clot volume intersected with each corresponding artery it traverses, allowing us to associate the hierarchical level with each corresponding edge. . 100
6.16 Illustration of the artifact cleaning step on patient 0028 from the annotated PERSEVERE sample. (A) The arteries mask (green) and the graph before cleaning, visualized with 3D Slicer. (B) Only the graph from before the cleaning process, with a highlight of the Main Pulmonary Artery (MPA) with one spurious branch. (C) The graph from `VesselGraph` is visualized from their interface. The highlight shows many remaining artifacts on the MPA. (D) Our graph, after the correction, shows no artifact remains on the MPA. 102
6.17 Bland-Altman for Qanadli manual VS automatic scores. 105
6.18 Bland-Altman for Mastora manual VS automatic scores. 105
6.19 TEV repartition by lobes and hierarchical levels provides a spatial description of embolic burden that can distinguish proximal from distal PE patterns with identical global clot volume. 107

7.1 Overview of the 4 image modalities. (A) Dual Energy polychromatique (B) SPCCT polychromatique (C) Dual energy 40 keV (D) SPCCT 40 keV 114
7.2 Segmentations in green of (A) Dual Energy polychromatique and (B) SPCCT polychromatique with the four ROIs in highlighted in red. 115
7.3 Axial Maximum-Intensity Projections (MIP) with 15 mm slab thickness, and zoomed slices for DECT-p, DECT-40keV, SPCCT-p and SPCCT-40keV with corresponding segmentations. Red arrow indicated a perforating artery of the basilar trunk. 117
7.4 Coronal maximum-intensity projections (MIP, 15mm slab thickness), and zoomed slices for DECT-p, DECT-40keV, SPCCT-p and SPCCT-40keV with corresponding segmentations. Arrows in blue highlights the possibility of better discrimination between arteries and veins with 40kev. 119
7.5 Sagittal oblique maximum-intensity projections reconstruction (15 mm slab thickness) and zoomed slices for arterial lumen visualization with a coil in DECT-p, DECT-40keV, SPCCT-p, and SPCCT-40keV images with corresponding segmentations. Arteries are in red and coils in blue. 120

A.1 Visualisation globale des objectifs du projet PERSEVERE 142
A.2 Illustration des 10 patients sélectionnés pour l'échantillon avec une vue d'ensemble de la localisation des thrombus. Chaque lobe est représenté par une couleur, et la portion interlobaire correspond à la zone où une artère traverse d'un lobe à un autre. Les artères périphériques correspondent au niveau segmentaire. 143
A.3 Illustration de deux représentations par graphes basées sur la ligne centrale, extraites d'une segmentation vasculaire. 145
A.4 Protocole d'annotation itératif inspiré de l'apprentissage actif, utilisé pour obtenir des segmentations fiables de l'arbre artériel pulmonaire et des thrombus sur l'ensemble de la base de données PERSEVERE. 148

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

A.5 Schéma général du pipeline proposé. A) Segmentation des artères, des poumons, des lobes et des thrombi décrite dans le Chapitre 5. B) Pipeline de modélisation par graphe et d'extraction de caractéristiques. Les étapes (1-3) génèrent le graphe et produisent des biomarqueurs locaux, qui sont ensuite agrégés lors de l'étape (4) pour produire des biomarqueurs à l'échelle du patient. 150
A.6 Reconstructions en projection d'intensité maximale (MIP) selon un plan sagittal oblique (épaisseur de dalle de 15 mm) et coupes agrandies pour la visualisation de la lumière artérielle en présence d'un coil, sur les images DECT-p, DECT-40keV, SPCCT-p et SPCCT-40keV avec les segmentations correspondantes. Les artères sont représentées en rouge et les coils en bleu. 153

B.1 Illustration of the lungs and lobes segmentation from Lungmask and Totalsegmentator on patient 0028 from the PERSEVERE dataset 158
B.2 Shape distribution on the 100 CAS patients 162
B.3 Representative example of a TOF MRA volume with its corresponding cerebral artery annotations for patient 000 from the CAS training dataset, displayed in 3D Slicer. 163
B.4 Illustration of the preprocessing on the patient 0000 of the training set. 164
B.5 Illustration of a point cloud between 0.8 and 1.2mm from a cylinder axis, and the impact of the *Threshold (%)* parameter on the *Inlier Points (%)* parameter. Figure done by Gabriel Husak 169

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

List of Tables

1.1 Variety in the number of segmental arteries per lobe in a cohort of 420 patients, as reported by Ma et al. 2023. . . . 8
1.2 Summary of the main thrombotic burden scoring systems. . . . 20

2.1 Distribution of patients per manufacturer and scanner model. . . . 27
2.2 Global intensity statistics per scanner model across the 414 retained patients. . . . 28
2.3 Patient distribution according to the ESC 2019 risk stratification categories (Konstantinides et al. 2020). . . . 31
2.4 Sample of annotated patients with clinical and image quality criteria. Troponine and BNP values are reported upon admission to the Emergency unit. Artifact scores: 0 = absent, 1 = present but not affecting interpretation, 2 = present and affecting interpretation. Injection quality: 0 = does not meet criteria, 1 = meets criteria. Global image quality: 0 = non-interpretable, 1 = interpretable with reservations, 2 = interpretable without reservations. . . . 32
2.5 Confusion Matrix for the reference A and the prediction B . . . 36

4.1 Average segmentation times per patient for a pulmonary vascular network with 269 vessels on average. . . . 69

5.1 Training data and architecture used for each model. . . . 72
5.2 Performance comparison of all four models on the PARSE test set ($n = 10$ patients) and the PERSEVERE test set ($n = 2$ patients). Values are shown as mean ± std. . . . 73
5.3 Mean foreground Dice scores of each nnU-Net configuration and ensemble for the first iteration (validation results). Bold values indicate the best configuration for each model. . . . 81
5.4 Performance of artery and thrombi segmentation models: mean Dice score on test datasets (first iteration). Bold values indicate the best-performing model per class. 81
5.5 Mean foreground Dice scores of each nnU-Net configuration and ensemble for the second iteration (validation results). Bold values indicate the best configuration for each model. . . . 82
5.6 Performance of artery and thrombi segmentation models: mean Dice score on test datasets (second iteration). Bold values indicate the best-performing model per class. 83

6.1 Graph construction main statistics from voxel-level to branch-level representation. . 101
6.2 Construction and cleaning statistics on the 24 PERSEVERE annotated sample. . . 101
6.3 Topological statistics of the final arterial segmentations (353 patients). . . . 103
6.4 Comparison of mean detected segmental arteries (± Std) for different morphological thresholds. . . . 104
6.5 Mean number of segmental arteries for 24 annotations and to the 353 segmentations 104
6.6 Bland-Altman and Spearman correlation . . . 106
6.7 Per-patient execution time across the PERSEVERE dataset (353 patients). . . . 108

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

7.1 Technical specifications of the images used in the study. . . . 113
7.2 Subjective analysis of image quality in a 68-year-old man who underwent cerebral CT angiography using dual-energy CT (DECT) and a clinical prototype spectral photon-counting CT (SPCCT). . . . 116
7.3 Circle of Willis: Post-correction metrics. With SPCCT modality we detected 2× more vessels, down to 0.62 mm. . . . 118
7.4 Lenticulostriate arteries (SPCCT only; undetectable on DECT). . . . 118
7.5 Artery and coil near implant. . . . 121

A.1 Temps de segmentation moyens par patient pour un réseau vasculaire pulmonaire contenant en moyenne 269 vaisseaux. . . . 147

B.1 Distribution of reconstruction kernels by SOMATOM Definition AS and SOMATOM Definition scanner model. . . . 159
B.2 Descriptive statistics of arterial segmentation volumes and topological Betti numbers across the full dataset ($n = 353$ patients). . . . 160
B.3 Descriptive statistics of arterial segmentation volumes and topological Betti numbers across the PARSE dataset($n = 90$ patients). . . . 161
B.4 Descriptive statistics of cerebral artery segmentation volumes and topological Betti numbers across the CAS2023 dataset ($n = 100$ patients). . . . 163
B.5 Results Summary for Single Models folds, Single-Model Ensembling, Multi-model per fold ensembling, and Multi-model per-model Ensembling on 20-patient test set . 166
B.6 Post-Processing Optimization: Impact of Small Object Removal on Tiny+Base+WNet Ensemble . . . 166
B.7 Team rankings and performance metrics of CAS 2023 challenge. . . . 167
B.8 Training efficiency comparison across four models . . . 168

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Acronyms

BNP Brain Natriuretic Peptide

CAS Cerebral Artery Segmentation

CFD Computational Fluid Dynamics

CI Confidence Interval

clDice centerline Dice

CT Computed Tomography

CTPA Computed Tomography Pulmonary Angiography

DECT Dual-energy CT

Dice Dice similarity coefficient

LAs Lenticulostriate Arteries

LLL Left Lower Lobe

LPA Left Pulmonary Artery

LUL Left Upper Lobe

LV left ventricle

MPA Main Pulmonary Artery

MRA Magnetic Resonance Angiography

OR Odds Ratio

PARSE Pulmonary ARtery SEgmentation Challenge

PAT Pulmonary Arterial Tree

PE Pulmonary Embolism

PERSEVERE Pulmonary Embolism Risk Stratification basEd on Vascular nEtwoRk modEling

PESI Pulmonary Embolism Severity Index

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

RLL Right Lower Lobe

RML Right Middle Lobe

ROI region of interest

ROIs regions of interest

RPA Right Pulmonary Artery

RUL Right Upper Lobe

RV right ventricle

RV/LV right-to-left ventricular diameter ratio

SPCCT Spectral photon-counting CT

sPESI Simplified Pulmonary Embolism Severity Index

TEV Total Embolic Volume

TOF-MRA Time-of-Flight Magnetic Resonance Angiography

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Introduction

Context and objectives

Pulmonary Embolism (PE) is a life-threatening cardiovascular condition caused by the obstruction of pulmonary arteries by one or several blood clots, also called thrombi. The severity of PE is highly dependent on the location, extent, and distribution of clots within the Pulmonary Arterial Tree (PAT). The thrombi can cause right ventricle (RV) dysfunction, until failure, which can lead to obstructive shock and death. PE is the third most frequent acute cardiovascular syndrome after myocardial infarction, and longitudinal studies have shown an upward trend in the annual incidence rates of PE over time. Accurate severity assessment is essential to guide treatment but remains challenging in clinical practice. Diagnosis primarily relies on Computed Tomography Pulmonary Angiography (CTPA), which provides detailed visualization of the pulmonary vasculature. Once PE is confirmed, patients are stratified into three risk categories (high, intermediate, low) based on their 30-day mortality risk. This stratification determines treatment strategy, ranging from anticoagulation alone to more aggressive interventions such as thrombectomy, as well as the appropriate care setting (outpatient, general ward, or intensive care unit).

Current guidelines first assess hemodynamic instability. In stable patients, RV dysfunction is evaluated using three biomarkers: one morphological marker, the right-to-left ventricular diameter ratio (RV/LV) measured on CTPA, and two functional markers, blood levels of troponin and Brain Natriuretic Peptide (BNP). However, this approach has several limitations. Blood biomarkers are not always available in emergency settings and reflect the consequences of PE rather than the thrombi themselves. For instance, RV dilation results from increased pressure in the pulmonary arteries, indirectly indicating clot presence. In addition, CTPA is not synchronized with the cardiac cycle, leading to motion artifacts that can impair accurate estimation of the RV/LV. Biomarkers are also released only after RV dysfunction has occurred.

To better characterize thrombotic burden, several scores have been proposed. These scores rely on the hierarchical organization of the PAT, where lobar arteries divide into segmental branches. For example, the Qanadli score considers the number of affected segmental arteries, while the Mastora score evaluates the degree of obstruction in each artery. However, these scores are based on simplified vascular models with a fixed number of segmental arteries per lobe. Although they provide a detailed assessment of clot burden, they are not used in routine practice due to their time-consuming manual computation and their inability to account for patient-specific anatomical variability. Indeed, the pulmonary vasculature exhibits substantial inter-individual variability, particularly at the segmental level, both in the number of arteries and in their branching patterns.

In this context, the **Pulmonary Embolism Risk Stratification basEd on Vascular nEtwoRk modEling (PERSEVERE) ANR project** aims to develop computational tools for

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

automated PE severity assessment from CTPA images. This project includes an in-house dataset of 431 clinical CTPA scans from patients with PE. A key limitation is the absence of annotations for both thrombi and the PAT. The objective of this thesis, conducted within the PERSEVERE project, is to design and evaluate an end-to-end pipeline for automated PE severity assessment. This pipeline combines deep learning-based segmentation, graph-based modeling of the PAT, and automated biomarker extraction. It aims to quantify both local biomarkers (vessel geometry, clot characteristics, anatomical location) and global biomarkers summarizing thrombotic burden at the patient level. This requires accurate 3D segmentation of arteries and thrombi, followed by the construction of a graph representation capturing the topology and hierarchy of the PAT. Such a representation enables the extraction of patient-specific, anatomically grounded biomarkers directly from CTPA, which is routinely available in clinical practice.

Achieving this goal requires addressing several interconnected challenges.

First, **the lack of annotations** in the PERSEVERE dataset prevents direct supervised training. This is further complicated by pathological alterations in PE, such as occlusions and perfusion defects, which make vascular structures harder to delineate. A dedicated annotation strategy is therefore required.

Second, even with annotations, **training robust segmentation models** remains difficult due to the limited amount of labeled data and the domain gap between public datasets (mostly healthy subjects) and pathological CTPA scans. This impacts generalization and requires adapted training strategies.

Third, **extracting clinically relevant biomarkers** requires a structured and anatomically consistent representation of the vascular network. Converting segmentation masks into a topologically correct graph is challenging, as errors may introduce disconnections or loops that violate the expected tree structure of the PAT. Ensuring topological consistency is therefore essential for reliable biomarker computation.

Beyond the pulmonary application, the **generalizability** of the framework can be evaluated in neurovascular imaging. The cerebrovascular network, with its tortuous geometry, provides a complementary setting to assess the robustness of the proposed annotation and graph-based vascular modeling approaches.

Contributions

In the previous section, we introduced the clinical and scientific context of this thesis, along with its main objective and challenges. Here, we summarize the contributions of this work, organized into four main axes:

- **Annotation methodology and dataset construction.** We developed VSMOD, an open-source plugin for 3D Slicer, enabling semi-automated segmentation and modeling of vascular structures in volumetric images. By combining interactive centerline model-

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

ing with region-growing segmentation, it reduces annotation time by about 80% while preserving vascular connectivity through a graph representation. Although designed for PAT in CTPA, it generalizes to other anatomies such as cerebrovascular structures. This work resulted in a JOSS publication, and the tool is publicly available. The annotations generated with VSMOD on the PERSEVERE dataset serve as ground truth for the rest of the thesis.

- **Deep learning-based segmentation of pulmonary arteries and thrombi.** We trained segmentation models on the annotated PERSEVERE dataset. Initial experiments on transfer learning from the Pulmonary ARtery SEgmentation Challenge (PARSE) dataset highlighted a domain gap between healthy and pathological data, leading to a dataset-specific approach. An iterative strategy combining model training, prediction, and manual correction improved annotation quality and quantity. This enabled training robust nnU-Net models, achieving Dice scores of 0.943 for arteries and 0.817 for thrombi.
- **Graph-based pipeline for PE biomarker extraction.** We proposed an automated pipeline that represents the PAT as a directed graph derived from segmentations and extracts clinically relevant biomarkers. The pipeline includes graph construction and correction, hierarchical vessel labeling, and aggregation of vessel-level features into patient-level metrics. It enables automatic computation of thrombotic burden scores such as Qanadli and Mastora, as well as Total Embolic Volume (TEV) distribution. Strong correlations with manual scores (Spearman $\rho = 0.83$ and 0.86) demonstrate clinical relevance. This work has been submitted to JBHI.
- **Application to neurovascular imaging with SPCCT.** We demonstrated the broader applicability of our graph-based framework to another vascular territory. The pipeline is applied to cerebrovascular segmentations from DECT and SPCCT, enabling quantitative comparison of segmentation quality across modalities in selected regions of interest (ROIs). This experiment shows that the proposed graph-based approach generalizes to vascular structures with different geometries and topologies. This work has been submitted and accepted for publication in the Journal of Neurobiology.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

1 Medical context

Table of contents

I Anatomical context . . . 6
I.1 Pulmonary anatomy . . . 6
I.2 Cerebral anatomy . . . 10
II Pulmonary Embolism . . . 12
II.1 Physiological mechanisms . . . 12
II.2 Clinical signs . . . 13
II.3 Comorbidities . . . 13
III Assessing PE severity . . . 14
III.1 Risk stratification . . . 14
III.2 Clinical Scores . . . 15
III.3 Thrombotic burden . . . 15
IV PERSEVERE Project . . . 22
V Conclusion . . . 23

This first chapter introduces the clinical and anatomical context necessary for the medical imaging analyses conducted during the thesis. We will first present the anatomical context of the lungs and pulmonary vasculature. We will describe the hierarchical structure of the Pulmonary Arterial Tree (PAT) and its anatomical variations as it motivates the graph-based modeling approach developed in Chapter 6. A brief introduction to the cerebrovascular anatomy follows, which serves as a topologically contrastive structure used in experiments. The chapter subsequently presents the medical context of Pulmonary Embolism (PE), covering its physiological mechanism, clinical consequences, and current diagnosis approaches. Finally, we discuss the challenges surrounding PE severity assessment, especially the measurement of thrombotic burden, which motivated the Pulmonary Embolism Risk Stratification basEd on Vascular nEtwoRk modEling (PERSEVERE) project and frames this thesis's contributions.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I Anatomical context

This section explores the vascular anatomy of the lungs and brain, highlighting key structural features visible in Computed Tomography Pulmonary Angiography (CTPA) and Time-of-Flight Magnetic Resonance Angiography (TOF-MRA) .

I.1 Pulmonary anatomy

This section describes the pulmonary anatomy and the vascular structures visible on CTPA images.

I.1.1 Hierarchical anatomy

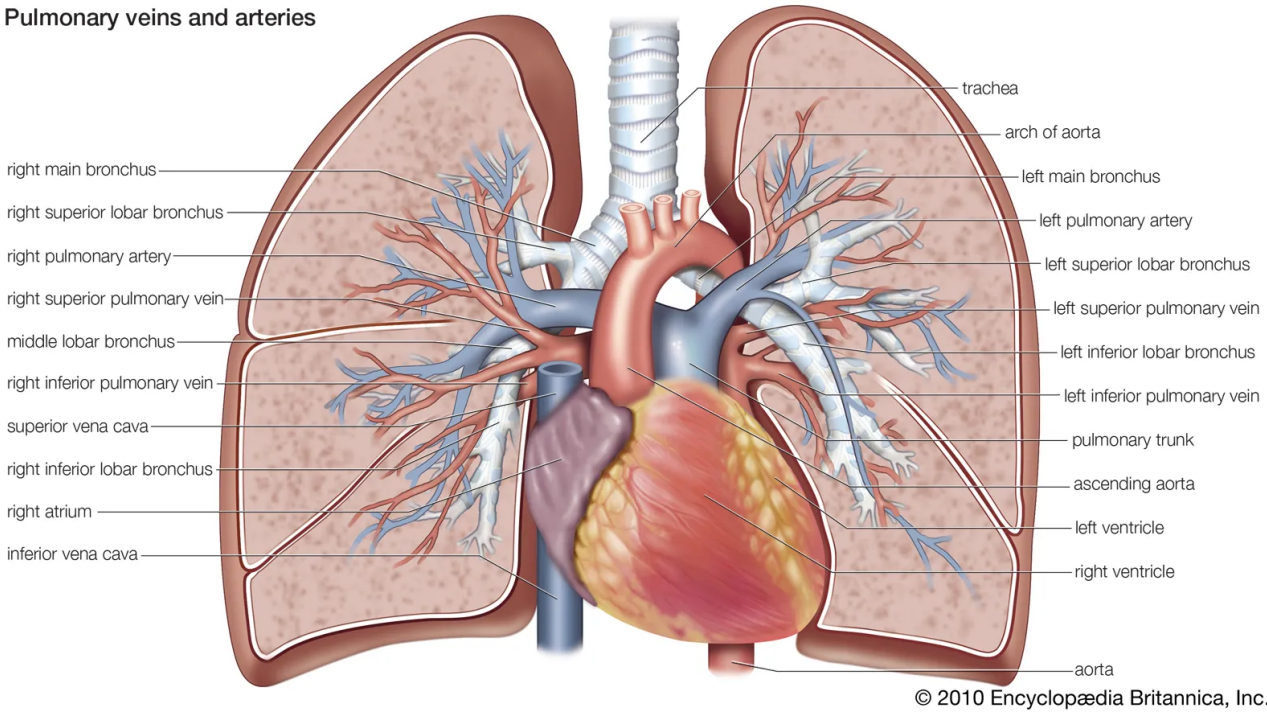


Fig. 1.1: Illustration of the pulmonary anatomy. Encyclopædia Britannica, Inc.

The general anatomical context of the respiratory system is presented in Figure 1.1. The right lung is divided into three lobes (on the left of the figure) and the left lung into two lobes (on the right of the figure). Each lobe is separated by interlobar fissures. The deoxygenated blood from the body is pumped by the right ventricle (RV) into the Main Pulmonary Artery (MPA) (the pulmonary trunk represented in blue in Figure 1.1), and from there to the whole PAT. As a size reference, the MPA diameter for a healthy subject is 25.1 ± 2.8mm (Truong et al. 2012). The MPA divides into the Right Pulmonary Artery (RPA) and the Left Pulmonary Artery (LPA), which each penetrate the right and left lungs at what is called the hilium. Inside each lung, the RPA and LPA divide into lobar arteries to irrigate each lobe. The lobar arteries then divide into smaller segmental arteries, whose number varies by lobe. The classic anatomical representation of the PAT adopted by the clinicians assigns a fixed number of segmental arteries per lobe,

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

corresponding to the subdivision of each lobe into segments: 3 in the Right Upper Lobe (RUL), 2 in the Right Middle Lobe (RML), 5 in the Right Lower Lobe (RLL), 5 in the Left Upper Lobe (LUL), and 4 in the Left Lower Lobe (LLL), for a total of 19 segments (Michaud et al. 2023). Segmental artery nomenclature depends on the segment of the lung it irrigates: apical, posterior, anterior, lateral, medial, or superior, which can be simplified as A1, A2, A3, etc., as illustrated in Figure 1.2. After the segmental level, the arteries continue to branch into smaller subsegmental arteries and continue until they become capillaries. The successive division of arteries from the MPA to the capillaries forms a hierarchical, arborescent vascular network, with mostly linear, straight vessels and, more importantly, no cycles.

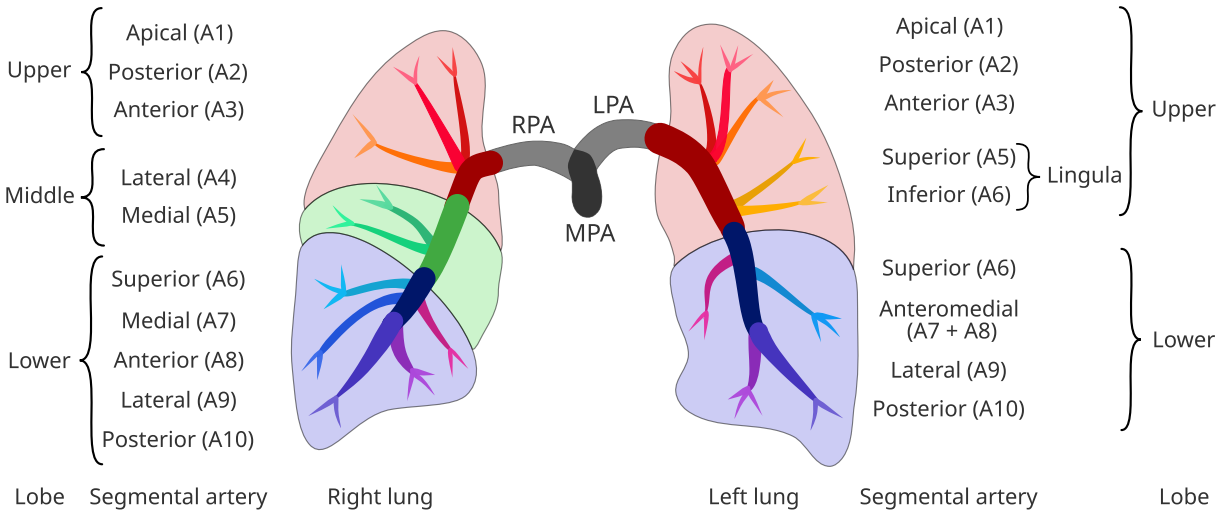


Fig. 1.2: Illustration of the pulmonary arterial tree hierarchical anatomy.

I.1.2 Anatomical variations

While the classical PAT representation assigns 19 segmental arteries, some clinical scores simplify the PAT segmental organization to 10 segmental arteries per lung, counting one additional in LLL, for a total of 20 (see III.3.1).

Ma et al. 2023 cataloged the variations by the number of possible segments per lobe in a larger cohort of 420 patients as presented in Table 1.1. Notably, the LUL exhibits the most variability in the number of segmental arteries, ranging from 2 to 7, while all other lobes show a more limited range from 1 or 2 up to a maximum of 4 segmental arteries.

In addition to the possible variation of the number of segmental arteries, the PAT can have a wide variety of branching patterns at the segmental level. A study by Fourdrain et al. 2018 done on 44 patients reported 11 variations for the RUL, 4 in the RML, 17 in the RLL, 25 in the LUL, and 18 in LLL. They study the variations in common trunks for different segmental arteries, sometimes not even of the same lobe. For example, as we can see on the second illustration in Figure 1.3, we can observe a common trunk between the Anterior segmental artery (A3) of the RUL with an artery going into the RML .

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Number of segmental arteries	Left		Right		
	Upper	Lower	Upper	Middle	Lower
	Number of patients (%)	Number of patients (%)	Number of patients (%)	Number of patients (%)	Number of patients (%)
1	–	–	32 (7.6)	128 (30.5)	–
2	6 (1.4)	334 (79.5)	327 (77.9)	269 (64.0)	332 (79.0)
3	107 (25.5)	82 (19.5)	59 (14.0)	21 (5.0)	86 (20.5)
4	212 (50.5)	4 (1.0)	2 (0.5)	2(0.5)	2 (0.5)
5	91 (21.6)	–	–	–	–
6	2 (0.5)	–	–	–	–
7	2 (0.5)	–	–	–	–

Tab. 1.1: Variety in the number of segmental arteries per lobe in a cohort of 420 patients, as reported by Ma et al. 2023.

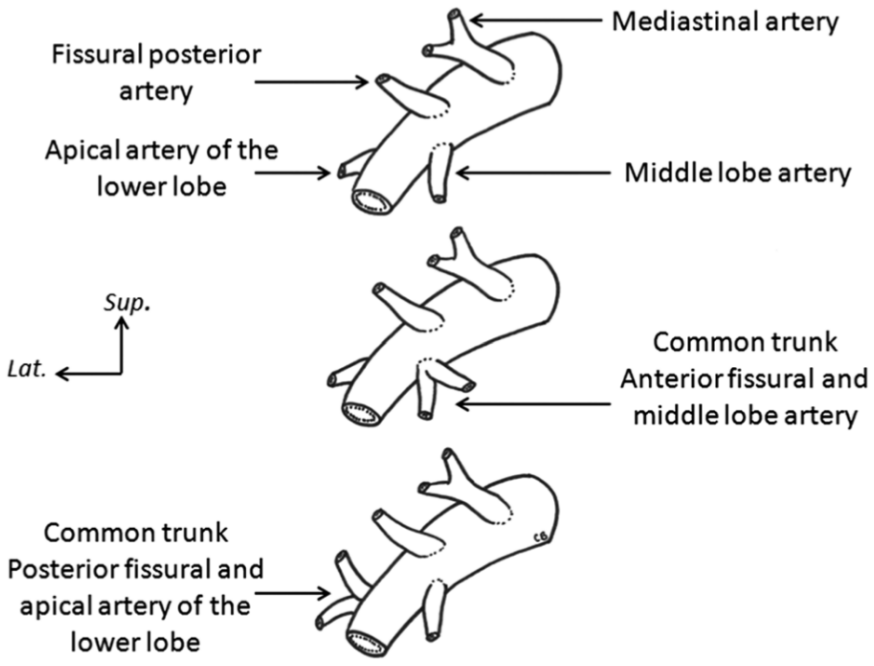


Fig. 1.3: Variation in the common trunks of the RUL segmental pulmonary arteries. Figure extracted from Fourdrain et al. 2018.

Michaud et al. 2023 documented inter-individual variations of segmental artery branching patterns inside each lobe, based on a cohort of 179 patients. In the Figure 1.4, we can see three possible branching patterns among the seven they observed for the three RUL segmental arteries. These three patterns account for 92.9% of their cohort. They do not describe the four remaining possible branching patterns.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

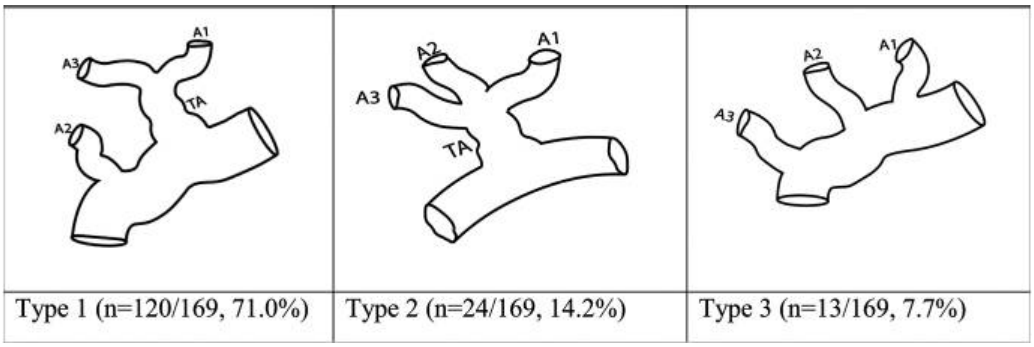


Fig. 1.4: Most common anatomic variations in the origins of the RUL segmental pulmonary arteries (A1, A2, A3). TA, truncus arteriosus. Figure extracted from Michaud et al. 2023.

Finally, pathology-induced deformations such as proximal dilatation upstream of large clots can alter the apparent geometry of the arteries (Raju et al. 2021).

I.1.3 PAT in contrast-enhanced CTPA

In clinical practice, the vascular structures of the PAT can be observed using contrast-enhanced CTPA (see Figure 1.5). Computed tomography is based on X-rays. Multiple X-ray beams are sent as the scanner rotates around the patient. Depending on the density, the tissues absorb the X-ray differently. The lower the density, the more the radiation will pass through, and the darker the tissue will appear. The higher the density, the more X-rays are absorbed, and the brighter the tissue will appear. Consequently, bones appear in hyper-intensity, whereas the air in the lungs appears in hypo-intensity. In contrast-enhanced angiography, a contrast agent is injected into the patient's bloodstream to make the vessels appear in hyper-intensity.



Fig. 1.5: Illustration of pulmonary anatomical structures on an axial view on a contrast-enhanced Computed Tomography Pulmonary Angiography from the patient PA000005 of the PARSE 2022 dataset (see Chapter 2 I.2).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

The tree-shaped, acyclic, and hierarchically labeled structure makes the PAT particularly well suited to a graph-based representation, where each vessel segment corresponds to an edge, and each bifurcation or end of an artery, a node. This key property motivates a graph modeling approach. Given the variability in the number of segmental arteries and the complexity of their branching patterns, assigning a specific segmental anatomical label to each segmental artery (e.g., A1, A2, A3) is not always feasible. Our approach, therefore, focuses on identifying the hierarchical depth level of each artery, distinguishing lobar, segmental, and sub-segmental levels.

I.2 Cerebral anatomy

This part describes the cerebral arterial network, and illustrates it on Magnetic Resonance Angiography (MRA). Although the primary focus of this thesis is the PAT, the cerebrovascular network is introduced here because it serves as a topologically contrasting structure used to evaluate if the methods developed for the PAT can be generalized to topologically and geometrically different networks (see Chapter 4 and 7).

While the PAT is a cycle-free tree, the cerebral arterial network contains a cycle: the Circle of Willis (see Figure 1.6). The circle of Willis connects the anterior and posterior cerebral circulation at the base of the brain.

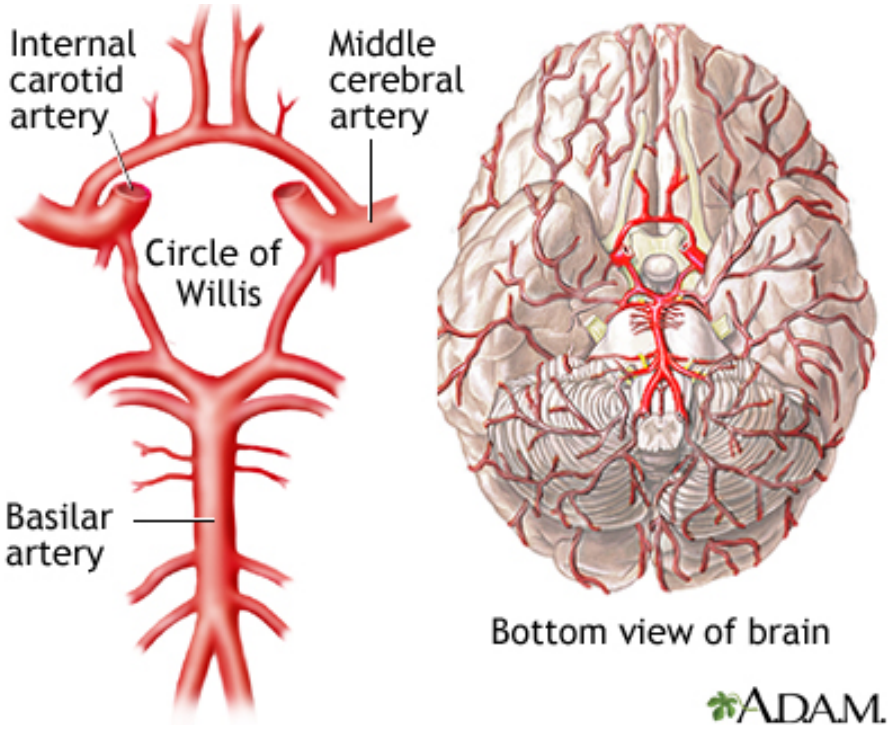


Fig. 1.6: Illustration of the Circle of Willis from MedlinePlus Medical Encyclopedia Image.

In the PAT, the arterial caliber decreases rapidly and consistently from the MPA (approximately 25mm) down to subsegmental arteries (diameter of 1-2 mm). In contrast, in the cerebral arterial network, the reduction of diameter is more gradual. The main cerebral arteries, such as the

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

internal carotid arteries, start with a mean diameter of 5 ± 0.6 mm, the middle cerebral artery has a mean diameter of 3.1 ± 0.4mm (Rai et al. 2013), and the caliber of distal branches decreases much more slowly.

In contrast to the predominantly straight vessel trajectories observed in the PAT, cerebral arteries follow highly curved, winding paths, exhibiting greater tortuosity. This geometric difference provides a challenging setting in which to evaluate the generalizability of two contributions developed in this thesis. First, the annotation plugin VSMOD (Chapter 4), initially designed and validated on pulmonary arteries, was also applied to cerebral vasculature to validate whether its centerline-based annotation workflow remains usable when vessel paths are no longer approximately straight but more tortuous. Second, the graph-based modeling and metric extraction step of our developed pipeline (Chapter 6) was applied to cerebral artery segmentation to verify that the results remain reliable for more tortuous, non-linear structures. Also, the quantitative information extracted from these cerebral segmentations formed the basis of an analysis comparing vessel-detection capabilities across two image modalities (Chapter 7).

In this work, we visualized cerebral arteries with MRA. Contrary to CTPA, MRA does not rely on radiation but on the magnetic field and radio waves. In figure 1.7 we observe a Time-of-flight MRA of a patient from the Brava dataset [1] (Wright et al. 2013). The static tissues absorb the radio wave pulses and appear saturated (dark) on the image. On the contrary, because the blood flow is constantly moving, it is not saturated, so it produces a stronger signal (white) on the image.

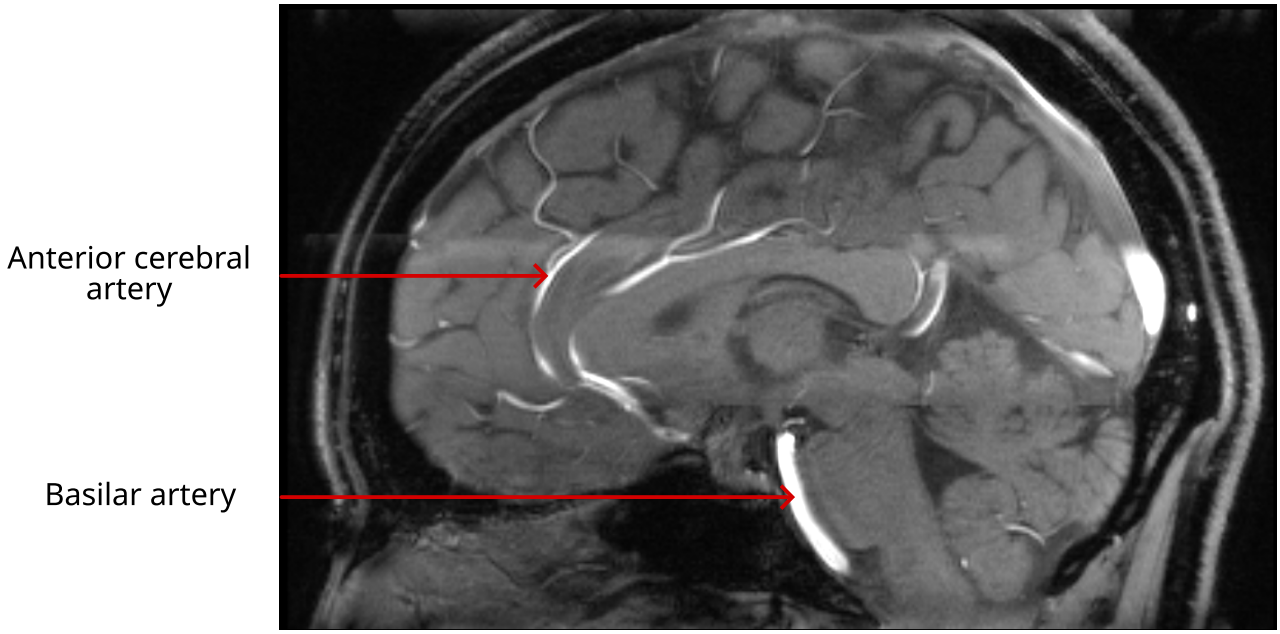


Fig. 1.7: Illustration of cerebral vascular structures on sagittal view on an MRA from patient 001 of the Brava dataset.

[1] http://cng.gmu.edu/brava

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

II Pulmonary Embolism

In this section, we will present the Pulmonary Embolism (PE), a medical condition caused by obstruction of one or more pulmonary arteries by a blood clot (thrombus). This condition represents the third most frequent acute cardiovascular syndrome, behind myocardial infarction and stroke (Konstantinides et al. 2020). Incidence estimates vary widely across populations, ranging from 39 to 115 cases per 100,000 (Wendelboe and Raskob 2016).

II.1 Physiological mechanisms

A thrombus forming in a peripheral vein detaches and travels through the venous circulation, passes through the right heart chambers, and ultimately obstructs a pulmonary artery (see Figure 1.8). These clots most commonly originate from deep vein thrombosis in the lower extremities, though they can also form in other parts of the venous system.

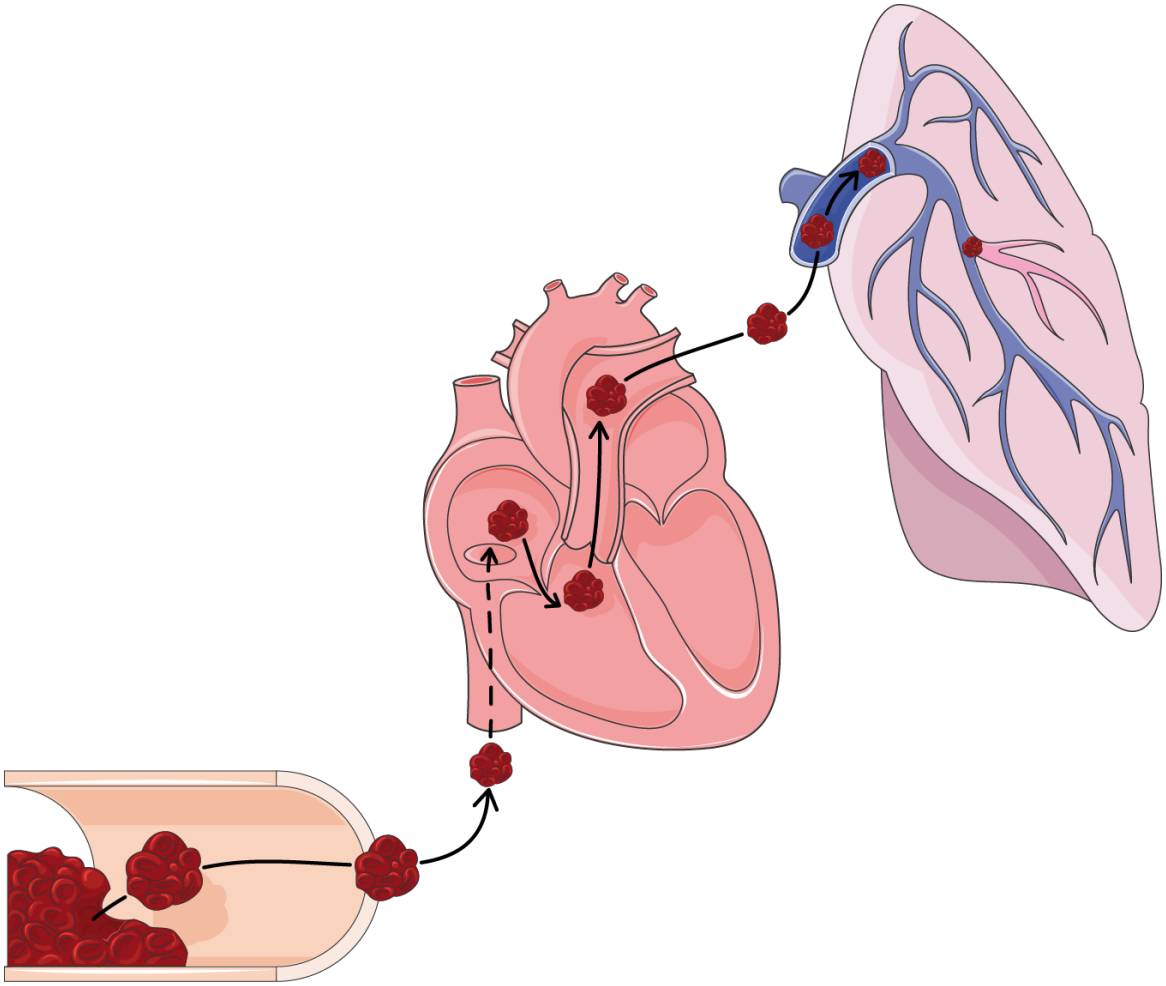


Fig. 1.8: Illustration of pulmonary embolism mechanism.
From Servier Medical Art by Servier, licensed under a Creative Commons Attribution 3.0 Unported License.

When a thrombus obstructs one or more pulmonary arteries, it increases resistance in the pulmonary vasculature, forcing the RV to work against an elevated pressure afterload. This

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

can rapidly progress through three successive stages of cardiac involvement. First, because of the increased pulmonary vascular resistance, the RV is forced to generate higher pressure to maintain the blood circulation, leading to RV dysfunction, characterized by RV dilation and impaired filling (the RV has trouble adequately filling up with blood). Second, if the pressure overload persists, the RV dysfunction repercussion transforms into a RV failure. The RV has trouble sustaining adequate cardiac output, and the consequences extend beyond the right heart: the arterial pressure drops, and the blood's ability to oxygenate (hematosis) is impaired. In the most severe case, the RV failure transforms into right heart failure, when the drop in cardiac output and downstream pressure (hypotension) is critical, and the hematosis is severely impaired. The most life-threatening complication of PE arising from this cascade is the combination of cardiac failure with obstructive shock, defined as a sustained drop in systolic blood pressure, which constitutes a medical emergency.

II.2 Clinical signs

The physiological cascade described before produces a range of detectable clinical signs and measurable biological markers, which together form the basis for assessing PE diagnosis. CTPA has become the primary diagnostic modality for PE (Mehdipoor et al. 2020), with its use rising from 47% to 90% of patients assessed with it between 2001 and 2013 (Jiménez, de Miguel-Díez, et al. 2016). RV dysfunction can be detected on CTPA, it appears as RV dilation quantified by right-to-left ventricular diameter ratio (RV/LV), while on echocardiography, it is characterized by impaired RV filling. An RV/LV ratio greater than one indicates abnormal RV enlargement and is associated with a significantly increased risk of PE-related mortality (Meinel et al. 2015). RV failure is clinically detected with lower limb edema, jugular venous reflux, liver pain, elevated Brain Natriuretic Peptide (BNP) levels, and sometimes elevated troponin levels. BNP is released by the myocardium in response to increased wall stress and elevated filling pressures. Elevated BNP levels reflect the degree of RV pressure overload and are associated with worse clinical outcomes. Troponin is a marker of cardiomyocyte injury, with elevated levels reflecting subendocardial ischemia caused by sustained RV pressure overload. Elevated troponin levels are independently associated with increased short-term mortality in PE. In the most severe cases, right heart failure results in a decreased cardiac output that can be observed on an echocardiography. Finally, obstructive shock is clinically detected by a sustained drop in systolic blood pressure below 90 mmHg.

II.3 Comorbidities

Comorbidities increase the risk of mortality in patients. The principal ones being left heart failure, respiratory failure, and cancer. Right heart failure can lead to left heart failure, and vice versa. Left heart failure arises from dysfunction of the left heart chamber, which can stem from several distinct mechanisms : ischemic causes, dilated causes, or rhythmic causes. Ischemic causes involve impaired myocardial contraction, reducing the cardiac output when the heart muscle can no longer function properly. Dilated causes refer to excessive weakening of the heart muscle, preventing the heart from emptying adequately due to overfilling. Finally, rhythmic causes correspond to abnormal cardiac rhythms that disrupt the regular blood flow required to adequately perfuse the tissues.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

III Assessing PE severity

In the presence of thrombi, mortality can result from either direct cardiac complications or from underlying comorbidities. Following diagnosis, physicians focus on determining the patient's prognosis by estimating the likelihood of survival over the next 30 days. This process, known as **risk stratification**, guides all subsequent therapeutic decisions.

III.1 Risk stratification

Early mortality risk		Indicators of risk			
		Haemodynamic instability[a]	Clinical parameters of PE severity and/ or comorbidity: PESI class III–V or sPESI ≥I	RV dysfunction on TTE or CTPA[b]	Elevated cardiac troponin levels[c]
High		+	(+)[d]	+	(+)
Intermediate	Intermediate–high	-	+[e]	+	+
	Intermediate–low	-	+[e]	One (or none) positive	
Low		-	-	-	Assesment optional; if assessed, negative

©ESC 2019

Fig. 1.9: Table of risk stratification criteria to categorize patients. Figure extracted from Konstantinides et al. 2020.

In clinical practice, since 2019, the European Society of Cardiology guideline (Konstantinides et al. 2020) indicates how to categorize the 30-day mortality risk into 3 categories: high, intermediate-high or low, and low (see Figure 1.9). This risk stratification forms the basis for all subsequent therapeutic decisions. After assessing hemodynamic instability (cardiac arrest, obstructive shock or persistent hypotension), risk stratification should rely on three biomarkers: the RV/LV, and the blood levels of BNP and troponin. Although the guidelines specify that the RV/LV ratio should ideally be measured manually using a transthoracic echocardiography, this exam is not routinely performed in emergency settings. In practice, the ratio is therefore measured directly from the diagnostic CTPA.

High-risk patients are defined as those with hemodynamic instability. Hemodynamic instability is characterized by the presence of one of the following events:

- Cardiac arrest.
- Obstructive shock defined by systolic blood pressure < 90 mmHg or the need for vasopressors to achieve systolic blood pressure ≥ 90 mmHg despite adequate filling status, and hypoperfusion of terminal organs.
- Persistent hypotension: systolic blood pressure < 90 mmHg or a drop in systolic blood pressure ≥ 40 mmHg lasting more than 15 minutes and not caused by arrhythmia, hypovolemia, or recent-onset sepsis.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Intermediate-risk patients are defined by patients without hemodynamic instability but with signs of RV involvement, indicated by:

- Elevated cardiac enzyme levels, B-type natriuretic peptide > 600 pg/mL and troponin > 14 ng/L.
- Evidence of dysfunction and/or enlargement of the left ventricle (LV), by echocardiography or CTPA.

Intermediate-risk patients are further divided into two subcategories: high-intermediate risk if both enzymes and the RV/LV ratio are elevated, and low-intermediate risk if only one (or neither) is positive.

Low-risk patients exhibit no hemodynamic instability and no signs suggesting cardiac involvement, particularly of the RV.

III.2 Clinical Scores

In clinical practice, the Pulmonary Embolism Severity Index (PESI) (Aujesky et al. 2005) is a prognostic score that integrates information on all possible causes of mortality. The PESI score is based on 11 criteria, half related to cardiac impact (blood pressure, heart rate, oxygen levels) and the rest to comorbidities that increase mortality risk. The PESI score ranges from I to V and a PESI score of Class I or II indicates a 30-day mortality risk of 0–3.5%, while a score of Class IV or V corresponds to a risk of up to 24.5%.

The complexity of computing the original PESI score limited its adoption in clinical settings, particularly in emergency departments. To address this need for simplification, the Simplified Pulmonary Embolism Severity Index (sPESI) (Jiménez, Aujesky, et al. 2010) was developed and is now the version used in routine practice. The sPESI score is based on 6 of the original 11 criteria; each is worth one point, so the score now ranges from 0 to 6. Patients with a sPESI of 0 are low-risk patients. Patients with a sPESI > 0 are considered at least intermediate risk, and can be high risk as well.

III.3 Thrombotic burden

However, both the clinical scores and the biomarkers used in risk stratification share a common limitation: they characterize the consequences of PE on the heart rather than the thrombi directly. The RV/LV ratio is an indicator of the RV dilatation in consequence to the increased pressure to maintain blood flow, and blood biomarkers such as BNP and troponin are similarly released as downstream consequences of RV failure. Furthermore, in routine clinical practice, CTPA is not synchronized with the heart rhythm, resulting in image blurring around the heart and toward the outer edges of the lungs due to cardiac and respiratory motion. This limits the accuracy of RV/LV ratio estimation. These limitations motivate the development of scoring systems that directly quantify the thrombotic burden by characterizing the location and extent of clot obstruction within the PAT. Over the decades, several scoring systems have

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

been developed to standardize and improve the assessment of thrombotic burden, each offering unique advantages and addressing the limitations of its predecessors.

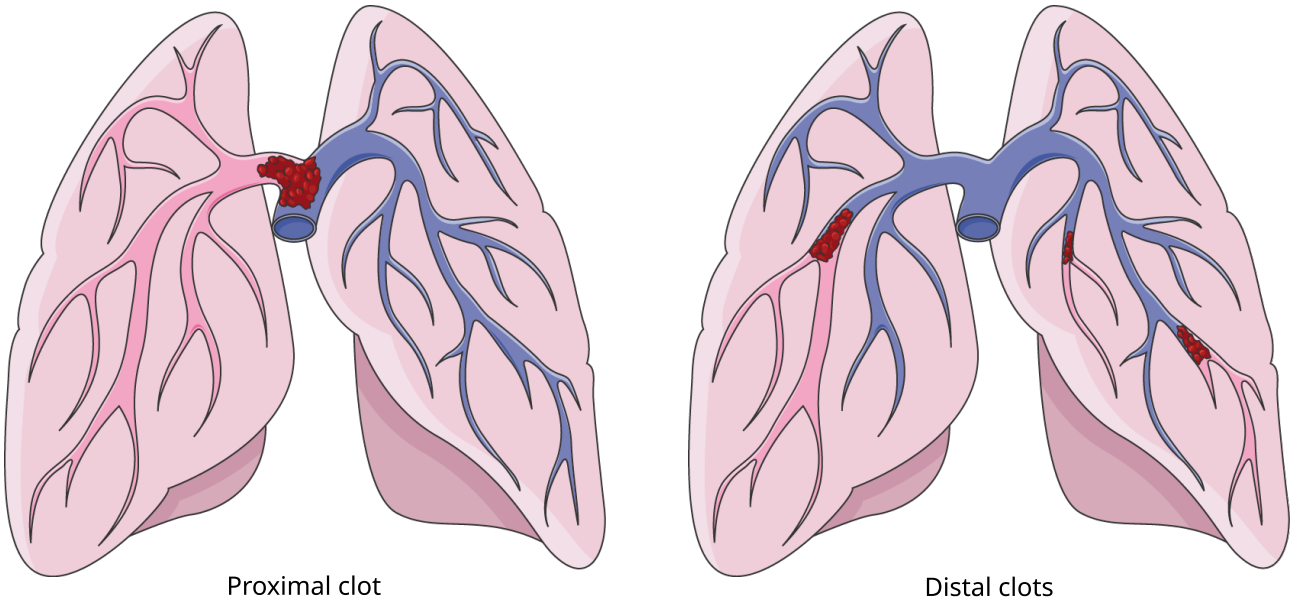


Fig. 1.10: Illustration of a proximal clot and a distal clot.
From Servier Medical Art by Servier, licensed under a Creative Commons Attribution 3.0 Unported License.

A fundamental concept underlying all these scoring systems is the anatomical localization of the clot within the PAT (see Figure 1.10). A clot is defined as proximal if located in the MPA the RPA, the LPA, or the lobar arteries (hierarchical levels of depth 1, 2, or 3, respectively). From the segmental arteries onward, including the sub-segmental arteries and below (level 4, 5+), a clot is considered distal. The proximal/distal distinction is a key criterion, as it reflects both the obstruction's proximity to the heart and the extent of the affected vascular territory. The different scoring systems presented in the next section often refer to the hierarchical level of the segmental arteries.

III.3.1 Thrombotic burden scores

Miller score The Miller score (G. A. H. Miller et al. 1971), introduced in 1971, was the first attempt to characterize thrombotic burden. This score assumes that the PAT is composed of nine segments on the right and seven on the left, which is less than today's anatomical references. It is divided into two sub-scores: an involvement score and a flow reduction score. For the involvement score, a thrombus located within a segmental artery is equal to one point. If a thrombus is positioned proximally to a segmental artery, the contribution equals the number of segmental arteries arising distally from the obstruction. The maximum involvement score is therefore 9 for the right lung and 7 for the left lung, yielding a combined maximum of 16 points. For the flow reduction score, each lung is divided into three zones (upper, middle, and lower), and the reduction in pulmonary arterial flow within each zone is assessed using a four-level scale: 0 for normal flow, 1 for mildly reduced flow, 2 for severely reduced flow, and 3 for absent flow. The maximum flow reduction score is 9, yielding a combined maximum of 18 points. The global maximum Miller score is therefore 34 points. While groundbreaking, the Miller score

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

treats partial and complete obstructions equally when computing the involvement score, which limits its precision.

Walsh score In 1973, the Walsh score (Walsh et al. 1973) refined this approach by standardizing the number of segments to nine for each lung, which is closer to today's anatomical references but still misses one segmental artery in the RLL. Unlike the Miller score, the Walsh score introduces rules to differentiate between complete and incomplete obstructions, providing a more nuanced assessment. At this time, the Miller and Walsh score relied on the pulmonary angiography to assess the perfusion : a contrast agent is injected directly into the pulmonary circulation, and a series of X-ray images are acquired to visualize the arterial tree and assess blood flow distribution within the lung. However, this image modality carries non-negligible risks and has therefore been progressively replaced in clinical practice by Computed Tomography (CT) pulmonary angiography. Both the Miller and Walsh scores were later adapted for use with CT scans. Research by Bankier et al. 1997 demonstrated a strong correlation between these two scores and a moderate correlation with heart rate, although the study noted that perfusion speed, a key factor, could not be assessed using CT imaging. This work was pivotal, as it was the first to show that the severity of a PE could be evaluated using CT scans, based on a cohort of 31 patients.

Qanadli score By 2001, the Qanadli score (Qanadli et al. 2001) emerged to address the shortcomings of its predecessors. The adaptation of the Miller score on CT scan no longer accounted for perfusion data, and the Walsh score, while more detailed, was rarely used in clinical practice due to its complexity. The Qanadli score redefines the PAT with ten segments per lung. The authors do not explain their decision to increase the number of segmental arteries per lower lobe, but today's anatomical representation shows only four in the LLL. For each thrombus, this score is calculated as the number of distal segments emerging from the obstruction, multiplied by a weight reflecting the degree of obstruction: 0 for no obstruction, 1 for partial obstruction, and 2 for complete obstruction. Isolated subsegmental clots are uniformly assigned 1 point. The calculation of this score is presented in Figure 1.11. This approach combines the segmental focus of the Miller score with the Walsh score's distinction between partial and complete obstructions, offering a more practical and clinically relevant tool.

Mastora score In 2003, the Mastora score (Mastora et al. 2003) introduced an even more precise method for calculating obstruction on an individual vessel basis. It employs a five-level obstruction grading system (1: 0–24%, 2: 25–49%, 3: 50–74%, 4: 75–99%, 5: 100%) and categorizes the PAT into mediastinal, lobar, and segmental arteries. Specifically, it identifies five mediastinal arteries (pulmonary trunk, right and left main arteries, and right and left interlobar arteries), six lobar arteries (right and left upper and lower lobes, right middle lobe, and left lingula), and 20 segmental arteries. The calculation of this score is presented in Figure 1.12. The Mastora score claims greater precision than previous systems and that it was designed to be more amenable to semi-automated calculation, potentially improving consistency and efficiency in clinical settings. Nevertheless, it was not widely adopted in a semi-automatic setting, and because it takes too much time to calculate manually, it is also not used in daily practice.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

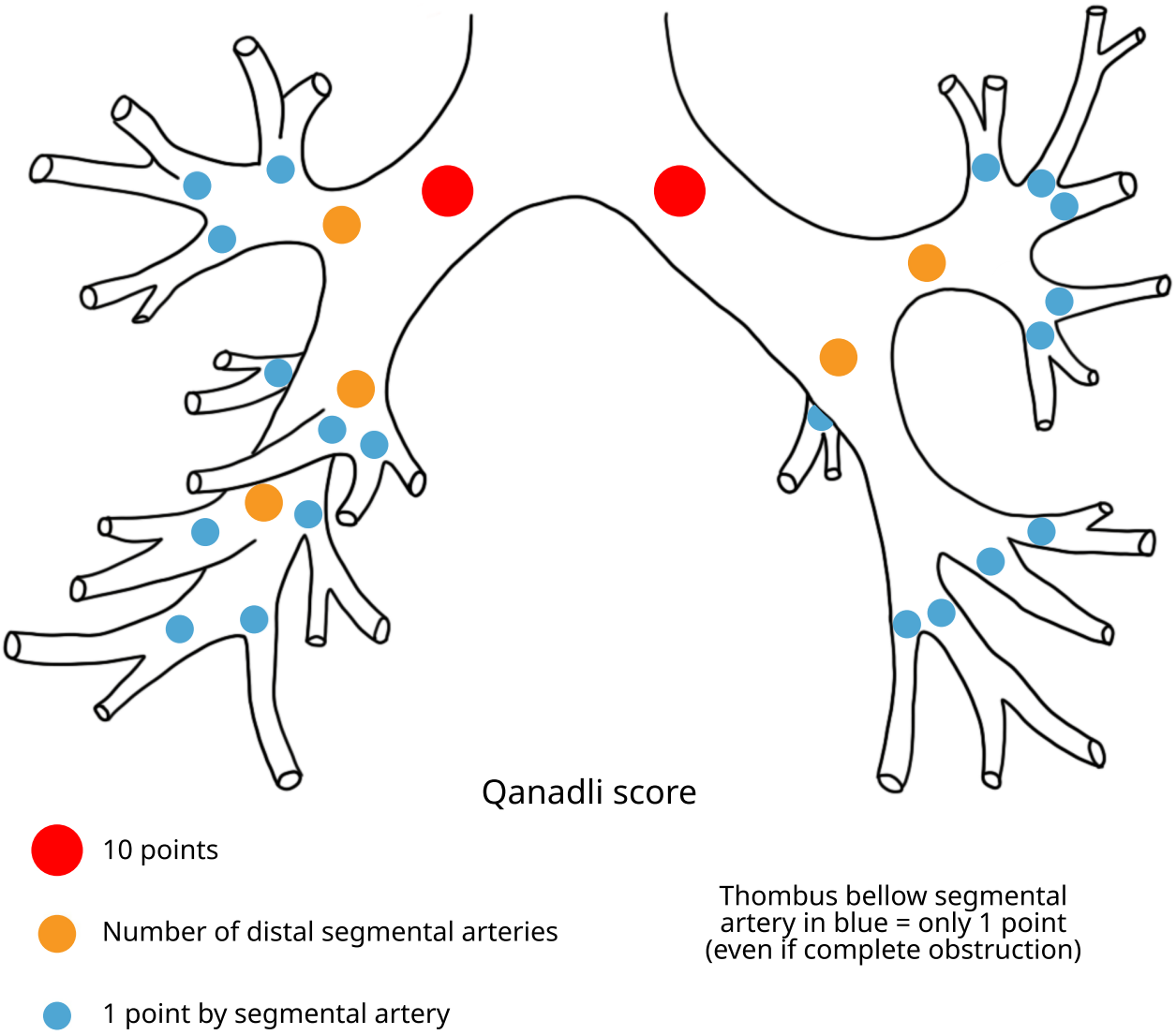


Fig. 1.11: Illustration of the location considered in the calculation of the Qanadli score. Modification of a Figure extracted from Alis and Haramati 2023

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

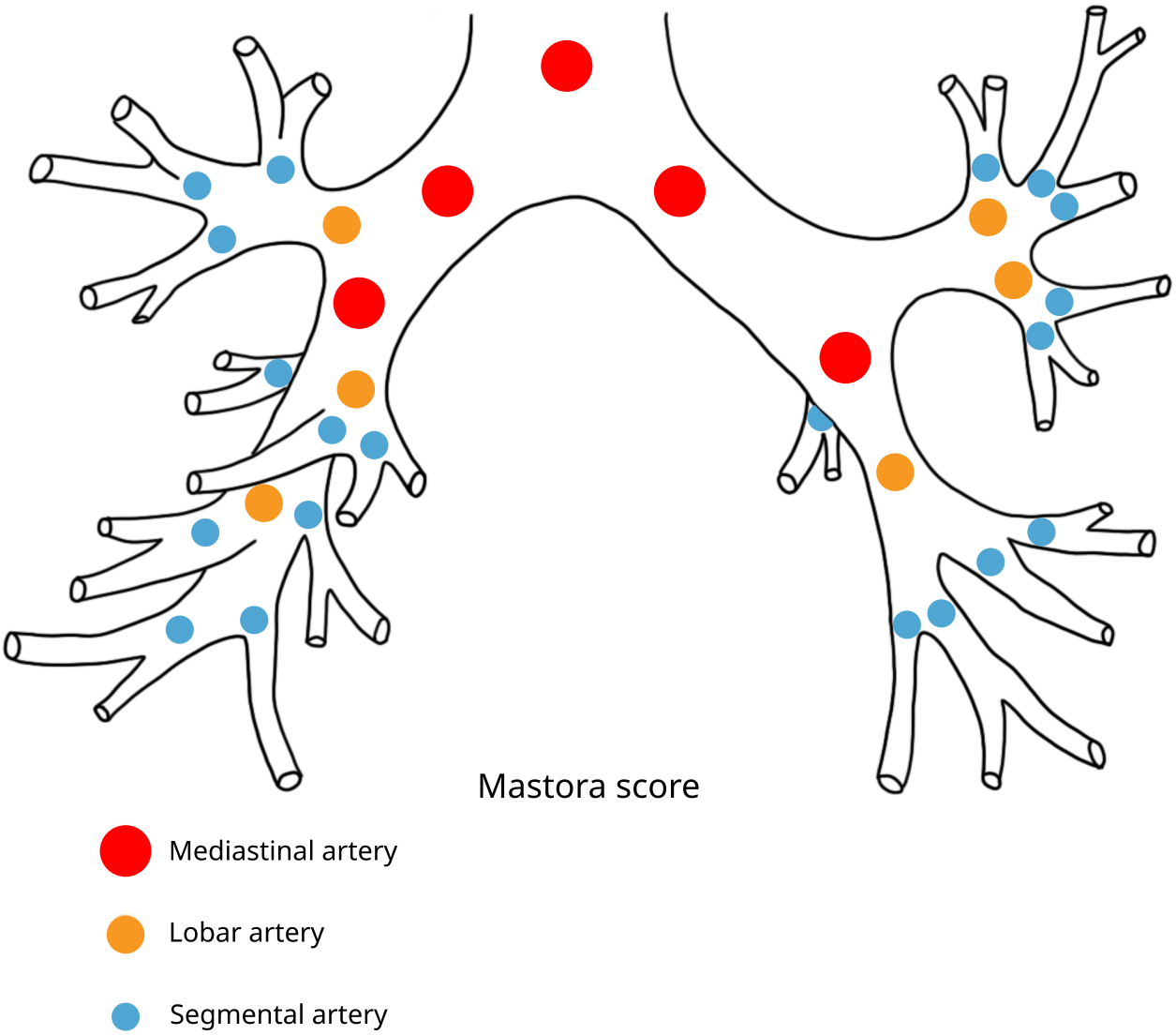


Fig. 1.12: Illustration of the location considered in the calculation of the Mastora score. Modification of a Figure extracted from Alis and Haramati 2023

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Embolic Volume Finally, several studies attempted to evaluate the blood clot burden through the Total Embolic Volume (TEV). This biomarker measures the total number of voxels of thrombi visible on CTPA either through manual segmentation (Nakada et al. 2010) or semi-automatic 3D segmentation method (Furlan et al. 2011; W.-M. Huang et al. 2022).

To conclude on the thrombotic burden estimation, Table 1.2 summarizes the main characteristics of the scoring systems described before. The successive refinements from the Miller score to the Mastora score reflect a progressive effort to improve the characterization of thrombotic burden, moving from involvement criteria toward vessel-level obstruction grading. The TEV represents a complementary approach, quantifying the raw volume of thrombi rather than their anatomical localization within the PAT. Despite these methodological advances, all of the existing scoring systems share a critical practical limitation: they are too time-consuming to compute manually in routine clinical practice. The Qanadli score requires the physician to identify each affected vessel and estimate the degree of obstruction across the entire PAT. The Mastora score, while more precise, demands an even finer vessel-level assessment, making it largely impractical outside of research settings. As a consequence, thrombotic burden is rarely quantified systematically in clinical practice, which in turn limits the availability of large annotated cohorts needed to rigorously assess its prognostic value. This gap between methodological and clinical feasibility underscores the need for automated approaches that can reliably compute these scores directly from CTPA scans. In Chapter 6 I.4.1 we will discuss how we adapted those state-of-the-art scores into automatic calculation.

Score	Year	Modality	Segmental arteries	Obstruction grading
Miller	1971	Angiography / CT	16 (9R + 7L)	Binary
Walsh	1973	Angiography / CT	18 (9R + 9L)	Partial / Complete
Qanadli	2001	CT	20 (10R + 10L)	3 levels
Mastora	2003	CT	20 (10R + 10L)	5 levels
TEV	2011	CT	N/A	Continuous (voxel volume)

Tab. 1.2: Summary of the main thrombotic burden scoring systems.

III.3.2 Controversy on thrombotic burden

The role of thrombotic burden in the risk stratification of PE remains a subject of debate. Luijten et al. 2023a argue that thrombotic burden has a limited role in risk stratification for hemodynamically stable patients, noting that many predictors of PE-related shock are interrelated and dynamic, making it difficult to rely on a single clinical feature to identify patients at risk of deterioration. In a response letter, Casey et al. 2023 argues that this claim requires additional context and that the evidence is more nuanced than suggested. The studies cited by both sides illustrate this complexity, with some reporting significant associations and others finding none.

Evidence against the predictive value of thrombotic burden. Multiple studies have failed to establish a significant association between thrombotic burden and short-term mortality. Vedovati et al. 2013, in a meta-analysis of 16 studies with a minimum of 14 patients to a

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

maximum of 1993 patients, found no correlation between the obstruction index (measured using the Qanadli score) and 30-day mortality (Odds Ratio (OR) 1.22, 95% Confidence Interval (CI) 0.99–1.51, $I^2 = 27\%$). Rotzinger et al. 2020 similarly reported no significant relationship between the Qanadli index and mortality ($P = 0.156$) in a cohort of 705 patients. More recently, Tajeri et al. 2024 corroborated the idea that the Qanadli score is not significantly correlated with mortality, on a cohort of 102 patients.

Evidence supporting the predictive value of thrombotic burden. Several studies suggest that thrombotic burden carries meaningful prognostic information, at least for certain outcomes or patient subgroups. Meinel et al. 2015 reported that, while thrombotic burden (Qanadli score) was not predictive of all-cause mortality (OR 1.6, 95% CI 0.7–3.9; $P = .2896$), it was significantly predictive of PE-related mortality (pooled OR 17.6, 95% CI 5.0–61.7; $P < .0001$), although this result was derived from only three small studies totaling 260 patients. Thrombotic burden was also predictive of composite adverse clinical events (pooled OR 2.4, 95% CI 1.3–4.5; $P = .0059$). Hariharan et al. 2016 showed an association between thrombotic burden (measured by the CT-PASS score, a modification of the TEV[2]) and adverse clinical events in a multivariate regression analysis adjusted for right ventricular strain, although this association was not statistically significant in hemodynamically stable patients specifically (OR 3.34, 95% CI 0.75–14.85). W.-M. Huang et al. 2022 reported that thrombotic burden (measured by TEV) and RV dysfunction are associated with shock in initially hemodynamically stable high-risk patients. Based on these findings, Casey et al. 2023 argues that thrombotic burden appears to be an emerging relevant factor for stratifying risk in certain initially stable high-risk patients and should continue to be explored.

Source of contradictions These contradictory findings can be attributed to several factors. First, the studies differ in their outcome definitions: some evaluate all-cause 30-day mortality, others focus on PE-related mortality specifically, and others use composite endpoints such as adverse clinical events or hemodynamic deterioration. As noted by Casey et al. 2023, the studies by Hariharan et al. 2016 and W.-M. Huang et al. 2022 do not clearly define the notion of shock, particularly regarding its onset and timing, which complicates interpretation. Second, the scoring methods used to quantify thrombotic burden are not uniform: the Qanadli score is the most commonly used, but Hariharan et al. 2016 relies on the CT-PASS score, and W.-M. Huang et al. 2022 uses TEV. Third, cohort composition varies across studies: some include exclusively hemodynamically stable patients, whereas others include both stable and unstable patients or focus on specific subgroups such as high-risk patients with initial hemodynamic stability. These differences in population, exposure measurement, and outcome definition are likely the major sources of the observed heterogeneity.

Emerging consensus In their response, Luijten et al. 2023b agree that thrombotic burden cannot be used as the sole criterion for risk stratification. They emphasize that, while thrombotic burden is associated with RV dysfunction, this does not justify its use as a standalone predictor of shock or mortality, especially given the heterogeneous definitions of thrombotic

[2]The CT-PASS is a simplified variant of the TEV, replacing the full volumetric measurement of all thrombi with the sum of the largest ones in each lung only.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

burden and clinical events across studies. They suggest that thrombotic burden could play a more prominent role in the future if studies demonstrate that its inclusion in RV dysfunction assessment improves health outcomes, and they caution against using it alone for therapeutic decision-making.

Despite the lack of consensus on mortality prediction, thrombotic burden appears consistently correlated with other clinically relevant parameters. Both Rotzinger et al. 2020 and Tajeri et al. 2024 reported significant correlations between the Qanadli score and troponin levels, D-dimer levels, heart rate, and pulmonary hypertension. Furthermore, An et al. 2022 showed that, while thrombotic burden alone cannot determine prognosis from CTPA, combining it with other clinical factors enables the assessment of pulmonary arterial pressure and RV dysfunction, which are independent predictors of chronic thromboembolic pulmonary hypertension. Thrombotic burden is also relevant beyond the acute phase: the PADIS-PE study Tromeur et al. 2018 showed that patients presenting with at least 40% thrombotic burden (Qanadli) at initial diagnosis carry an increased risk of recurrence if anticoagulants are discontinued, as well as an increased risk of long-term complications, including pulmonary hypertension.

Taken together, these findings suggest that thrombotic burden is not a reliable standalone predictor of short-term mortality but carries meaningful prognostic information through its associations with cardiac biomarkers, hemodynamic parameters, and long-term outcomes. However, the Qanadli and Mastora scores remain too time-consuming to compute in routine clinical practice, and there are currently no large-scale studies systematically examining the correlation between thrombotic burden and cardiac biomarkers such as BNP and troponin. This underscores the need for automated methods to quantify thrombotic burden and to rigorously investigate its added value in risk stratification using well-defined cohorts and standardized outcome definitions.

IV PERSEVERE Project

As discussed in the previous section, there remains significant uncertainty regarding the prognostic value of thrombotic burden: while it is not a reliable standalone predictor of short-term mortality, it correlates with cardiac biomarkers, hemodynamic parameters, and long-term outcomes. Resolving this uncertainty requires large-scale studies with standardized definitions, which in turn depend on the ability to systematically and reproducibly quantify thrombotic burden across patients. However, the existing thrombotic burden scores, such as Qanadli and Mastora, remain too time-consuming to compute manually, preventing their routine use in clinical practice and limiting the availability of large annotated cohorts. Furthermore, the blood biomarkers on which current risk stratification relies, notably BNP and troponin, are not always available, particularly in emergency settings where time is limited. Clinicians have reported the need for a tool capable of analyzing the CTPA scan directly to flag patients at significant risk and remind them to take the blood samples for BNP and troponin measurement when necessary.

The PERSEVERE project fits into this context, aiming to address these challenges by automating the extraction of new biomarkers and reproducible quantification of thrombotic burden

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

directly derived from CTPA scans, which are always routinely available since the diagnosis of PE is established on the CTPA. These new biomarkers would fall into two categories : local biomarkers that capture vascular and embolic features and patient-level biomarkers that aggregate these local measurements into global ones, such as the thrombotic burden.

Achieving this goal requires several successive technical steps. First, the pulmonary arteries and thrombi must be accurately segmented from the CTPA images. These segmentation masks then serve as the basis for a graph-based PAT modeling, which encodes the vascular and embolic features. Finally, new biomarkers are extracted from this graph representation. The PERSEVERE project is therefore organized into three steps: the segmentation of the PAT and thrombi from the CTPA, the graph-based modeling of the PAT and the extraction of imaging-derived biomarkers from this graph (see Figure 1.13).

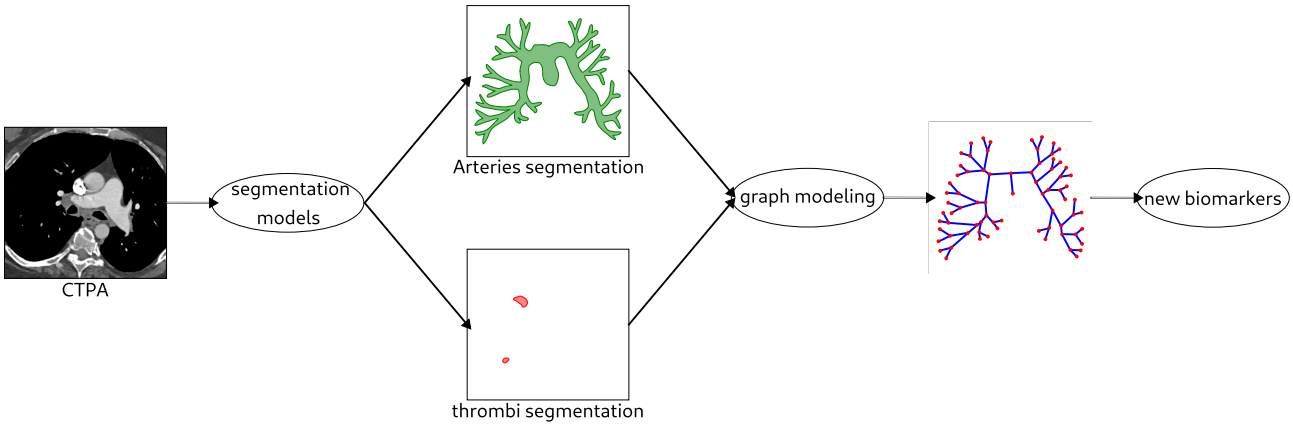


Fig. 1.13: PERSEVERE project overview of the three main steps.

V Conclusion

This chapter has established the medical and anatomical foundations underlying the contributions of this thesis. We first described the PAT, highlighting its tree-shaped, acyclic, and hierarchically organized structure, as well as the inter-individual variability in branching patterns, especially at the segmental level. These properties motivate the graph-based modeling approach developed in Chapter 6. We then briefly introduced the cerebrovascular anatomy, which presents a contrasting topology through its tortuous geometry and cyclic nature, making it valuable for testing the generalizability of our methods. We then presented PE, a life-threatening cardiovascular condition caused by the obstruction of pulmonary arteries by blood clots. We described its physiological mechanisms, clinical consequences, and the current diagnosis workflow based on CTPA imaging. We also reviewed existing thrombotic burden scoring systems, from the Miller and Walsh scores to the Qanadli and Mastora scores, and discussed their respective limitations regarding clinical applicability and routine use.

Finally, we identified the core challenges motivating this work. Existing thrombotic burden scores are too time-consuming to compute manually, and large-scale studies correlating thrombotic burden with cardiac biomarkers such as BNP and troponin remain scarce. These limita-

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

tions call for automated, imaging-derived methods capable of quantifying thrombotic burden, which is precisely the goal of the PERSEVERE project within which this thesis is inscribed.

The objective of this thesis is therefore to extract new imaging-derived biomarkers directly from the CTPA, which is always available as it serves as the primary diagnostic tool for PE. This extraction relies on a graph-based modeling of the PAT, which in turn requires accurate and topologically correct segmentations of the pulmonary arterial tree. These two requirements, segmentation and graph modeling, constitute the core technical contributions of this work, developed within the PERSEVERE project. In the following chapter 2, we introduce the datasets used throughout this work, including our in-house dataset, along with the evaluation metrics used to assess the performance of the proposed segmentation methods.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

2 Datasets and segmentation metrics

Table of contents

I Datasets . . . 26
I.1 In-house dataset: the PERSEVERE dataset . . . 26
I.2 Pulmonary Artery Segmentation Challenge 2022 dataset . . . 33
II Segmentation evaluation metrics . . . 36
II.1 Overlap metrics . . . 36
II.2 Topological metric . . . 38
III Conclusion . . . 39

This chapter describes the in-house clinical dataset collected within the PERSEVERE project framework, as well as two public vascular segmentation datasets used during the initial experimentation phase to train segmentation models. The PERSEVERE dataset serves as the primary resource for training, validating, and evaluating the vascular segmentation models (see Chapters 4 and 5) and the downstream PAT modeling and analysis pipeline (see Chapter 6). The patient inclusion criteria and scanner characteristics are first detailed, followed by a description of the image preprocessing steps applied to ensure consistency across the PERSEVERE cohort. The key characteristics of a publicly available annotated dataset are then observed. Finally, the set of metrics used to assess the segmentation quality throughout this thesis is presented, combining overlap and topological metrics, each capturing a complementary aspect of segmentation quality adapted to the specificity of vascular structures.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I Datasets

This section presents the data resources used throughout this thesis. We first describe the in-house PERSEVERE dataset, collected within a clinical research framework, which constitutes the core material for both model development and downstream graph-based analysis. We then introduce a public vascular dataset, namely the Pulmonary ARtery SEgmentation Challenge (PARSE) 2022 challenge dataset, which provided complementary resources during the early stages of deep learning experimentation (see chapter 5).

I.1 In-house dataset: the PERSEVERE dataset

The PERSEVERE dataset includes 431 patients diagnosed with acute PE diagnosed on CTPA at the emergency department of CHU Saint-Étienne between April 2014 and June 2020. All patients were admitted within 72 hours to the Vascular and Therapeutic Medicine unit, headed by Pr. Laurent Bertoletti. The dataset also provides extensive clinical information, including PE severity level (see Chapter 1 III.1), sPESI score (see Chapter 1 III.2), and, for most patients, troponin and BNP levels. Patients with hemodynamic instability were not included in the cohort, meaning that high-risk patients are not included in this study. The final cohort received approval from the local ethics committee (IRBN262017/CHUSTE).

CTPA images were reconstructed using soft convolution kernels (e.g., I40f, I30f, B30f, B40f, see additional information in annex II) with a mediastinal window setting. An iodinated contrast agent (Xenetix) was administered at the arterial phase. Data were transmitted anonymized on the CREATIS GIRDER platform, in DICOM format, and converted to NIfTI format. For 22 patients, the mediastinal series could not be identified through automated parsing of the DICOM metadata and required manual review. Once converted and reviewed, the full set of 431 patients was subjected to a series of exclusion criteria to ensure the consistency and quality of the final cohort, as detailed in the following subsections.

I.1.1 Exclusion criteria

Starting from an initial cohort of 431 patients, several exclusion criteria were applied to obtain the final database of 353 patients : (i) Manufacturer homogeneity, (ii) incomplete clinical blood biomarkers, and (iii) image quality. The following subsections detail each criterion and the corresponding exclusion steps, as summarized in Figure 2.1.

Manufacturer: Among the 431 patients, the CT scanner manufacturer and model used for each CTPA acquisition were analyzed to assess the homogeneity of the imaging equipment across the cohort. Results are summarized in Table 2.1. The scanner distribution across the dataset is highly imbalanced: 387 patients (89.8%) were acquired on the SIEMENS SOMATOM Definition AS, 29 patients (6.7%) on the SIEMENS SOMATOM Definition, 7 patients (1.6%) on the Siemens Sensation 16, and the remaining 8 patients across various GE and Philips systems (1.8%). Given that 89.8% of acquisitions were performed on SIEMENS scanners, specifically on the SOMATON Definition or Somaton Definition AS models, only patients scanned on these two devices were kept. Patients acquired on other scanner models were excluded, as

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

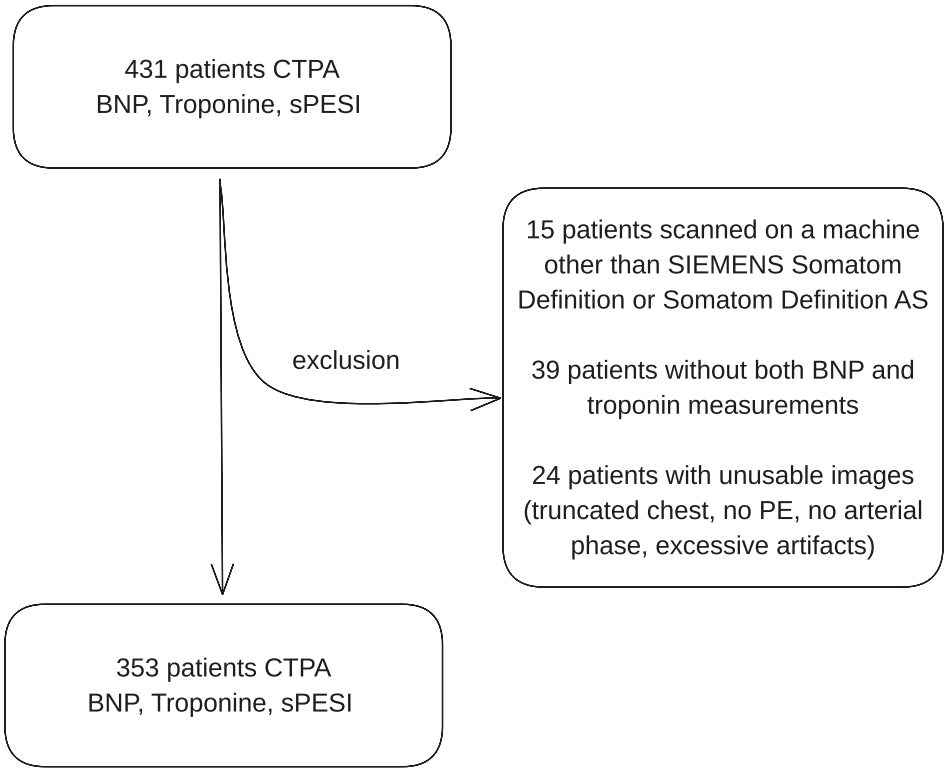


Fig. 2.1: Summary of the number of excluded PERSEVERE patients by criteria.

their limited numbers would have been insufficient to ensure a representative and statistically balanced contribution to the cohort. The database at this step is 416 patients after we removed 15 patients. The complete list of excluded patients, along with their corresponding scanner model, is provided in Appendix I.1.

Manufacturer	Model	Number of patient
Philips	Brilliance 40	1
GE MEDICAL SYSTEMS	BrightSpeed QX/i	1
	Optima CT540	1
	Optima CT660	2
	Revolution EVO	3
SIEMENS	Sensation 16	7
	SOMATOM Definition	29
	SOMATOM Definition AS	387
Total		431

Tab. 2.1: Distribution of patients per manufacturer and scanner model.

In order to ensure that the coexistence of two scanner models within the cohort did not introduce an intensity bias, global intensity statistics were compared between patients acquired on each

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

device. In all 414 patients[1], the mean lung intensity was -501.1 +/- 87.2 HU, range: [-790.0, -247.2] HU. The intensity distributions are represented with violin plots in Figure 2.2.

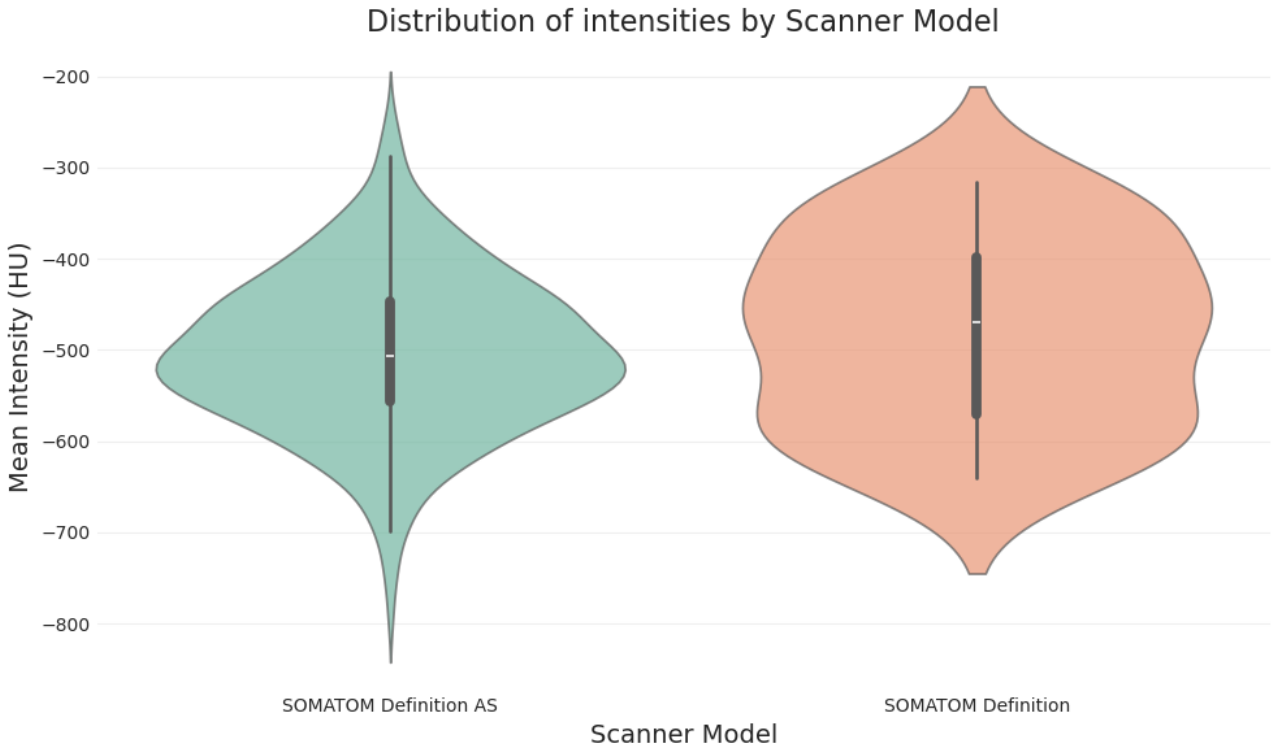


Fig. 2.2: Distribution of intensities of the SOMATOM Definition AS and SOMATOM Definition scanner.

The SOMATOM Definition AS yielded a mean of -502.7 +/- 85.8 HU, while the SOMATOM Definition yielded -480.2 +/- 102.9 HU, which is a difference of 22.5 HU (see Table 2.2). A student t-test confirmed that this difference was not statistically significant (t = -1.343, p = 0.180), supporting the compatibility of the two acquisition protocols.

Tab. 2.2: Global intensity statistics per scanner model across the 414 retained patients.

Scanner	**Number of patient**	**Mean (HU)**	**Std (HU)**	**Median (HU)**	**Min (HU)**	**Max (HU)**
SOMATOM Definition AS	385	−502.7	85.8	−505.9	−790.0	−247.2
SOMATOM Definition	29	−480.2	102.9	−469.9	−640.0	−316.1

These results indicate that, despite the use of different reconstruction algorithms, both scanner models produce images with statistically equivalent intensity distributions. Therefore, no additional intensity normalization was deemed necessary prior to model training.

[1]The patient 0097 did not have a MEDIASTINAL image and the patient 0146 had already been removed for missing arterial phase

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

BNP and Troponine dosage: Patients with missing BNP or troponin levels were excluded to ensure consistency in the available data for the rest of the ANR project, so we excluded 39 patients without complete blood biomarkers. The database now has 377 patients. The complete list of excluded patients is provided in Appendix I.2.

Image quality: Image quality was assessed using a set of predefined criteria derived from standard practices in CTPA acquisition and PE diagnosis. A scan was considered acceptable only if it met all of the following conditions:

- Complete anatomical coverage, with no truncation of the heart or lungs.
- Presence of a visible arterial phase, ensuring adequate contrast enhancement of the pulmonary vasculature.
- Absence of severe imaging artifacts, including motion artifacts, beam hardening artifacts, and injection-related artifacts, that would compromise image interpretation.
- Absence of significant lung damage or parenchymal abnormalities in the surroundings of the pulmonary arteries that would render vessel delineation unreliable.
- Sufficient contrast between the arterial structures and surrounding tissues to allow consistent vessel segmentation.

Applying these criteria to the cohort, 24 patients were excluded, and the complete list of excluded patients with details on the corresponding exclusion motif is provided in Appendix I.3. Briefly, three scans presented incomplete anatomical coverage, either due to a missing mediastinal acquisition or image truncation. Ten patients were excluded because no PE could be identified by the radiologist, making their inclusion in the cohort clinically inconsistent. One patient had no visible arterial phase, and two patients had insufficient contrast between the arterial structures and surrounding tissues and were excluded. Four patients presented severe artifacts caused by motion, beam hardening, or contrast injection issues. Finally, four additional patients exhibited extensive lung damage or parenchymal abnormalities compromising arterial delineation. Representative examples of scans excluded on the basis of these quality criteria are illustrated in Figure 2.3.

I.1.2 CTPA preprocessing

The original DICOM volumes were converted into NIfTI format and resampled to the dataset median voxel spacing (0.72x0.72x0.50 mm) with a trilinear interpolation, which provides a smooth approximation suited to continuous intensity fields such as CT images. The volumes were then cropped with a 10-pixel margin around the lung masks to reduce memory size and focus on the relevant thoracic region. While initial experiments relied on the `Lungmask` model, the final pipeline used `TotalSegmentator`, which provided more accurate lung segmentations and lobes delineation (see Annex I.4).

Volumes range from 292 to 470 voxels (mean 381.7 ± 34.8) along the x-axis, 213 to 371 voxels (mean 291.3 ± 28.7) along the y-axis, and 357 to 727 voxels (mean 478.1 ± 51.8) along the z-axis, reflecting the large anatomical variability in lung height between patients (see Figure 2.4).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

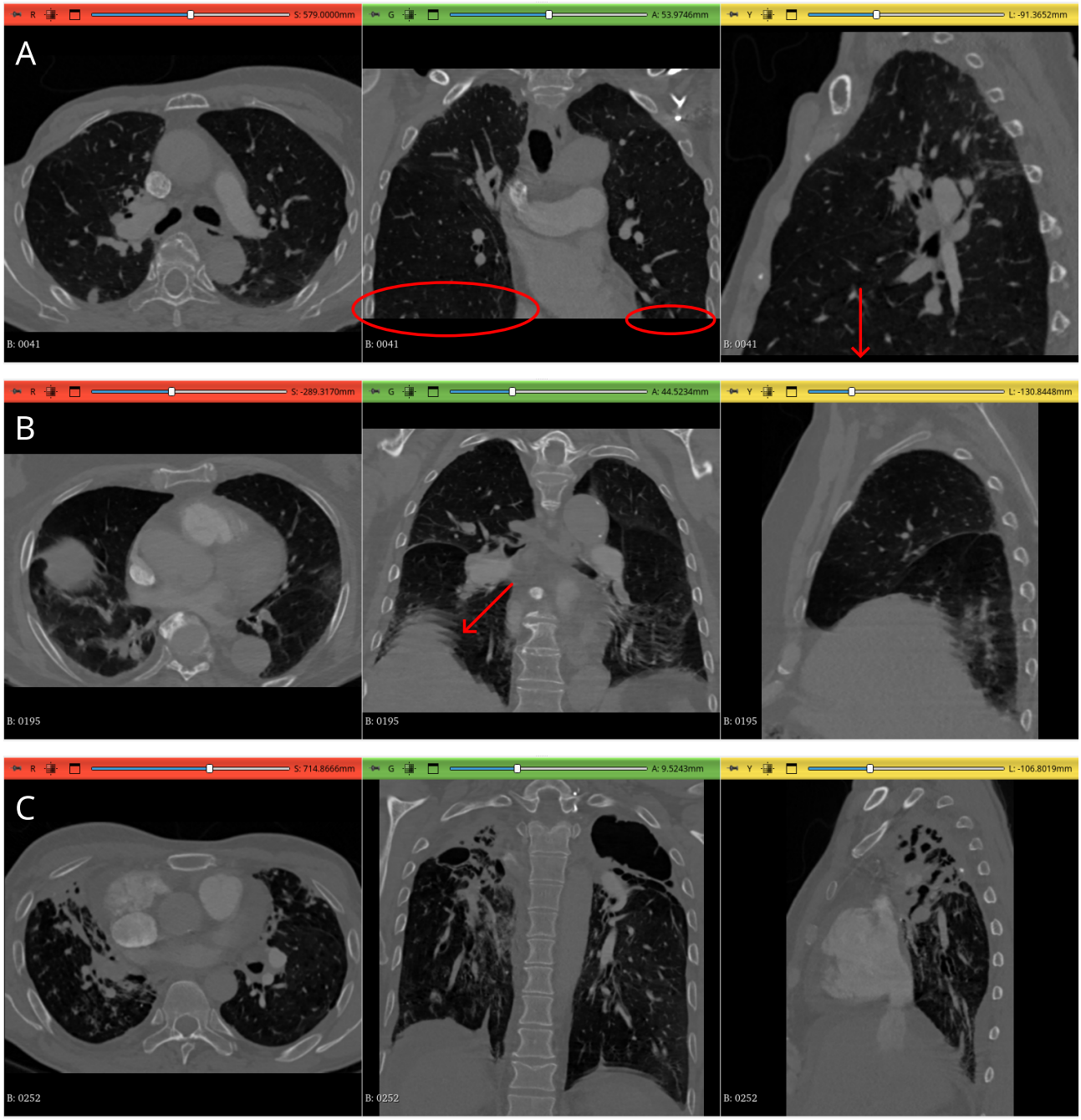


Fig. 2.3: Representative examples of images excluded from the cohort due to insufficient quality: (A) truncated field of view in patient 0041, (B) motion artifact in patient 0195, and (C) damaged lung parenchyma in patient 0252, displayed in 3D Slicer.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

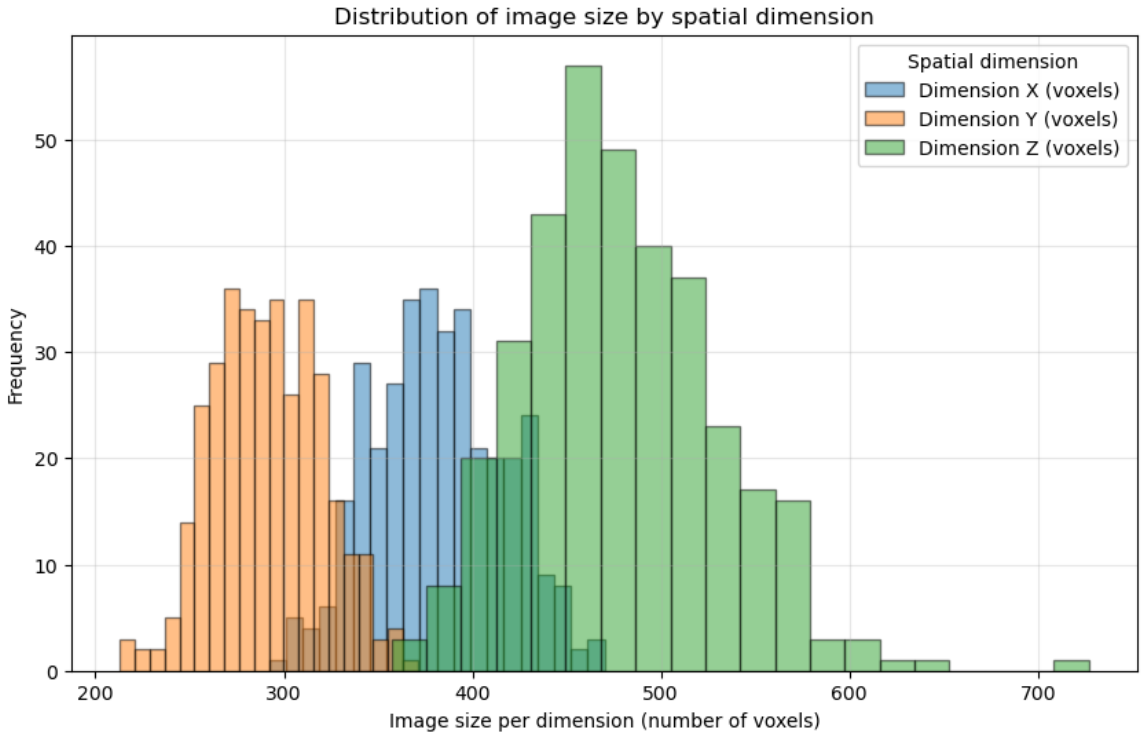


Fig. 2.4: Images sizes distribution on the final 353 patients of the SOMATOM Definition AS and SOMATOM Definition scanner.

I.1.3 Severity distribution

The distribution of patients in the severity categories of the 2019 ESC is shown in Table 2.3. High-risk patients were not included in the cohort, as their severity is immediately identifiable upon arrival at the emergency department through clinical signs of hemodynamic instability. The cohort is dominated by intermediate-risk patients, who represent approximately 86% of the total population (intermediate-low: 45.0%, intermediate-high: 41.1%). Low-risk patients account for only 13.9% of the cohort, making this class underrepresented compared to the two intermediate classes.

ESC 2019 Severity	**Number of patients (%)**
Low	49 (13.9)
Intermediate-low	159 (45.0)
Intermediate-high	145 (41.1)

Tab. 2.3: Patient distribution according to the ESC 2019 risk stratification categories (Konstantinides et al. 2020).

I.1.4 Selection of a subset of patients

Manual annotations are required to train and evaluate the segmentation models developed in this work. To this end, a subset of patients from the PERSEVERE dataset was selected for

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

manual annotation. Since manual annotation is a time-consuming process, a set of selection criteria was defined to ensure that only images of sufficient quality were included, so that the resulting annotations would be reliable and usable for model training and evaluation.

Selection was based on two main image quality criteria. First, the quality of the contrast injection was assessed, as poor injection can result in insufficient opacification of the pulmonary arteries, making them difficult to delineate accurately. Second, the impact of imaging artifacts was evaluated, covering motion artifacts, beam hardening artifacts, and artifacts linked to the injection itself. Each of these artifact types was scored on a three-level scale: 0 indicating absence, 1 indicating presence but without impact on interpretation, and 2 indicating presence with impact on interpretation. Injection quality was scored as a binary criterion: 0 if the injection did not meet the required standard, and 1 if it did. From these individual criteria, a global image quality score was derived: 0 for non-interpretable images, 1 for images interpretable with reservations, and 2 for images interpretable without reservations. Only patients reaching a global quality score of 2 were considered eligible for annotation.

Patient ID	sPESI	Severity	Troponine	BNP	Motion artefact	Hardening artefact	Injection artefact	Injection Quality	Global Quality
0028	2	2	NF	3578	1	1	1	1	2
0128	2	2	205	404	1	1	1	0	2
0222	2	2	95	9865	0	0	0	1	2
0031	1	1	39	238	1	1	1	1	2
0032	1	1	76	785	1	1	0	1	2
0055	1	1	10	769	0	0	0	1	2
0060	1	1	9	42	1	2	0	1	2
0216	0	0	8	106	2	1	1	1	2
0330	0	0	7	56	1	1	1	1	2
0233	0	0	7	30	1	1	0	1	2

Tab. 2.4: Sample of annotated patients with clinical and image quality criteria. Troponine and BNP values are reported upon admission to the Emergency unit. Artifact scores: 0 = absent, 1 = present but not affecting interpretation, 2 = present and affecting interpretation. Injection quality: 0 = does not meet criteria, 1 = meets criteria. Global image quality: 0 = non-interpretable, 1 = interpretable with reservations, 2 = interpretable without reservations.

Based on these criteria, 10 patients were selected, distributed across the three ESC 2019 severity categories, with at least three patients per category to ensure a representative sample. The clinical characteristics[2] and image quality scores of the selected patients are reported in Table 2.4. Ultimately, we aim to increase this annotation subset to improve segmentation models' performances (see Chapter 5). An illustration of these 10 patient samples is presented in Figure 2.5.

[2]In the annotated sample of 10 patients, patient 0028 has no Troponin value recorded. This is because the sample selection was based on Troponin levels measured in the Vascular and Therapeutic Medicine unit (VTM), rather than those collected at the Emergency unit (EU), both of which are available in the clinical data spreadsheet. For the final cohort of 353 patients, Troponin values collected at the EU were used, which is why this patient was excluded from the final cohort despite being included in the annotation subset.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

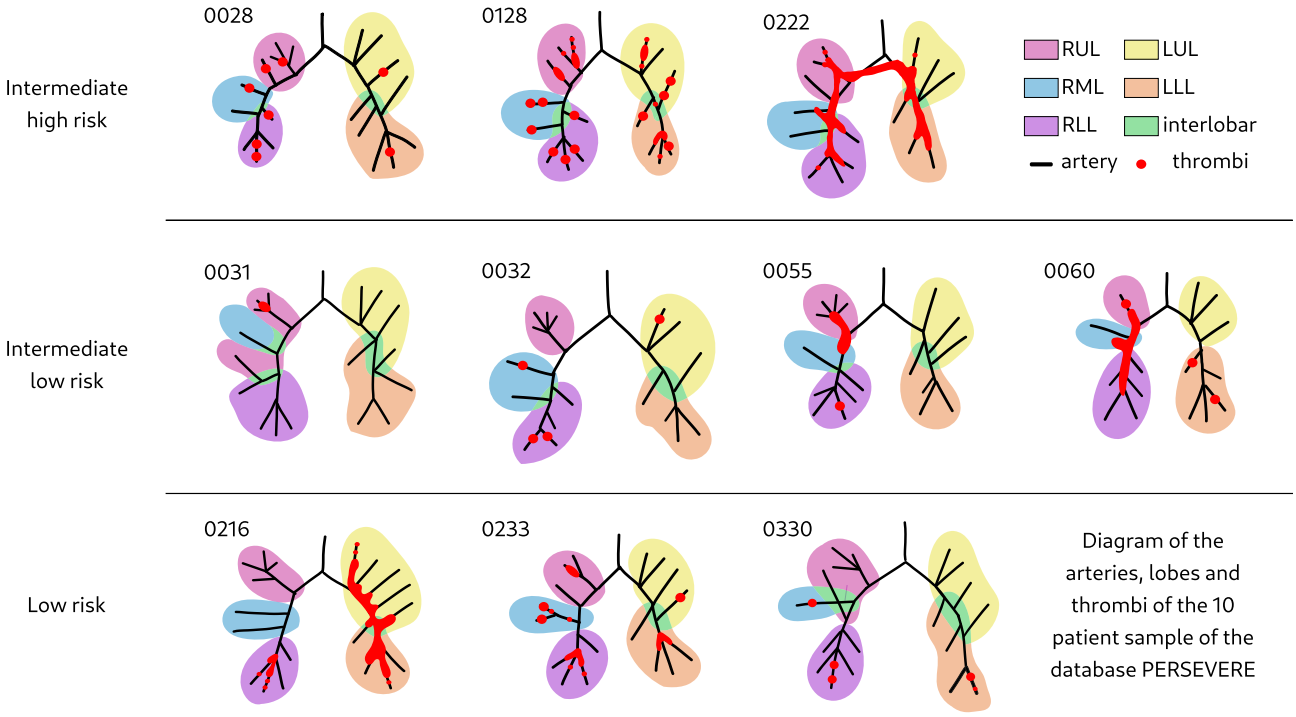


Fig. 2.5: Illustration of the 10 sample patient case with an overview of the location of the thrombi. Each lobe is color-coded, and the interlobar portion corresponds to the part where an artery crosses from one lobe to another. The peripheral arteries correspond to the segmental level.

Having described the PERSEVERE dataset and the preprocessing steps applied to ensure cohort consistency, we now turn to the public dataset used during the early stages of experimentation. This dataset, drawn from MICCAI challenge on pulmonary artery segmentation, provided a valuable starting point for developing and testing vascular segmentation models.

I.2 Pulmonary Artery Segmentation Challenge 2022 dataset

In this part, we will present the Pulmonary ARtery SEgmentation Challenge (PARSE) dataset from a challenge that took place during the MICCAI conference in 2022 (Luo et al. 2024), on which we began experimenting with segmentation. This challenge focuses on segmenting pulmonary artery structures, with the clinical motivation of enabling quantitative assessment of morphological changes in the pulmonary vasculature. Such measurements are particularly relevant for the diagnosis and follow-up of pulmonary hypertension, as well as for surgical planning in thoracic interventions. To support this task, the challenge relies on contrast-enhanced CTPA, which provides sufficient vascular contrast to delineate arterial structures. This dataset is an attempt to make a benchmark for the pulmonary artery segmentation task.

The dataset comprises 200 3D contrast-enhanced CTPA volumes (100 training, 30 validation, 70 testing) with pulmonary artery annotations. Only the 100 training cases (30340 slices total) with paired images/labels are publicly available in NIfTI format; validation images and labels are not available. Annotations were performed by 10 clinical experts with more than 5 years of

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

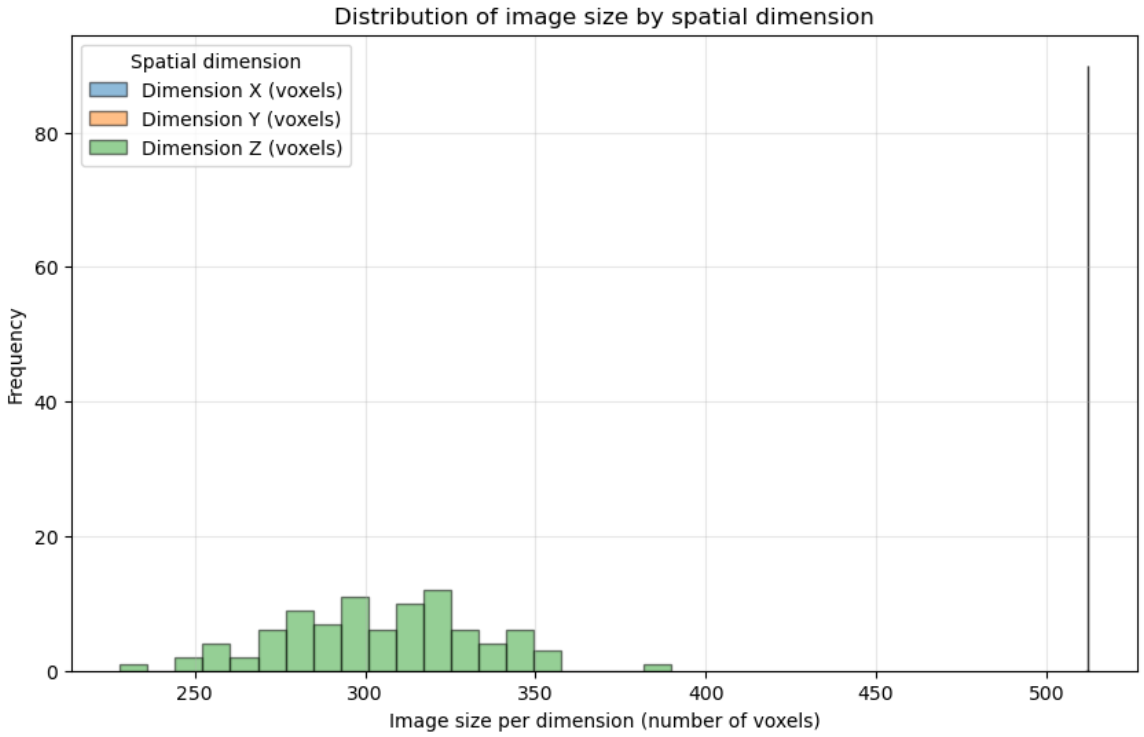


Fig. 2.6: Images size distribution on the 90 PARSE patients.

clinical experience using a semi-automatic region growing tool in the MIMICS software. The dimensions of this dataset are 512×512×[228-390] voxels (median: 512×512×301)(see Figure 2.6), voxel spacing 0.50-0.92×0.50-0.92×0.8-1.0 mm (median: 0.65×0.65×1.0 mm).

We use the training dataset in Chapter 5, section I. Figure 2.7 illustrates an example of an image and an annotation on the first patient of the PARSE training dataset. See more statistical information on this dataset in Annex IV.

The PERSEVERE and PARSE datasets presented provide the image data and reference annotations on which the segmentation models will be trained and evaluated. To quantify the quality of these segmentations, a set of evaluation metrics is introduced in the following section.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

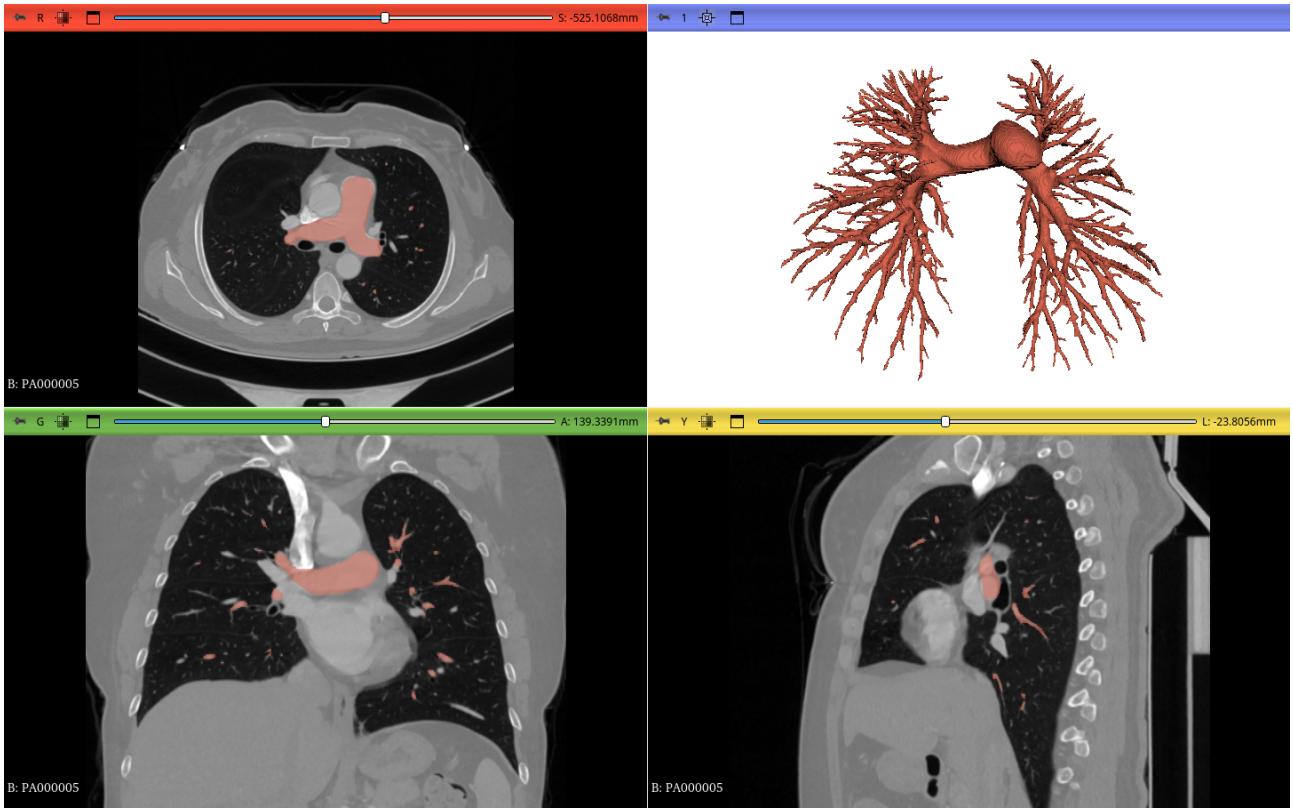


Fig. 2.7: Representative example of a CTPA volume with its corresponding pulmonary artery annotations for patient PA000005 from the PARSE training dataset, displayed in 3D Slicer.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

II Segmentation evaluation metrics

In this section, we present several evaluation metrics used to assess the quality of segmentation outputs in this work. These metrics are organized into two categories: overlap-based and topology-based measures. Since all segmentations are evaluated on 3D binary masks, the following definitions are expressed in terms of voxels rather than pixels. Through this section, we denote A the reference standard (the manual annotation) and B the predicted segmentation produced by the model. Both A and B are 3D binary masks defined over the same voxel grid dimensions, where a value of 0 indicates background, and a value of 1 indicates a structure of interest.

II.1 Overlap metrics

To assess the performance of a segmentation model, the degree of overlap between A and B can be measured using a confusion matrix, presented in Table 2.5.

		B : predicted segmentation	
		Positive or class (1)	Negative or background (0)
A : reference	Positive or class (1)	True Positive (TP)	False Negative (FN)
	Negative or background (0)	False Positive (FP)	True Negative (TN)

Tab. 2.5: Confusion Matrix for the reference A and the prediction B

Four categories of voxels can be distinguished:

- The True Positives (TP): voxels labeled as the structure of interest (1) in both A and B.
- The True Negatives (TN): voxels labeled as background (0) in both A and B.
- The False Positives (FP): voxels labeled as the structure of interest (1) in B, but as background (0) in A.
- The False Negatives (FN): voxels labeled as background (0) in B, but as the structure of interest (1) in A.

From this confusion matrix, several evaluation metrics can be derived, which are presented below.

II.1.1 Precision

The precision measures the proportion of voxels predicted as the structure of interest that are correctly identified:

$$Precision = \frac{TP}{TP + FP} \tag{2.1}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

II.1.2 Recall

The recall, also referred to as sensitivity or true positive rate, measures the proportion of reference voxels belonging to the structure of interest that are correctly predicted :

$$Recall = \frac{TP}{TP + FN} \tag{2.2}$$

II.1.3 Accuracy

The accuracy measures the proportion of correctly classified voxels over all voxels:

$$Accuraccy = \frac{TP + TN}{TP + TN + FN + FP} \tag{2.3}$$

II.1.4 The Dice Similarity Coefficient (Dice)

The Dice similarity coefficient (Dice) is commonly used in segmentation evaluation to assess the volumetric agreement between A and B. It is defined as the harmonic mean of the precision and recall, which corresponds to two times the overlap area by the total area of both masks:

$$Dice = 2 \times \frac{Precision \ \times \ Recall}{Precision \ + \ Recall} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{2.4}$$

II.1.5 The centerline Dice (clDice)

clDice was introduced by Shit et al. 2021 as an overlap metric specifically designed to evaluate the connectivity of tubular structure segmentations, such as vascular trees. It relies on the skeleton of both the reference A and the segmentation B, where the skeleton of a binary mask is defined as a one-voxel-wide representation of the centerline of the mask, preserving its topology.

Let S_A and S_B denote the skeletons of A and B, respectively, TP_{SB} the S_B voxels that are overlapping with A, and FP_{S_B} the S_B voxels that are outside of A. The Topology Precision measures the proportion S_B that fall within A:

$$\text{Topology Precision} = \frac{TP_{S_B}}{TP_{S_B} + FP_{S_B}}. \tag{2.5}$$

Let TP_{SA} the S_A voxels that are overlapping with B and FN_{S_A} the S_A voxels that are not within B. The Topology Sensitivity measures the proportion of S_A covered by B :

$$\text{Topology Sensitivity} = \frac{TP_{S_A}}{TP_{S_A} + FN_{S_A}}. \tag{2.6}$$

The centerline Dice similarity coefficient (clDice) is defined as the harmonic mean of Topology Precision and Topology Sensitivity:

$$\text{clDice}(M_A, M_B) = 2 \times \frac{Topology \ Precision \times Topology \ Sensitivity}{Topology \ Precision + Topology \ Sensitivity}. \tag{2.7}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

It should be noted that while clDice is designed to be sensitive to connectivity, it remains an overlap-based metric and does not directly measure topological properties in a strict mathematical sense. Its sensitivity to connectivity errors depends on whether those errors affect the skeleton, which is not always guaranteed. All overlap metrics presented above range from 0 to 1, with 0 indicating no overlap and 1 indicating perfect overlap between A and B.

II.2 Topological metric

While overlap metrics provide a global measure of volumetric agreement between A and B, they do not capture how well the predicted segmentation preserves the reference's topological structure. The Betti numbers are topological descriptors that may describe the topology of segmentations. Figure 2.8 illustrates their geometric interpretation across 2D and 3D shapes. For 3D binary masks, three Betti numbers are relevant:

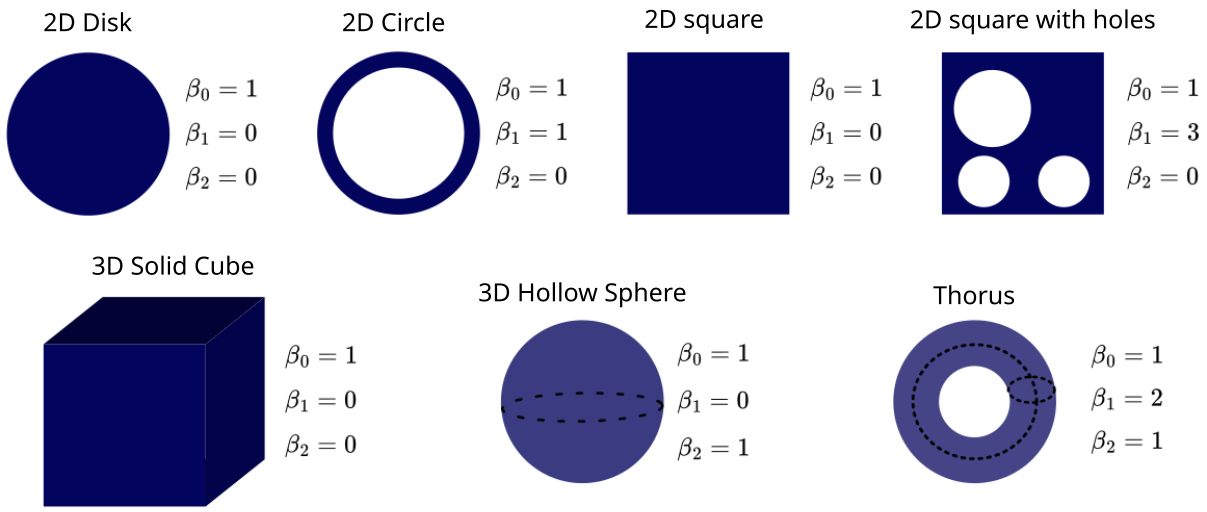


Fig. 2.8: Illustration of Betti numbers for 2D and 3D shapes. Figure by Pierre Rougé inspired by Stenseke 2021.

The first Betti number β_0 corresponds to the number of connected components of a segmentation. A value of $\beta_0 = 1$ indicates that the mask forms a single connected structure, which is the expected configuration for the PAT. A value of $\beta > 1$ indicates the presence of additional connected components, which may correspond either to disconnected vessel segments or to isolated predictions that do not correspond to any anatomical structure. The higher this number, the higher the number of disconnected components.

The second number of Betti β_1 corresponds, for 3D structures, to the number of independent cycles or tunnels in the segmentations. Since the PAT is anatomically a tree-shaped, acyclic structure, the expected value is $\beta_1 = 0$. In a voxelized segmentation, however, non-zero values of β_1 can arise as topological artifacts even in the absence of true anatomical cycles. Such artifacts may occur when two nearby vessel walls are connected through diagonal voxel adjacency, forming an unintended loop in the discrete representation of an otherwise cycle-free structure.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

A value of $\beta_1 > 0$ in a segmentation of the PAT therefore signals a topological inconsistency with respect to the expected tree-like model, and should be interpreted as a segmentation error regardless of its anatomical cause. In contrast, for vascular structures that are anatomically cyclic, such as the cerebral arterial network, which contains the circle of Willis, a non-zero value of β_1 is expected and does not constitute an error.

The third Betti number is β_2 counts the number of enclosed cavities within a 3D mask. A value of $\beta_2 > 0$ may result from segmentation errors that leave unfilled voids within the interior of a vessel, or from noise that introduces hollow regions within otherwise solid structures.

These three Betti numbers together provide a complementary characterization of the topological quality of a segmentation, and are particularly well suited to vascular structures, where they allow the detection of disconnected components, spurious cycles, and enclosed cavities. Their interpretation is structure-dependent and should always be considered in light of the expected topology of the anatomical structure under study. For vascular structures such as the PAT, where connectivity and acyclicity are anatomically expected properties, the ideal segmentation should satisfy $\beta_0 = 1$, $\beta_1 = 0$, and $\beta_2 = 0$, making a deviation from these values an indicator of a topological segmentation error.

III Conclusion

This chapter presented the datasets and evaluation metrics used throughout this thesis. The PERSEVERE dataset is the primary in-house resource for training and evaluating the proposed pipeline. After applying exclusion criteria based on manufacturer homogeneity, image quality, and completeness of clinical biomarkers, a final cohort of 353 patients was retained. The public dataset PARSE 2022 provides a set of annotations that we will use during the early stages of deep learning experimentation, as we will see later in Chapter 5. The metrics presented in this chapter combine two complementary perspectives on segmentation quality. Overlap metrics such as the Dice capture volumetric agreement between the reference annotation and the predicted segmentation, while clDice gives the same importance to the small and large vessels through the skeleton. And topological metrics based on Betti numbers provide insights into PAT segmentation, including the number of components, unintended cycles, or holes that would not be captured by overlap measures alone. Having defined the datasets and segmentation metrics, the next chapter reviews the existing literature on deep-learning segmentation methods for vascular structures and thrombi. It then covers existing attempts to automatically compute thrombotic burden and graph-based modeling of vascular structures, providing the background to motivate the methodological choices made in this thesis.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

3 State of the art in pulmonary structures segmentation and PAT modeling

Table of contents

I Segmentation methods . . . 42
I.1 Deep learning architectures for medical image segmentation . . . 42
I.2 Lung and lobe segmentation . . . 47
I.3 PAT segmentation . . . 48
I.4 Thrombi segmentation . . . 50
II From segmentation to graph-based modeling . . . 52
II.1 Centerline-based vascular modeling . . . 52
II.2 Graph construction and representation . . . 53
II.3 Applications of vascular graph models . . . 54
III Conclusion . . . 55

The previous chapter concerned the datasets and the metrics used during this thesis. As established in Chapter 1, the primary objective of my thesis is to develop a PAT model derived from CTPA. The graph-based representation of the PAT forms the basis for biomarker extraction. This graph will encode the spatial distribution and diameters of the pulmonary arteries, together with the location and volume of thrombi within the arterial hierarchy. Building such a model requires both vascular and thrombi segmentation from CTPA images. This chapter explores state of the art in vascular, pulmonary thrombi, and lung segmentation, before addressing PAT graph modeling.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I Segmentation methods

This section introduces the deep learning segmentation approaches for pulmonary structures. Rather than exhaustively reviewing the state of the art on existing segmentation models, we first present the principal deep learning architectures that form the foundation of the methods explored throughout this thesis. This state-of-the-art relies on three major reviews of segmentation techniques in pulmonary imaging (Shi et al. 2025; Van Rikxoort and Van Ginneken 2013; Zhao et al. 2019). We then review their application to the specific pulmonary structures of interest: the lungs and lobes, the arteries, and finally the pulmonary thrombi.

I.1 Deep learning architectures for medical image segmentation

In this section, we introduce the principal deep learning architectures referenced throughout this thesis. As most reviewed models are built on the U-Net architecture, we trace the evolution of its principal variants, from the original 2D U-Net and its 3D extension, through cascaded and mean teacher frameworks, to the self-configuring nnU-Net pipeline. This evolution is described with a focus on vascular segmentation.

I.1.1 From CNNs to U-Net

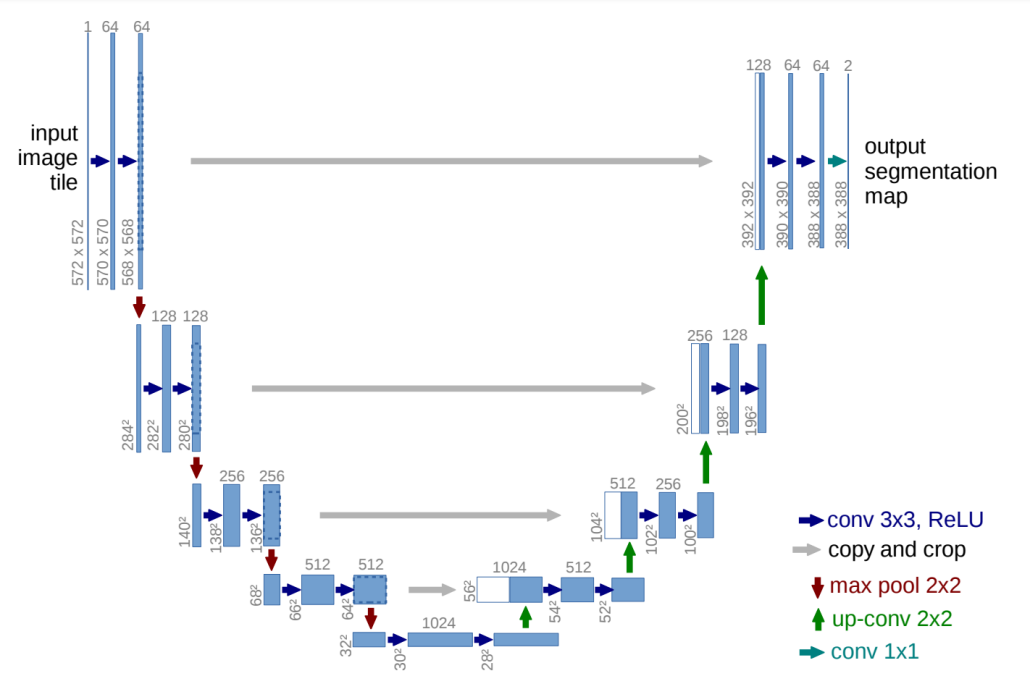


Fig. 3.1: Illustration of the 2D U-Net architecture. Figure extracted from Ronneberger, Fischer, and Brox 2015.

The use of deep learning for medical image segmentation was largely driven by the success of convolutional neural networks (CNNs) in learning task-specific multi-scale image representations directly from data. The 2D U-Net architecture (see Figure 3.1) of Ronneberger, Fischer, and Brox 2015 has established the dominant paradigm for medical image segmentation. Its

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

encoder-decoder structure consists of 5 depth levels. The encoder for each of the 4 downsampling steps applies two 3x3 convolutions followed by a max pooling operation with a stride of 2. Starting from 64 feature channels, each downsampling step doubles it until reaching 1024 features in the bottleneck, progressively capturing global context at the cost of spatial resolution. The decoder upsamples the resulting latent space through 4 successive upsampling steps to recover spatial resolution, halving the number of feature channels at each level. Skip connections concatenate feature maps from the encoder to the decoder to preserve fine-grained local details.

Trained end-to-end with a Dice loss or cross-entropy loss, the U-Net achieves strong performance with relatively small annotated datasets, which is critical in medical imaging. However, processing CT volumes slice by slice discards inter-slice spatial context, which is particularly limiting to preserve tubular structures such as pulmonary vessels. The 3D U-Net extension, introduced in the following part, addresses this limitation.

I.1.2 3D U-Net

The volumetric extension of 2D U-Net, the 3D U-Net (Çiçek et al. 2016), replaces the 2D convolutions by 3D convolutions, giving the model access to inter-slice spatial context in volumetric data such as our CTPA. This is important for tubular structures such as pulmonary vessels, whose continuity and branching topology can be better captured in 3D when sufficient inter-slice resolution is available. One of the limitations of 3D U-Net is its high memory footprint, which requires patch-based training.

To reduce memory and training time requirements, lightweight U-Net variants can decrease network depth or the number of feature channels, while preserving the encoder-decoder structure with skip connections, as we will see in the next part.

I.1.3 Lightweight variants

TinyU-Net is a compact custom adaptation of the U-Net architecture by Rougé, Passat, and Merveille 2024 used for cerebrovascular segmentation. It uses four levels of depth, instead of 5 and downsampling is performed using strided convolutions with a stride of 2, replacing the max-pooling operations used in the original U-Net. The number of feature channels is reduced from a range of 64 to 1024 down to 32 to 256. With this configuration, TinyU-Net is suited to settings where GPU memory is limited or where rapid training iterations are needed. TinyU-Net offers a computationally efficient baseline, well-suited for rapid training on small datasets. In the next section, we will review another type of architecture that can be used when we have a small dataset, a semi-supervised solution.

I.1.4 Semi-supervised learning: Mean Teacher

Obtaining large quantities of voxel-level annotations for CTPA is expensive. Semi-supervised learning addresses this by leveraging both labeled and unlabeled images during training, reducing dependence on costly manual annotations.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

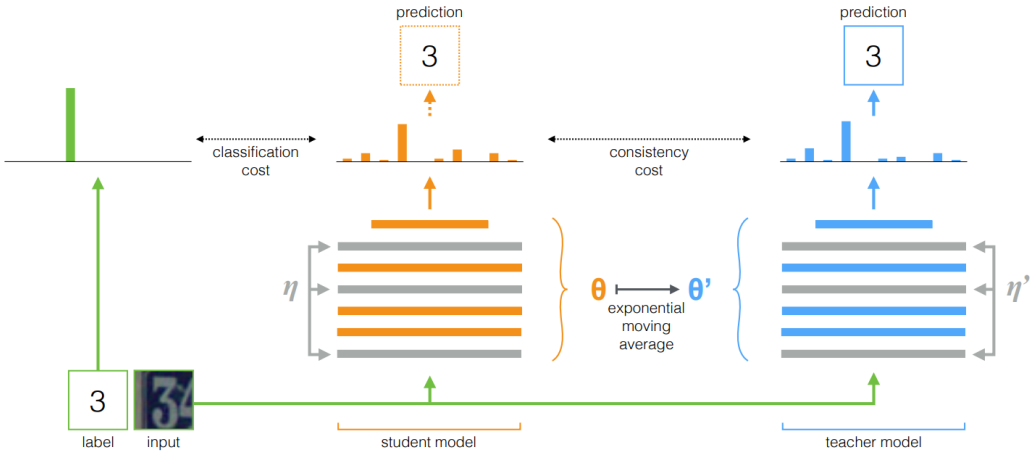


Fig. 3.2: Illustration of the Mean teacher architecture. Figure extracted from Tarvainen and Valpola 2018. The student model weights are θ and noise η, and the teacher model weights are θ' and noise η'.

The Mean Teacher framework from Tarvainen and Valpola 2018 implements this idea by maintaining two networks with identical architectures: a student and a teacher. When processing labeled images, the student is trained with a standard classification loss: the cross-entropy loss between the softmax output and the one-hot ground-truth label. On unlabeled images, no classification loss is applied. In both cases, a consistency loss penalizes disagreement between the student and teacher predictions, providing an additional training signal without requiring ground truth annotations. Let θ and θ' denote the weights of the student and teacher network, respectively, x the input image, and η and η' independent noise perturbations applied to the student and teacher inputs, respectively. The consistency cost $J(\theta)$ is defined as the expected mean squared error between the student prediction $f(x, \theta, \eta)$ and the teacher prediction $f(x, \theta', \eta')$:

$$J(\theta) = \mathbb{E}_{x,\eta',\eta}\left[\|f(x,\theta',\eta') - f(x,\theta,\eta)\|^2\right] \tag{3.1}$$

The weight of this consistency loss is not fixed: it ramps up from zero to its final value over the first 80 epochs, so that the model initially focuses on the classification signal from labeled examples before gradually incorporating consistency regularization over unlabeled data. The teacher model weights θ' are not updated by gradient descent like the student model weights θ, but instead computed as an exponential moving average of the student weights θ over training steps. Because the teacher weights aggregate student weights over time, the teacher produces more stable predictions than the student, making it a reliable source of pseudo-supervision on unannotated data. This framework is relevant in the context of CTPA segmentation, where fully annotated datasets are scarce, and a large pool of unannotated scans may be available. The next part introduces a cascaded architecture that also relies on two networks, but organizes them sequentially to decompose the segmentation task into distinct subproblems.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I.1.5 Cascaded architectures

The cascaded segmentation architecture principle is to apply multiple network stages sequentially, with each stage either refining the output of the previous one or solving a complementary subproblem.

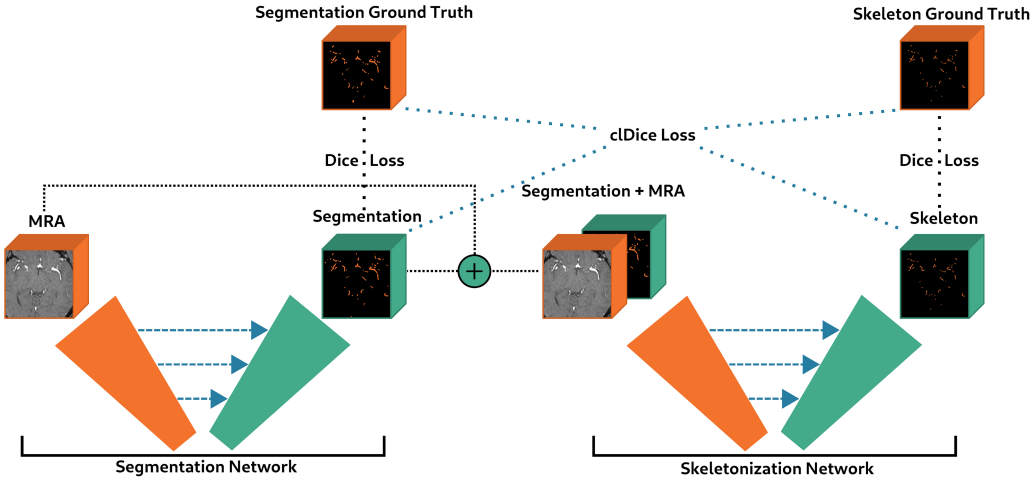


Fig. 3.3: Illustration of the cascaded architecture implemented for cerebrovascular segmentation. Figure extracted from Rougé, Passat, and Merveille 2024.

Rougé, Passat, and Merveille 2024 proposed a cascaded multitask U-Net of two successive TinyU-Net stages (see Figure 3.3) and incorporated topological priors. The first stage takes the image as input and produces an initial vessel segmentation supervised by a Dice loss. This output is then concatenated with the original image and fed into the second network, which performs the skeletonization task and is also supervised by a Dice loss on the predicted skeleton. Both stages are jointly supervised by a clDice loss, which takes as input the predicted segmentation and the predicted skeleton. Let S_P and S_G denote the predicted and ground truth segmentations, C_P and C_G the predicted and ground truth skeletons, and $\lambda_1, \lambda_2 \in R$ the weighting parameters The final training loss combines the three loss terms:

$$L(S_P, S_G, C_P, C_G) = \text{Dice}(S_P, S_G) + \lambda_1 \cdot \text{Dice}(C_P, C_G) + \lambda_2 \cdot \text{clDice}(S_P, S_G, C_P, C_G) \quad (3.2)$$

By jointly supervising both the segmentation and the skeleton, this multitask architecture explicitly encourages vessel correctness instead of relying only on volumetric overlap.

While task-specific architectures such as cascaded or lightweight designs require manual configuration and parameter tuning, self-configuring frameworks aim to achieve competitive performance while automating this process. The following part presents one of the most widely adopted self-configuring frameworks in medical image segmentation, nnU-Net.

I.1.6 Self-configuring frameworks: nnU-Net

The nnU-Net framework (Isensee et al. 2021) removes the need for manual tuning of training hyperparameters and architecture design by automatically configuring the full training pipeline from dataset properties alone (see Figure 3.4). nnU-Net's robustness on smaller datasets stems

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

from several built-in design choices. Given a new dataset, nnU-Net extracts a set of dataset fingerprint properties: image modality, voxel spacing, intensity statistic, image size, and class frequencies. From these fingerprint properties, a set of fixed heuristic rules infers the adapted pipeline parameters: patch size, batch size, network depth, and normalization strategy. The framework then selects the adapted configurations among a 2D U-Net, a 3D U-Net at low resolution, a 3D U-Net at full resolution, and a 3D U-Net cascade: first a 3D U-Net at low resolution followed by a 3D U-Net at full resolution.

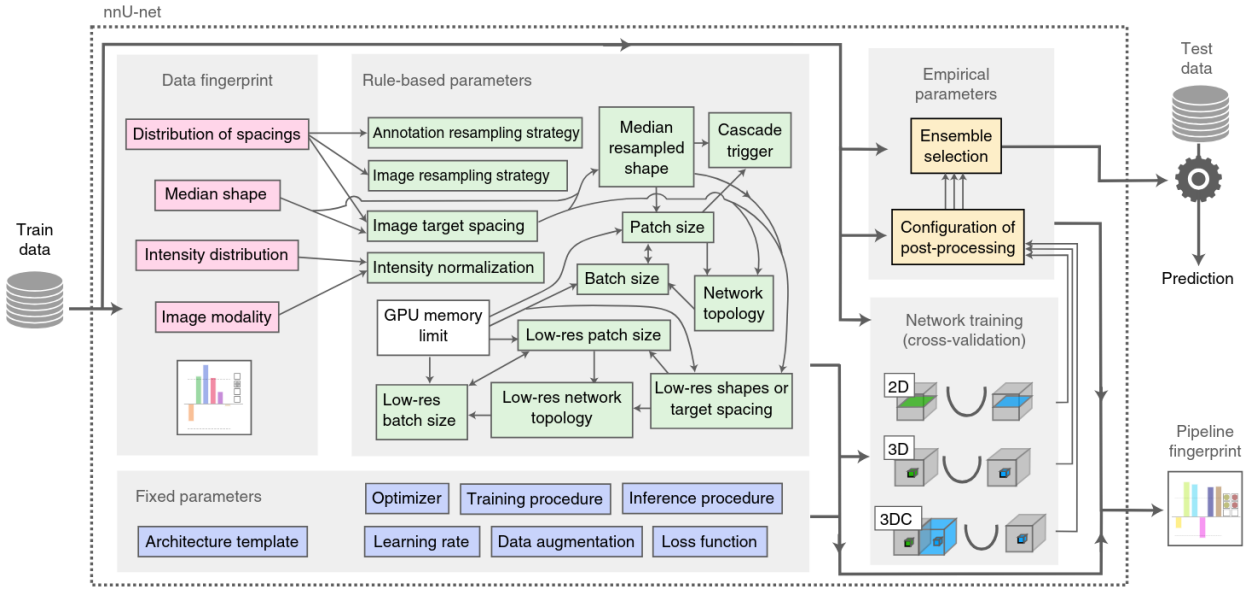


Fig. 3.4: Illustration of the nnU-Net automatic configuration. Figure extracted from Isensee et al. 2021.

Once the adapted configurations are created, the training pipeline incorporates an extensive data augmentation strategy, including random rotations, scaling, elastic deformations, gamma correction, and mirroring, which artificially increases the effective training set size. Training is performed using a five-fold cross-validation to fully exploit the available annotated data. The default training combines a soft Dice loss with a cross-entropy loss. Once each training based on the configurations is done, prediction and ensemble predictions are evaluated. The best-performing configuration is chosen based on cross-validation performance.

These design choices collectively contribute to nnU-Net's strong performance even in limited annotation settings, though the minimum dataset size required remains task and anatomy-dependent. nnU-Net has become a de facto gold-standard baseline across many biomedical segmentation tasks, often matching or outperforming hand-tuned architectures.

Having reviewed the main deep learning architectures of interest in our work, the following section moves to the segmentation of pulmonary anatomy, presenting two dedicated models for lung and lobe segmentation.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I.2 Lung and lobe segmentation

Lung and lobe segmentation serves three distinct roles in the analysis pipeline developed in this thesis. First, lung masks are used to define a bounding box in the PERSEVERE images, thereby constraining processing to the pulmonary region. Second, they provide anatomical context for PE and vascular segmentation as lung boundaries can help differentiate vascular structures from adjacent tissues. Third, lobe masks serve as anatomical references to assign a precise spatial location to each artery in the PAT, which is necessary for the computation of clinically relevant biomarkers. Several automatic tools are available for lung and lobe segmentation. In this thesis, we rely primarily on two of them: `Lungmask` and `TotalSegmentator`, described below.

I.2.1 Lungmask

Lungmask (Hofmanninger et al. 2020) is an automatic lung segmentation tool based on 2D U-Net models. Several U-net model variants are available:

- `R231`: a 2D U-Net trained on the R-231 dataset, which contains 231 cases collected from clinical routine acquisitions, covering 22 different combinations of manufacturer, convolution kernel, and slice thickness. This model segments the right and left lungs.
- `LTRCLobes`: a 2D U-Net trained on a subset of 36 patients from the Lung Tissue Research Consortium (LTRC)[1] dataset. This model segments the right and left lungs, as well as the lobes.
- `LTRCLobes_R231` : a fusion of the two previous models through an ensemble strategy. This model segments the right and left lungs and the lobes. Regions missed by `LTRCLobes` are recovered using `R231` predictions and assigned to the nearest anatomical lobe label, while detections from `LTRCLobes` that fall outside the `R231` lung mask are discarded.
- `R231CovidWeb`: a model that trained on COVID-19 slices from MedSeg[2] with reconstructed HU values, addressing the practical limitation of web-sourced COVID-19 data that typically lacks original HU information. This model segments the right and left lungs.

Among all these models evaluated in Hofmanninger et al. (2020) study, `R231` achieved the best results for lung segmentation on the public LTRC dataset, reaching a Dice of 0.99. This study emphasized that this model trained with data diversity of acquisition conditions was the main factor, outweighing architectural choices, for yielding better performances. The **R231** model was explicitly designed to include high-density anomalies within the lung mask, which is relevant in the context of PE where vessels and thrombi can alter local image intensity, but the paper does not directly test on PE cases.

[1] https://biolincc.nhlbi.nih.gov/studies/ltrc/
[2] http://medicalsegmentation.com/covid19/

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I.2.2 Totalsegmentator

`TotalSegmentator` (Wasserthal et al. 2023) is a widely used deep learning tool that segments more than one hundred anatomical structures on CT scans, including lungs, pleura, heart, great vessels, and specific vascular classes such as lung vessels and coronary arteries. The model was trained using the nnU-Net framework on a diverse dataset of 1204 CT examinations randomly sampled from routine clinical studies (2012, 2016, 2020), at University Hospital Basel, covering 104 structures (27 organs, 59 bones, 10 muscles, and 8 vessels). Images were resampled to 1.5-mm isotropic resolution, with ground truth generated via an iterative annotation workflow that combined initial model predictions with manual corrections by clinical experts. The dataset was split into 80% training (1082 patients), 5% validation (57 patients) and 5% test (67 patients). This approach achieved high accuracy (mean Dice coefficient 0.943) and robust generalization across scanners, protocols, and pathologies. This makes TotalSegmentator particularly relevant for providing lobar anatomical context in pathological cases such as PE, where lung morphology may be altered.

In the next section, we will now focus on segmentation models specifically designed for PAT segmentation. Our main concern for PAT segmentation is producing a 3D mask that is both voxel-level accurate and forms a continuous anatomically correct vascular tree. This section reviews deep learning approaches applied to this task, as well as loss functions and post-processing strategies designed to enforce topological consistency.

I.3 PAT segmentation

Pulmonary artery segmentation aims to delineate the complete PAT in CTPA. The review from Shi et al. 2025 categorizes deep learning–based pulmonary vessel segmentation into 2D slice-based, 2.5D orthogonal multi-view, and full 3D volumetric models, and reports that the majority of recent studies use 3D U-Net variants, frequently achieving Dice scores above 0.9 for arteries and veins. These results are, however, obtained on controlled research datasets, which may not reflect performance on clinical data with pathologies such as PE or motion artifacts. The following works illustrate this progression: Y. Xu et al. 2018 proposed a two-stage pipeline in which a Mask-RCNN first detects candidate vessel regions, before an adaptive U-Net refines the segmentation within those regions. Cui, X. Liu, and N. Huang 2019 proposed a fused 2.5D U-Net++ for lung vessel segmentation, exploiting multi-view information to improve robustness to anisotropic resolution. Nam et al. 2021 employed a full 3D U-Net for pulmonary vessel segmentation in CTPA, demonstrating the benefit of volumetric context for preserving vessel connectivity across slices.

Despite these advances in deep learning, several challenges remain in pulmonary artery segmentation. The detection of small, low-contrast peripheral vessels remains difficult, and performance can degrade in the presence of acquisition artifacts, which are common in routine clinical scans. In the context of this thesis, the presence of PE in the PERSEVERE dataset further complicates segmentation by locally disrupting the pattern of the vascular lumen. Studies report substantial inter- and intra-observer variability around bifurcations, in regions with emboli or motion artifacts, and at the interface between arteries and veins (Shi et al. 2025). This

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

variability directly impacts the reliability of ground-truth labels, a challenge that we examine in the following part.

I.3.1 Annotation quality and ground truth reliability

Annotation quality is itself a critical factor: inconsistent or incomplete ground truth labels propagate directly into model training, leading to segmentations that reproduce annotator errors rather than true anatomical boundaries. This is particularly concerning for PAT, where distal vessels are frequently under-annotated due to their small diameter and low contrast, and where the presence of thrombi introduces systematic ambiguity at the lumen boundary. In the case of PAT, this is challenging: the tree-shaped structure comprises numerous ramifications, a progressively decreasing diameter toward the periphery, and a complex branching architecture that makes obtaining complete and consistent annotation difficult even for expert annotators.

To address these issues, recent initiatives propose explicit criteria for annotation quality, advocating standardized labeling protocols, multi-expert consensus, and the use of uncertainty maps to reflect ambiguous regions in training and evaluation (Lekadir et al. 2025). In practice, such initiatives highlight the need for dedicated annotation tools and protocols that guide annotators through complex vascular structures while enforcing consistency criteria. This concern is central to the present thesis, in which the absence of existing annotations for the PERSEVERE dataset required the development of a dedicated annotation workflow, described in the next Chapter 4.

Beyond the reliability of annotations themselves, a second limitation of standard segmentation pipelines concerns the way model outputs are evaluated and optimized: even when ground truth labels are of high quality, volumetric metrics and losses can fail to capture errors that have an impact, such as disconnections or incorrect vessel delineation, for vascular tree analysis, which motivates the following discussion on topological consistency.

I.3.2 Topological consistency in artery segmentation

Standard segmentation losses such as the Dice loss (see II.1.4) measure volumetric overlap between the predicted and ground truth masks, but are insensitive to the topological structure of the segmentation. Small topological errors such as spurious branches or broken connections can have a disproportionate effect on downstream graph-based analyses while causing only small changes in voxel-wise Dice scores (see Figure 3.5). This makes Dice a potentially misleading evaluation criterion for vascular tree segmentation if used alone.

Topology-aware losses address this limitation by explicitly penalizing connectivity errors. Among these, the clDice loss (see II.1.5) computes the Dice overlap along the centerlines of the predicted and ground truth structures, giving equal importance to thin and thick vessels and, as a secondary effect, allows for better preservation of small vessels that volumetric losses would otherwise ignore.

Topological consistency can also be improved through post-processing applied after model inference. Keshwani et al. 2020 proposed a method that identifies broken vessel connections in the

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

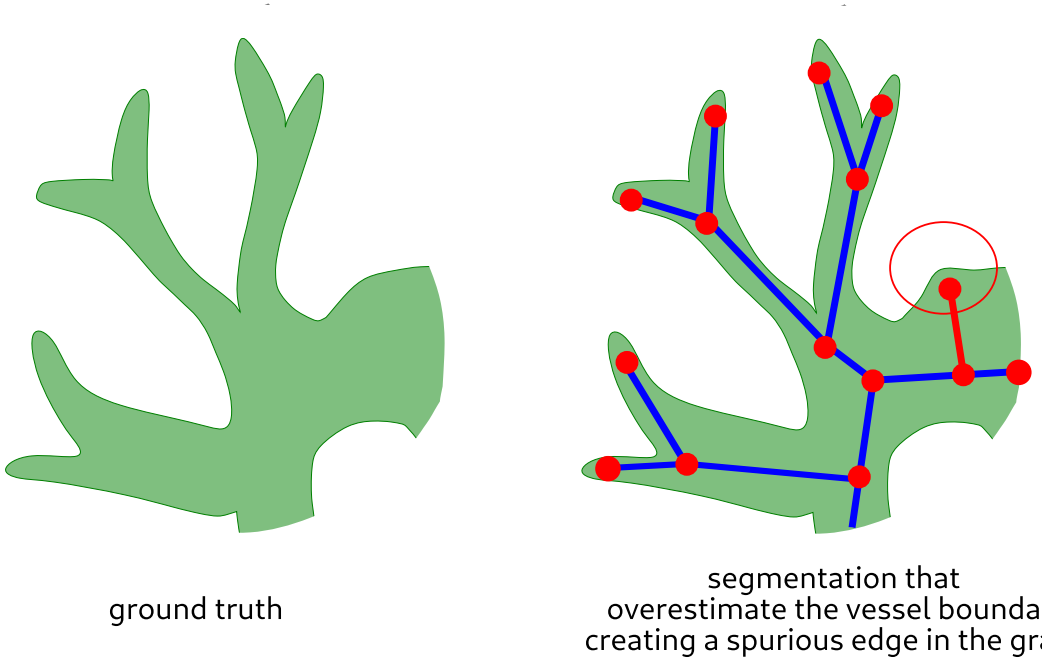


Fig. 3.5: Illustration of how a segmentation error, in this case, overestimating a vessel boundary, which will have little impact on the Dice score, while introducing a spurious branch in the reconstructed vascular tree.

segmentation and repairs them using shortest-path completion based on image intensity cues. Carneiro-Esteves, Vacavant, and Merveille 2025 introduced a learning-based post-processing approach that predicts and corrects topological errors in vessel segmentations, directly targeting the connectivity failures. Such post-processing steps are complementary to topology-aware losses and can be applied on top of any segmentation model. Several works, therefore, complement Dice with topological metrics such as Betti numbers (see II.2) (Hu et al. 2019; Sofi and Alsahanova 2023; Zhang et al. 2025), an approach that has also been applied to PAT segmentation in mice (M. Miller et al. 2023).

While the preceding discussion focused on the segmentation of the arterial tree itself, the detection of pulmonary emboli within it is a complementary and equally challenging problem. The following section presents the deep learning strategies proposed to segment thrombi.

I.4 Thrombi segmentation

PE segmentation presents distinct challenges from vessel segmentation. While vessels appear as bright tubular structures in contrast-enhanced CTPA, emboli manifest as hypo-attenuating filling defects within the enhanced vessel lumen, making their detection dependent on accurate vessel localization. Their size, shape, and contrast vary considerably depending on the degree of occlusion and the timing of contrast injection. The following subsections organize these contributions along two axes: whether they operate in 2D or 3D, and how much voxel-level supervision is required during training.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I.4.1 PE segmentation methods

A first family of methods reformulates PE segmentation as a localization problem. Long et al. 2021 proposed using a Mask-RCNN to define a bounding box around the region of interest before performing segmentation.This approach is primarily designed for 2D images and benefits from the object-detection capabilities of a region-proposal network. This method does not capture the extent of PE across slices, potentially introducing inconsistencies between them.

A second family of methods aims to reduce annotation effort while preserving segmentation quality. Pu et al. 2023 introduced a method that does not require voxel-level manual annotation. A computer vision pipeline first segments pulmonary arteries, partitions them by radius, and applies adaptive thresholding to identify hypoattenuating candidate regions consistent with PE. These candidates provide patch-level supervision for training an R2U-Net model. To assess the value of this annotation-free strategy, a second R2U-Net was trained on a smaller set of manually outlined cases for comparison. The model trained without manual annotation outperforms the conventionally supervised model, particularly on thrombus larger than 10mm, suggesting that extensive manual annotation is not always a prerequisite for effective PE segmentation.

Both families of approaches share common difficulties. Distinguishing partial filling defects from motion artifacts and vessel wall irregularities remains challenging, as these phenomena produce similar hypoattenuating patterns within the enhanced lumen. The high variability of clot appearance across patients, degrees of occlusion, and acquisition conditions further complicates generalization.

Quantifying thrombotic burden and its distribution across the vascular tree requires not only accurate segmentation of both the arteries and the thrombi, but also a coherent model that connects these two elements. The following section reviews methods that go beyond thrombi segmentation and attempt to automatically estimate the thrombotic burden.

I.4.2 Thrombotic burden automatic quantification

Quantifying the thrombotic burden in PE is essential for risk stratification and treatment planning. Existing approaches to automatically or semi-automatically compute thrombotic burden from CTPA can be organized into two categories: score-based methods and volume-based methods.

The first category of methods aims to reproduce established clinical severity indices, such as the Qanadli and Mastora scores, through automatic or semi-automatic analysis of CTPA. Peters et al. 2007 explored the semi-automatic computation of the Qanadli score by labeling segmented arteries on CT scans. While this approach directly targets a clinically validated index, their method requires artery segmentation extending well beyond the subsegmental level, which is computationally demanding and sensitive to segmentation errors in peripheral vessels. Score-based approaches are attractive because their output is directly interpretable by clinicians, but they inherit the limitations of the scores themselves: both the Qanadli and Mastora scores assign discrete weights to arterial segments regardless of the actual volume and spatial distribution of the thrombus.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

The second category of methods focuses on measuring the Total Embolic Volume (TEV) as an anatomically grounded descriptor of thrombotic burden. W. Liu et al. 2020 proposed a U-Net based framework to automatically extract TEV from CTPA, demonstrating correlation with the Qanadli score, the Mastora score, and RV functional parameters. W.-M. Huang et al. 2022 demonstrated that TEV correlates with several clinical parameters, including the RV/LV ratio on CT, right ventricular dysfunction on echocardiography, and hemodynamic instability. Importantly, TEV shows strong correlation with the Qanadli score when the embolic volume remains below 10 cm^3. However, because the Qanadli score assigns a single point to each subsegmental occlusion regardless of its actual size, it can underestimate thrombotic burden in peripheral PE.

This divergence between TEV and score-based indices in large embolic loads highlights that the two families of measures are complementary rather than redundant, and that neither alone is sufficient for complete severity assessment.

Regardless of the output representation, none of the reviewed methods integrates embolic localization into a structured model of the PAT. Jointly quantifying thrombotic burden and its spatial distribution across lobes and the arterial hierarchy requires moving beyond voxel-level segmentation and toward a graph-based representation of the PAT. The following section reviews the methods for constructing such models from segmentation outputs.

II From segmentation to graph-based modeling

Graph-based representations of vascular structures provide a compact, structured description of the vascular tree, enabling topological analysis and morphological quantification, with potential extensions toward hemodynamic simulation. Unlike voxel-based representations, graph models explicitly encode vessel connectivity and hierarchical organization, which are critical for downstream analysis of the PAT. Graph modeling of the vascular tree structure generally includes edges representing vessels and nodes marking vessel bifurcations or vessel ends. Our goal is to integrate volumetric vessel information, as well as the presence and impact of thrombi, directly into the edges. This type of graph is generally modeled by extracting a centerline from a binary mask.

II.1 Centerline-based vascular modeling

Most approaches to vascular graph construction rely on centerline extraction from a binary vessel segmentation. Centerlines provide a simplified representation of the vascular tree, reducing the volumetric structure to a one-dimensional skeleton while preserving its topology. Several methods have been proposed for centerline extraction, including iterative thinning algorithms and minimal path approaches. Thinning methods like Lee, Kashyap, and C. N. Chu 1994 iteratively remove boundary voxels while preserving connectivity, resulting in a one-voxel-thick skeleton. Minimal path approaches, such as those implemented in the Vascular Modeling Toolkit framework[3] (Antiga et al. 2008), formulate centerline extraction as an optimization

[3] http://www.vmtk.org/

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

problem guided by image intensity and geometric constraints.

While centerline extraction is simple and efficient, it is sensitive to segmentation noise and surface spurs, so pruning is typically required. In the next section, we will present how the centerline can be transformed into the graph representation.

II.2 Graph construction and representation

Once extracted, the centerline can be converted into a voxel graph representation or a branch graph representation (see Figure 3.6). The voxel graph representation simply transforms each voxel into a node, and the edges are formed between neighboring voxels (Nunez-Iglesias et al. 2018). In the branch graph representation, nodes can correspond to bifurcations and endpoints, and edges represent vessel segments.

The voxel-level graph can be simplified into a branch-level graph by classifying each node according to its degree: terminal nodes (degree 1), bifurcation nodes (degree greater than or equal to 3), and intermediate nodes (degree 2). Linear chains of intermediate nodes connecting two non-intermediate nodes are fused into a single edge, each representing one arterial segment. The resulting graph encodes arteries as edges, each characterized by its centerline coordinates, radius profile, and length.

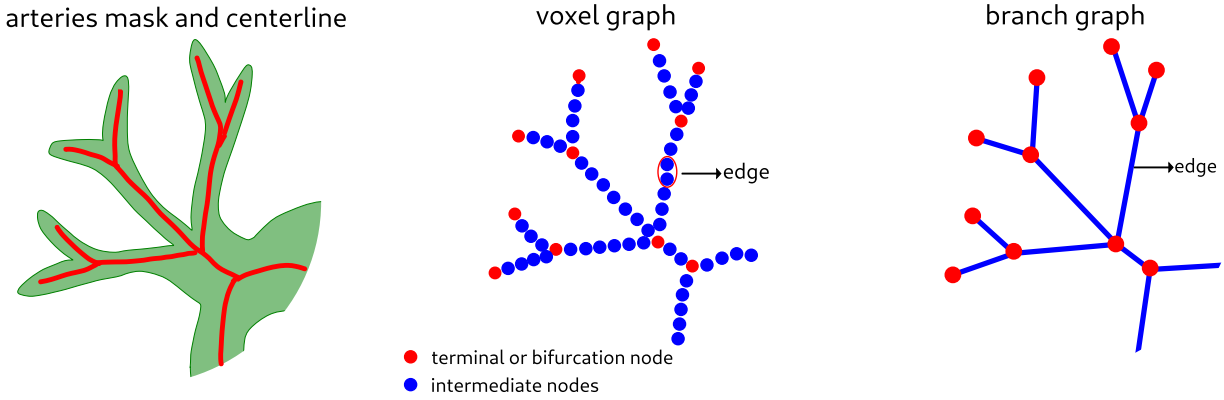


Fig. 3.6: Illustration of two centerline-based graph representations from a vessel segmentation. The voxel graph, where each centerline voxel forms a node and edges connect neighboring voxels. The branch graph, where nodes correspond to vessel bifurcations and endpoints, and edges represent the vessel segments connecting them.

Existing pipelines, such as `VesselGraph` Drees et al. 2021 from the Voreen framework Meyer-Spradow et al. 2009, provide examples of centerline extraction to a graph modeling pipeline. This pipeline takes a 3D binary vessel mask as input, extracts the vascular centerline, and produces a centerline-based branch graph representation (see Figure 3.7). In this figure, we can observe the mentioned spurious branches near the MPA area. This topic will be addressed in the Chapter dedicated to the graph modeling pipeline 6. In the following section, we will present existing applications of vascular graph models across different anatomical domains.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

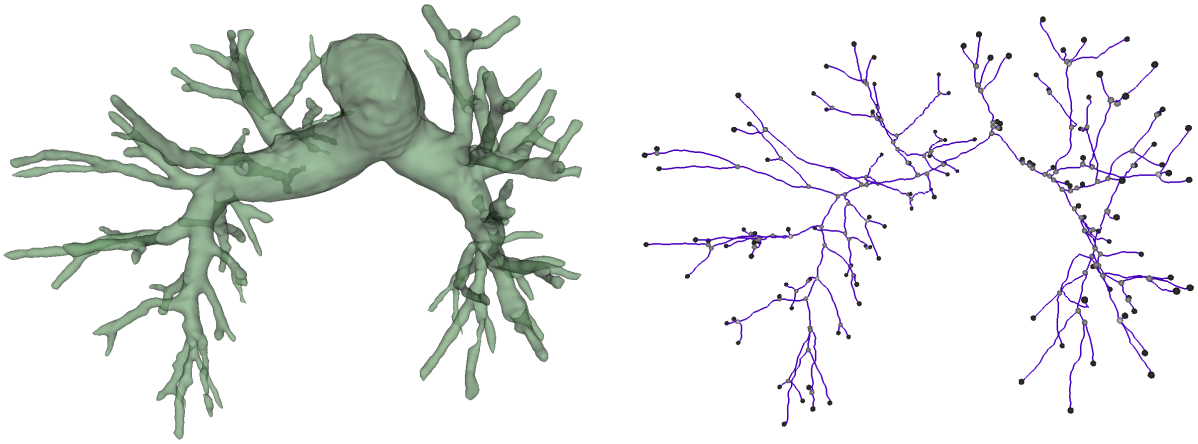


Fig. 3.7: Illustration of the graph construction pipeline applied to a PAT segmentation. Left: the 3D binary mask of the pulmonary arterial tree. Right: the corresponding graph representation obtained through the `VesselGraph` pipeline of the Voreen framework, where nodes mark vessel bifurcations and endpoints, and edges encode individual vessel branches.

II.3 Applications of vascular graph models

Vascular graph models have found applications across multiple domains beyond simple visualization. In hemodynamic simulation, graphs serve as 0D or 1D representations for Computational Fluid Dynamics (CFD). Reichold et al. 2009 developed a vascular graph model to simulate cerebral blood flow in realistic geometries. Similarly, Reymond et al. 2009 validates a 1D hemodynamic model for systemic circulation of the arterial tree. In mesh generation for CFD centerline-based graphs provide structured inputs for automated meshing. Decroocq et al. 2023 proposed a pipeline for modeling and hexahedral meshing of cerebral arterial networks from centerlines, demonstrating how graph topology guides the creation of high-quality CFD meshes while preserving vascular geometry.

Graph representations also enable morphological quantification across vascular beds. In retinal imaging, Mishra et al. 2021 developed VTG-Net, which transforms vessel segmentations into topology-preserving graphs for artery-vein classification, achieving superior accuracy by jointly learning CNN features and graph-based topological features. More recently, Li et al. 2026 introduced Graph-PAVNet for pulmonary artery-vein separation, constructing topology-aware vascular graphs that resolve fragmentation issues in voxel-based methods and improve clinical interpretability through branch-level inference.

Interactive annotation workflows further demonstrate the utility of graph-based approaches for reducing annotation burden Lamy et al. 2022 developed a semi-automatic framework for liver vessel annotation that leverages user-defined graphs. Users place nodes at key branches and bifurcations according to anatomical nomenclature, creating tree structures that guide automated vessel segmentation while preserving branching topology. This graph-guided propagation substantially reduces annotation time compared to fully manual approaches while maintaining anatomical consistency, demonstrating the practical value of vascular graphs for creating high-quality training datasets.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

III Conclusion

This chapter has reviewed the state of the art of the architectures of interest in pulmonary segmentation and PAT modeling, covering the full pipeline from image segmentation to graph-based representation and biomarker extraction. Deep learning approaches have substantially improved vascular segmentation at the voxel level, but pixel-wise accuracy does not guarantee anatomical correctness. Topological errors such as broken vessel connections or spurious branches will have impact on downstream graph construction, while causing only marginal changes in standard overlap metrics such as the Dice. For PE segmentation, distinguishing partial filling defects from artifacts remains an open problem. Graph-based modeling and thrombotic burden quantification further require not only accurate segmentation but also robust centerline extraction and anatomically consistent representation of the full arterial hierarchy, none of which is fully addressed by existing pipelines. These observations motivate the approach developed in this thesis, which aims to construct an anatomically consistent PAT graph model that integrates vessel geometry, spatial location, and thrombus positioning to enable the automatic computation of clinically relevant biomarkers, such as Qanadli and Mastora scores and TEV.

The annotation quality and consistency of manual vessel annotations being an important criterion, the next chapter 4 details the annotation methodology and protocol used in this thesis to obtain reliable ground truth labels for the PAT.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

4 Development of a plugin for semi-automatic annotation of vascular structures

Table of contents

I Limits of existing tools . . . 58
I.1 Preliminary annotation strategies . . . 58
I.2 Statement of need . . . 64
II VSMOD plugin . . . 64
II.1 Centerline module . . . 64
II.2 Segmentation module . . . 67
III Evaluation . . . 68
III.1 Reduction of annotation time . . . 68
III.2 Plugin usage on different vascular networks . . . 69
IV Conclusion . . . 69

As discussed in the previous Chapters 2 and 3, the connectivity and completeness of the extracted vascular tree are critical properties for downstream applications such as quantitative biomarker extraction. Deep learning methods have emerged as powerful tools to automate this extraction, but their development depends heavily on the availability of large, well-annotated datasets. Accurate volumetric annotation of vascular networks is a critical prerequisite for training and evaluating deep learning-based vessel segmentation methods (Chapter 3). In the context of the PERSEVERE dataset, which includes CTPA acquisitions from patients diagnosed with Pulmonary Embolism (PE), the challenge is that the dataset currently lacks manual annotations, and vascular segmentation in PE patients is further complicated.

In this chapter, we first evaluate several existing segmentation tools and highlight their limitations, which ultimately motivated our first contribution. We then present the development of VSMOD, an open-source plugin for 3D Slicer that enables semi-automated segmentation and modeling of vascular structures in volumetric medical images. VSMOD was initially designed to annotate the PAT in Computed Tomography Pulmonary Angiography (CTPA), but we will test its methodology to other vascular territories. This work led to a publication in the **Journal of Open Source Software (JOSS)** (des Ligneris et al. 2025), and the plugin is publicly available at https://github.com/omerveille/VSMOD.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I Limits of existing tools

The objective is to extract a detailed and connected Pulmonary Arterial Tree (PAT) binary mask from CTPA images. Despite recent progress, many state-of-the-art models still struggle to preserve the connectivity of complex vascular trees (Carneiro-Esteves, Vacavant, and Merveille 2025; Rougé, Passat, and Merveille 2024; Shit et al. 2021). In practice, connectivity errors (broken branches, spurious connections) are frequent and directly impact the reliability of subsequent analyses. Our in-house PERSEVERE dataset lacks annotation for the arteries, and we selected a subset of 10 patients (see Chapter 2 I.1.4) to obtain precise annotation that will be used to train segmentation models (see Chapter 5). In this section, we will evaluate several existing automatic, semi-automatic, or manual segmentation tools available in 3D Slicer to obtain PAT masks.

I.1 Preliminary annotation strategies

Because the PERSEVERE dataset lacks annotations, we tested the different methods on a patient from the PARSE dataset to qualitatively assess the segmentation results against the annotation. In this section, we summarize experiments to segment the lungs and pulmonary arteries using different models (`Lungmask`, `TotalSegmentator`) or tools such as `Grow from Seeds`, `Draw Tube`, and `RVesselX` from 3D Slicer. At the end of this section, we highlight the main limitations that motivated the design of our own plugin I.2.

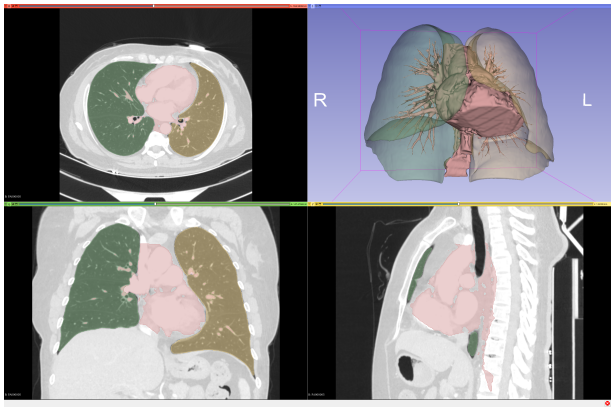
I.1.1 Automatic semgmentation with Lungmask and Totalsegmentator

We first tested to segment pulmonary structure with the `Lungmask` models with four variants available (see I.2.1) and `TotalSegmentator` (see I.2.2) models, as these allow to segment lungs and lobes, but also the lung vessels as part of the `Chest Imaging Platform`[1] extension in 3D Slicer. The segmentations are tested on patient PA000005 from the PARSE 2022 dataset (See Chapter 2 I.2). In Figure 4.1 we can see the result of the `Lungmask` segmentation models. The `R231` and `R231CovidWeb` models produced visually correct right and lung masks, and the `LTRCLobes` and `LTRCLobes_R231` models also produced visually correct lobe segmentations. But in all cases, the vessel mask (in light pink) produced did not distinguish arteries, veins, or smaller branches and fused everything into a single label: the arteries, veins, and all the heart cavities.

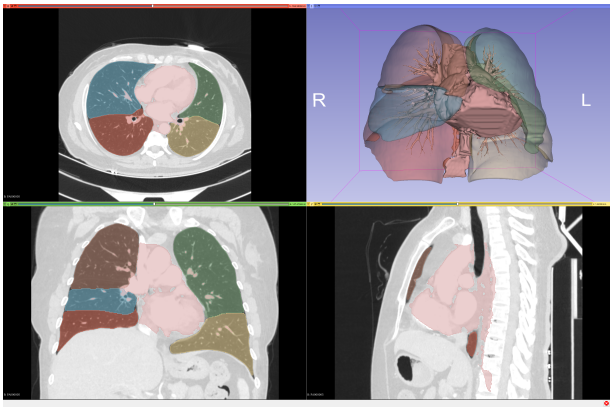
The `TotalSegmentator` proposes two preset `lung basic` or `lung extended` that include lobe, lungs, and vessels segmentation. The result of the segmentation from these presets on the PARSE patient is presented in Figure 4.2. The first preset `lung basic` produced masks of the lungs and surrounding thoracic organs, which were useful for context but did not directly provide a clean separation of the PAT. Only the beginning of the pulmonary artery is present while it is still in the heart area, but not once inside the lungs. The `lung extended` preset generated a vascular mask within the lungs, but this mask mixed arteries and veins and is disconnected from the main pulmonary artery as soon as we get near the hilum.

[1] https://www.chestimagingplatform.org/

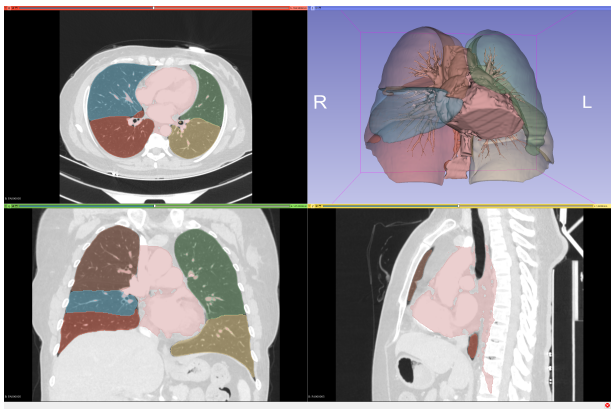
Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés



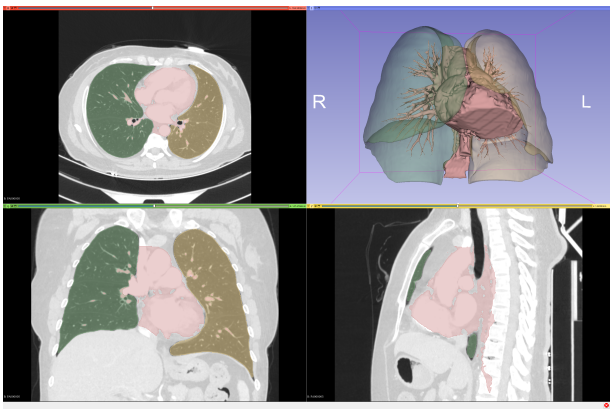
Lungmask R231



Lungmask LTRCLobes

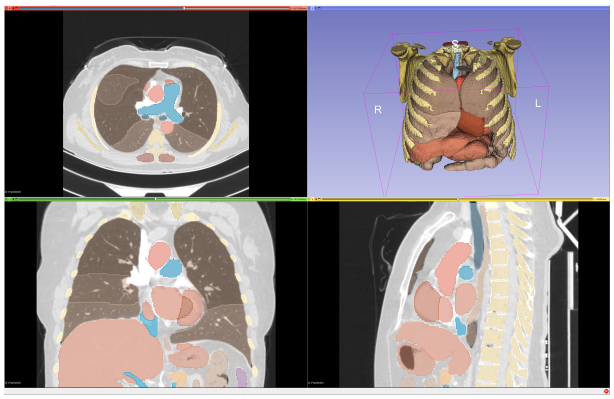


Lungmask LTRCLobes_R231

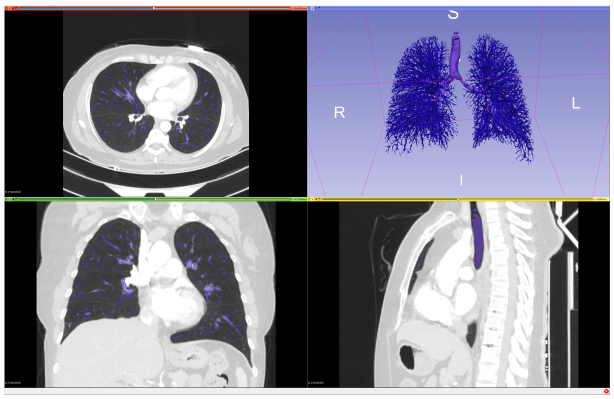


Lungmask R231 CovidWeb

Fig. 4.1: Example of the lungmask models' segmentation results on patient PA000005 from the PARSE dataset



TotalSegmentator total or lung basic



TotalSegmentator lung vessels or lung extended

Fig. 4.2: Example of the TotalSegmentator models' segmentation results on patient PA000005 from the PARSE dataset

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Because the automatic models tested did not yield satisfactory pulmonary artery segmentation results, in the following section, we explored combining the `Lungmask LTRCLobes_R231` vessel-mask output with manual Segment Editor operations (logical operations, margin operations), followed by a `region growing` segmentation constrained by in/out masks.

I.1.2 Semi-automatic segmentation with draw tube and region growing

Region-growing methods start from one or more seed points and iteratively expand into neighboring voxels satisfying an intensity, homogeneity, or gradient criterion (Van Rikxoort and Van Ginneken 2013). They are well-suited to capturing long, tree-like vascular structures in 3D but are prone to leakage or premature stopping due to intensity inhomogeneities (e.g., thrombi-induced filling defects). They are also sensitive to seed placement and typically require parameter tuning and user interaction. The `Grow from Seeds` tool L. Zhu et al. 2014, available in 3D Slicer, implements a multi-label region-growing strategy that addresses some of these limitations by incorporating explicit exclusion constraints. This multi-label formulation is relevant to pulmonary vessel segmentation, where arteries and veins share similar intensity profiles and must be distinguished through spatial constraints rather than intensity alone. For the label of the vessels, we thought of exploiting the `Draw Tube` tool to place a seed in each artery of interest, as a manually placed tubular structure imitating the centerline.

We illustrated the process of this experiment in Figure 4.3. To construct the OUT label, which constrains the extension of our arteries, we use an automatically segmented lung mask from which we subtract the automatically dilated vessel mask. We use the outputs of the model `Lungmask LTRCLobes_R231`: we extend the vessel mask by 2mm and subtract the resulting mask from the lungs mask (see Figure 4.3 A). The artery seed mask is constructed using the `Draw Tube` tool. We want to place a tubular structure inside each artery, similarly to a centerline. The limitation with the `Draw Tube` is that one tube corresponds to one label, so each vessel tube must be placed individually. We will fuse these to obtain our IN mask using the "ADD" logical operator, allowing overlap (see Figure 4.3 B). The placement of each tube is very time-consuming; it took several hours for our one example patient. To distinguish the arteries from the veins, additional vein labels were manually annotated, where the contrast does not allow the distinction to be made correctly, so we can prevent leakage in the surrounding veins (See Figure 4.3 C). With these seed masks, we can use region growing to obtain three labels extended to similar tissues: the lungs mask, the arteries mask, and the veins mask (see Figure 4.3 D).

In theory, we thought that this workflow would allow the region-growing algorithm to propagate the vessel labels along the pulmonary tree while being constrained by surrounding tissues. In practice, we observed several limitations. First, obtaining a dense and coherent set of `Draw Tube` segments across all relevant branches was laborious: after several hours of manual drawing and successive corrections, we could approximate the ground truth for a single case, but the process was not scalable to the full sample dataset. Second, despite corrections on multiple slices, the final segmentation remained less accurate than the expert ground truth, particularly at the subsegmental level and near veins in close proximity. Overall, this approach yielded only a partial solution, at the cost of time-consuming manual effort and without explicit control over

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

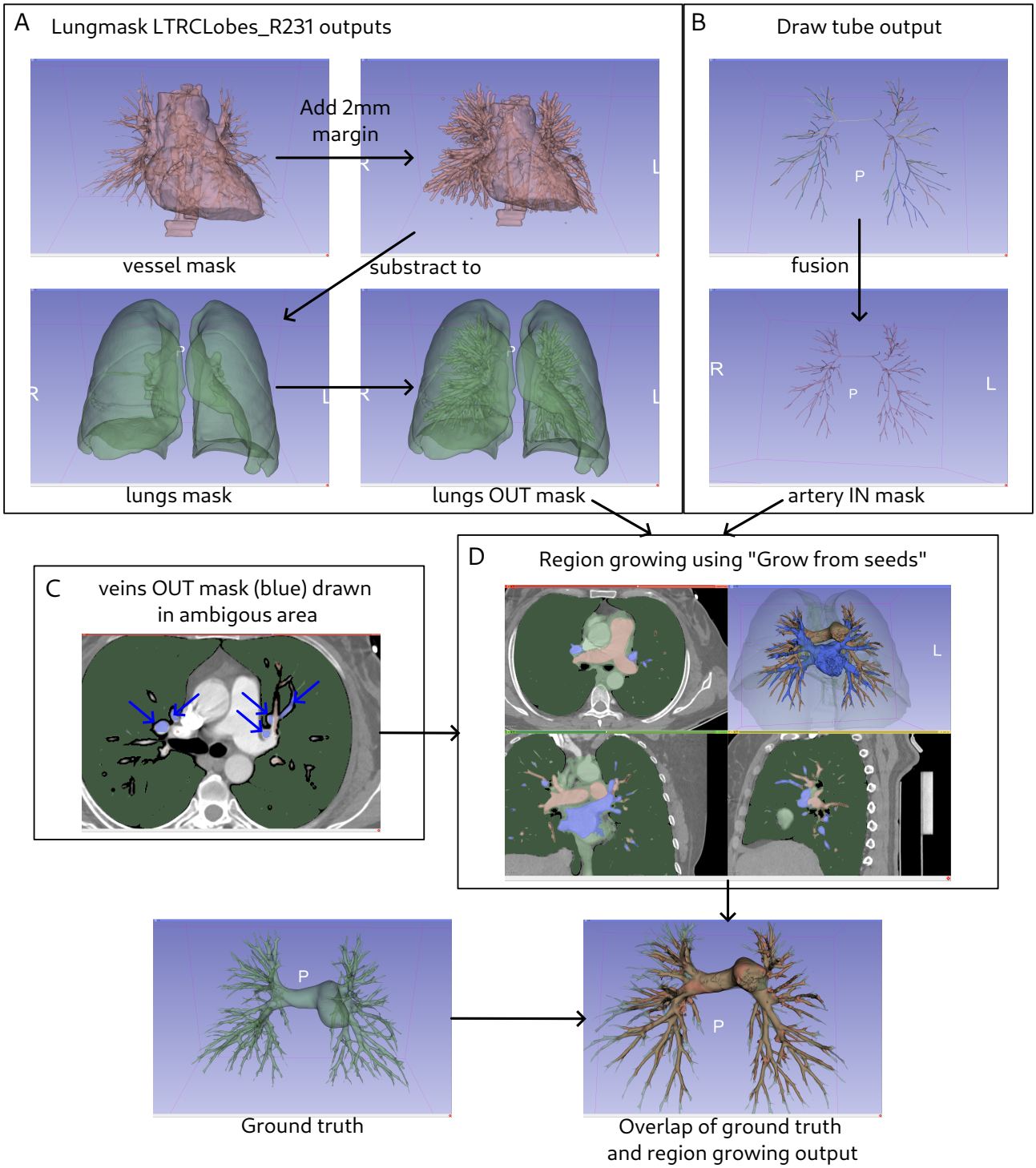


Fig. 4.3: Illustration of the region growing initialisation using `Draw tube` and operations on `lungmask` automatic segmentation to pulmonary vessel masks.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

the resulting tree topology.

I.1.3 RVesselX

Finally, we tested the `RVesselX` 3D Slicer plugin (see II.3), which provides fast, accurate, semi-automatic segmentation of liver anatomy by combining interactive tree annotation with automatic vessel extraction. This tool inspires our approach, as it couples interactive control over the vascular tree with automatic propagation of vessel labels.

Because the plugin requires interactively placing points within the vessels of interest, we labeled each arterial segment for patient PA000005 up to the segmental level to serve as a guide (See Figure 4.4).

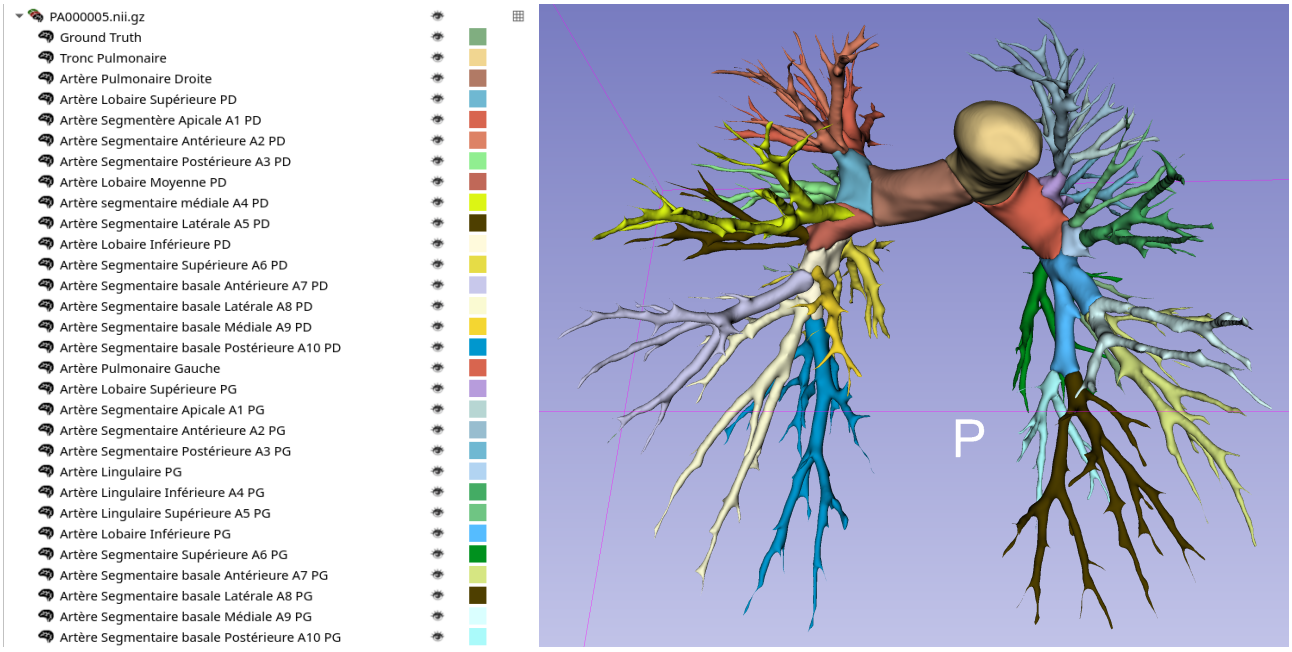


Fig. 4.4: Labeling arteries until the segmental level on patient PA000005 from the PARSE dataset

Even though the branching structure and pattern are specific to the liver, we hypothesized that the approach could be transferred to the pulmonary artery. The labeled ground truth hierarchical organization mask serves as a reference to help place the point of interest required to trace the branching pattern in the plugin.

When applied to pulmonary arteries, we found that the default parameters were not adapted to pulmonary artery (See Figure 4.5 A). Despite testing, including adding a larger number of control points (See Figure 4.5 B) or adjusting parameters such as maximum vessel diameter and growth factor (See Figure 4.5 C), the resulting segmentations remained incomplete and noisy, and did not go any further than the MPA. These experiments confirmed that, although `RVesselX` demonstrates the value of interactive, topology-aware tools, its current implementation could not be directly reused for pulmonary artery segmentation. This motivated the

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

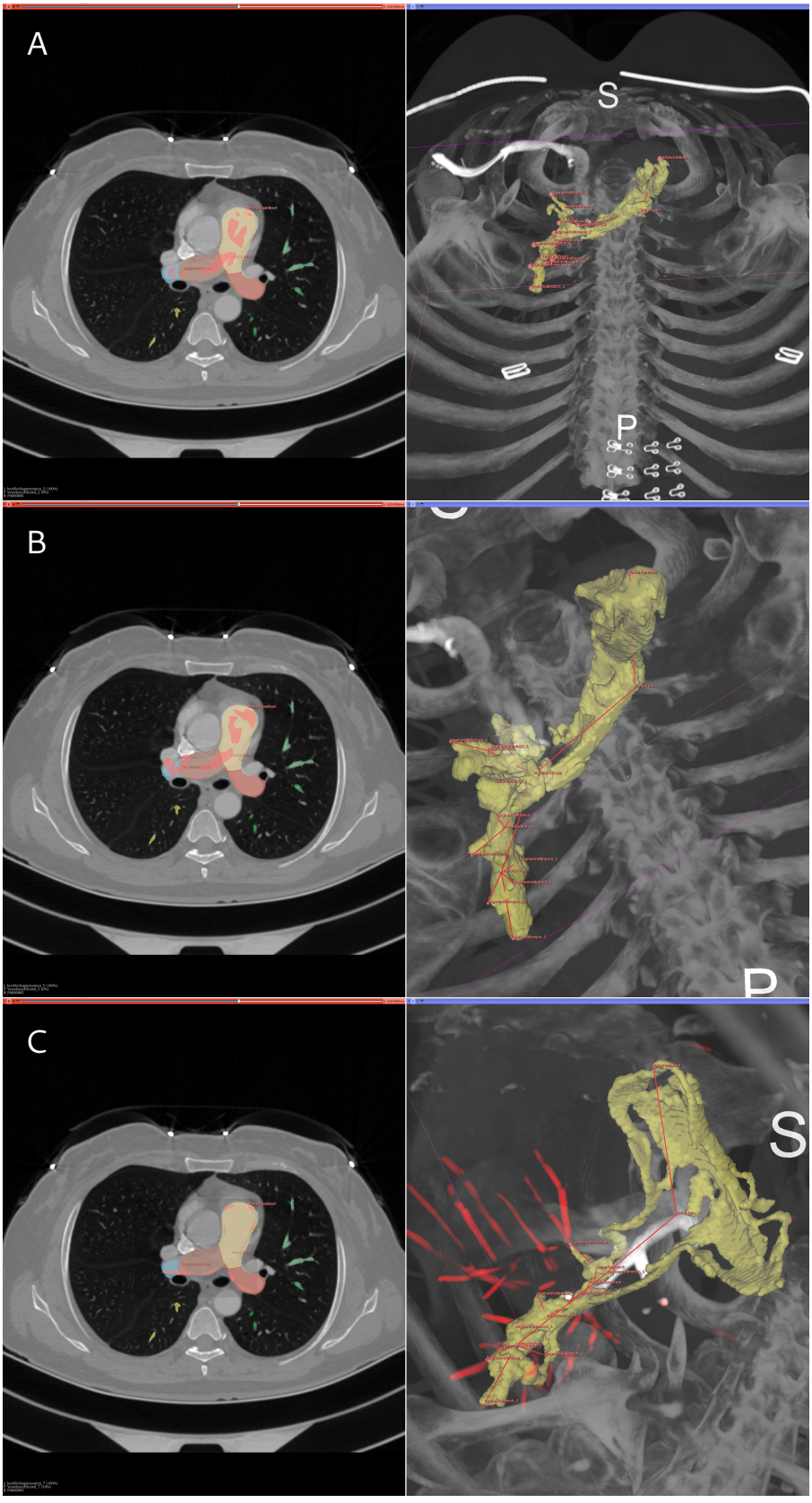


Fig. 4.5: Illustration of the test of PAT segmentation using RVesselX

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

development of a dedicated plugin, VSMOD, able to adapt to the characteristics of thoracic CT angiography while retaining the hierarchical organization of the vascular structure.

I.2 Statement of need

From a data curation perspective, obtaining high-quality vascular annotations remains extremely demanding. Manual delineation with standard tools (paint, brush, or slice-by-slice contouring) is time-consuming and subject to inter-observer variability. Existing tools with tubular shape priors, such as the `Draw Tube` in the SlicerSegmentEditorExtraEffects extension, reduce part of the burden but still require users to draw each branch individually and assume a simple cylindrical geometry, poorly adapted to arteries containing clots, resulting in varying radii. Some specialized plugins, such as `RVesselX` for liver vasculature in 3D Slicer, achieve efficient and accurate segmentation by constraining user interactions to predefined anatomical structures (Lamy et al. 2022). However, these tools are organ-specific and do not generalize easily to other vascular territories. VSMOD addresses this need by combining centerline tracking and region-growing segmentation in a single workflow. The plugin allows users to interactively construct a vascular graph and then derive a volumetric segmentation from this graph using region growing, thus ensuring consistency between geometry, radii and connectivity of the annotated vessels.

II VSMOD plugin

VSMOD is implemented as a 3D Slicer extension composed of two main modules: a **centerline module** and a **segmentation module**. The centerline module relies on a tracking algorithm to extract vessel centerlines and build an explicit graph representation of the vascular network (See Figure 4.6 A). The segmentation module then uses this graph to automatically place seeds (See Figure 4.6 B) and perform a region-growing segmentation, resulting in a complete volumetric annotation (See Figure 4.6 C).

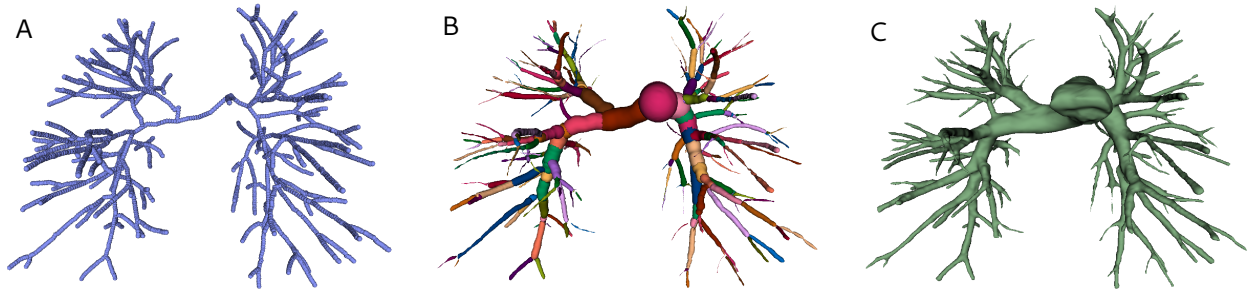


Fig. 4.6: Pulmonary artery annotation workflow with VSMOD. Left-to-right: (A) RANSAC centerlines (B) automatic region-growing seeds (C) final connected segmentation.

II.1 Centerline module

The first module of VSMOD is dedicated to the extraction and modeling of vessel centerlines (See 4.7). To initialize a new vessel, the user must first provide three information: a starting

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

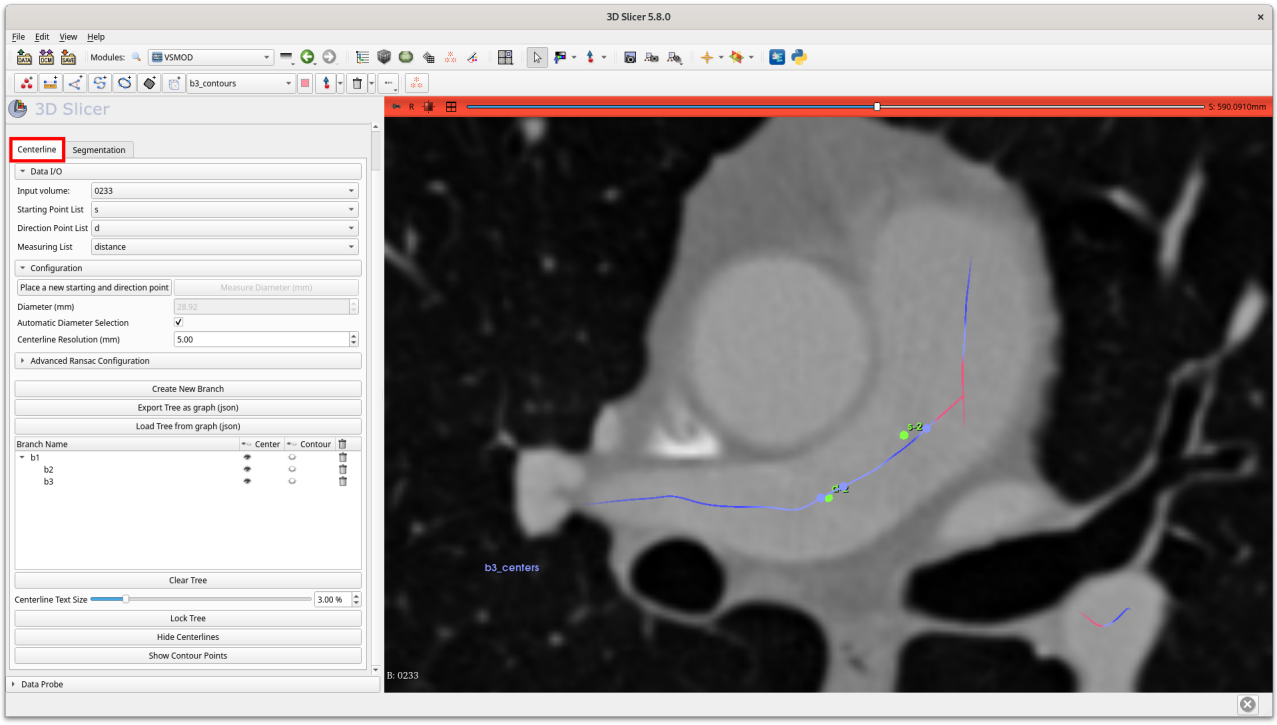


Fig. 4.7: Centerline module

point, a direction point, and an estimate of the local vessel diameter. The starting point is placed in the center of the lumen at the beginning part of the vessel to be annotated. The direction point is then placed downstream along the approximate center of the same vessel. This pair of points thus defines both the spatial position and orientation of the first tracking step. Once these two points are defined, the user measures the vessel diameter in the same region by placing two additional points across the lumen, on opposite sides of the vessel wall, which automatically sets the diameter estimation. This action is required only when launching the algorithm on the first vessel; subsequent branches' diameters will be estimated using the *Automatic Diameter Selection* option, which automatically detects the diameter of the closest previously fitted vessel, rather than requiring a new manual measurement.

Internally, we assume the local cylinder geometry shape and the intensity contrast between the lumen of the arteries and their walls. We use a RANSAC-based algorithm (Yureidini, Kerrien, and Cotin 2012) to sample image points in a local neighborhood around the starting point, using the estimated radius, to fit a sequence of small local cylinders within the vessel. Each candidate cylinder is evaluated based on how many sampled points lie close to its axis (inliers) versus farther away (outliers). The cylinder with the highest inlier count is selected, and its position and orientation define the next step of the centerline, so that a chain of cylinders progressively approximates the vessel path and its local radius.

By repeating this procedure, the algorithm places a sequence of centerline points at a configurable spatial resolution, through the *Centerline resolution (mm)* parameter: at least one point every x mm (5mm by default), together with an estimated radius at each location, until either

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

the maximum number of cylinders is reached or no valid cylinder can be fitted.

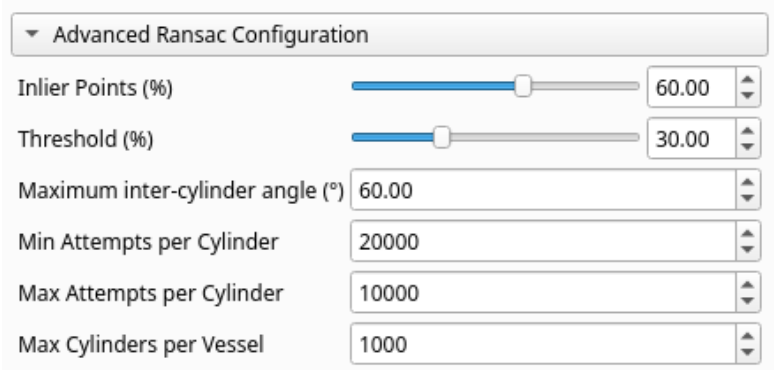


Fig. 4.8: Advanced RANSAC parameter available

To provide flexibility for different vascular structures, VSMOD exposes several advanced RANSAC parameters that control the cylinder fitting process (See. Figure 4.8). The *Inlier Points (%)* parameter (default: 60%) sets the minimum percentage of inlier points required to validate a cylinder fit. The *Threshold (%)* parameter defines the maximum allowed distance from the axis (as a fraction of the current radius) for a point to be considered an inlier. It acts as a margin of error we allow around the radius (see more details in Annex VI). The *Maximum inter-cylinder angle* parameter sets the limit of authorized curvature between consecutive cylinders. *Min/Max Attempts per Cylinder* parameters control RANSAC iterations, as in the number of attempts the algorithm makes if it does not find a suitable cylinder. Defaults are set to min = 20000 and max = 10000. The actual number of tries lie in $[\text{min}, \text{min} + \text{max}]$, so $[20000, 30000]$ for the default. And finally, the *Max Cylinders per Vessel* parameter caps the number of cylinders placed during tracking for a single vessel. All these defaults are optimized for the PAT but can be adjusted for other anatomies (see III.2).

The user can then either continue at the end of the detected vessel centerline or start a new vessel emerging near the centerline. In both cases, the user places two points in the lumen of a vessel, close to the already-placed centerline, to allow the algorithm to connect the new centerline to the existing vascular tree.

Internally, the module maintains a graph structure where nodes correspond to points along the centerlines and edges correspond to vessel segments. Parent–child relationships between branches are explicitly encoded, which preserves the hierarchical organization of the vascular tree. For each centerline point, a local radius is estimated and stored as an attribute of the graph. This information is later used by the segmentation module to adapt the size of the seeds to the vessel diameter estimation.

To facilitate long annotation sessions and external analyses, VSMOD can export and import the vascular graph as a JSON file using the NetworkX format. This functionality allows users to pause and resume annotation, or to process the graph with external tools while preserving the topology and geometric attributes of the vessels.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

II.2 Segmentation module

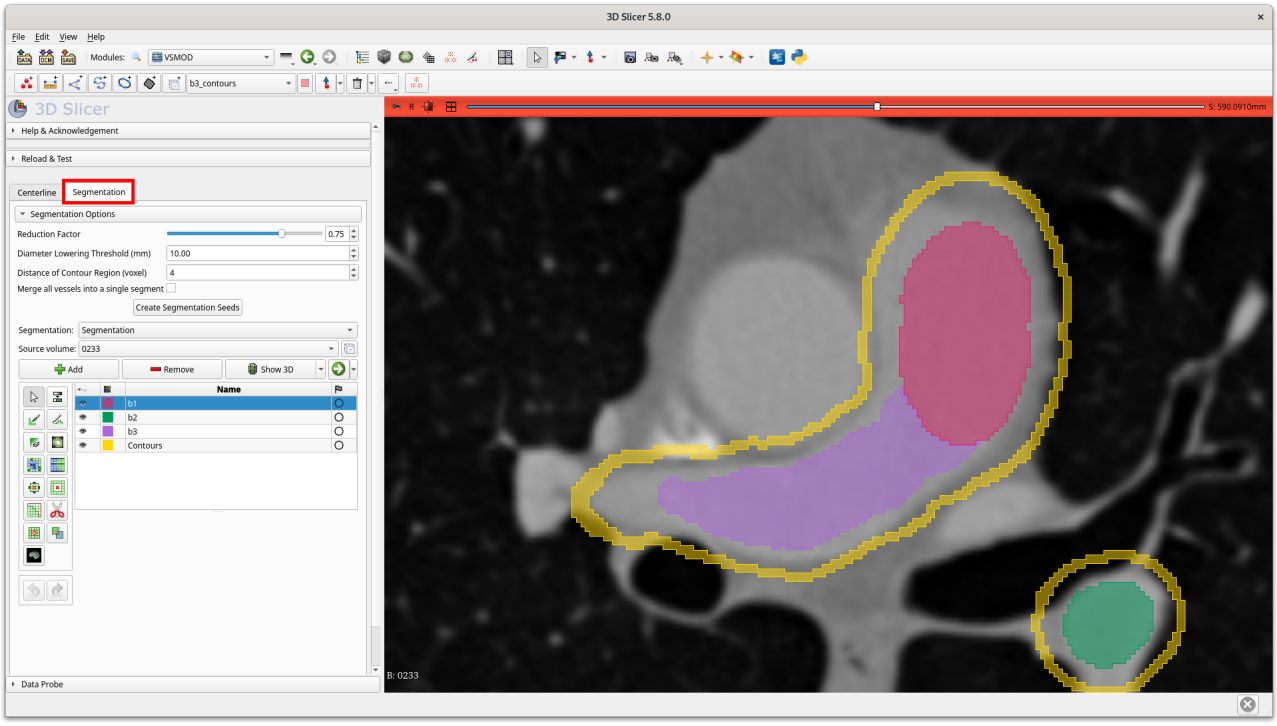


Fig. 4.9: Segmentation module

The second module transforms the centerline graph into a volumetric segmentation of the vascular network (See Figure. 4.9). It is built on top of the `Grow from Seeds` effect available in the 3D Slicer Segment Editor, which implements an efficient region-growing segmentation strategy. Instead of asking the user to manually draw seeds inside each vessel, VSMOD automatically generates seeds along the centerlines.

For each centerline point, a local seed is created with a size proportional to the estimated radius of the vessel at this location. Because we are dealing with patients who have pathologies, some radius estimation along the centerline can be faulty, and in our case, they were overestimated. So the *Reduction factor* parameter ensures that seeds remain inside the lumen while covering the vessel section, including branches with varying caliber. This parameter is by default set at 0.75, which means that a seed of 75% of the estimated diameter will be painted on the centerline point. It was extremely useful, especially for the vessel proximal to the heart, since their estimation of radius may be overestimated because of all the surrounding structures with similar contrast. To keep flexibility on this parameter, we added the *Diameter Lowering Threshold*, which means all vessels above this value will have the reduction threshold applied.

The background seeds are also automatically defined around the vessels to constrain the region-growing process and prevent leakage into surrounding tissues (see the yellow contours illustrated in Figure 4.9). To construct this outer border, first, a mask is created from all the vessel seeds. Then, two masks are created using morphological dilation with a ball-shaped structuring element: one with radius *Distance of Contour Region (voxel)* and another with radius *Distance*

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

of Contour Region + 2. This parameter sets the distance between the vessels and the external boundary. The contour border is obtained by subtracting the smaller dilated mask (distance) from the larger dilated mask (distance + 2), creating a hollow shell around the vessels. The user has the choice to keep all vessel labels unique or to merge all vessels into a single segment by choosing the option *Merge all vessels into a single segment*.

Once seeds are defined, the region-growing algorithm computes the multilabel segmentation, from which we obtain the vascular tree mask. As the segmentation is driven by the centerline graph, it inherits its topology: all branches present in the graph are represented in the volumetric mask, and the segmentation remains fully connected by construction. Users retain the possibility to edit seeds or the resulting segmentation manually to correct local errors, but the bulk of the annotation effort is automated.

III Evaluation

In this section, we evaluate VSMOD in terms of annotation efficiency and applicability to different vascular networks. The primary use case is the annotation of the PAT in CTPA images from our in-house PERSEVERE dataset, but we also test the plugin on brain vascular networks on a patient from the Brava dataset.

III.1 Reduction of annotation time

We first assessed the potential reduction in annotation time when using VSMOD compared to conventional tools available in 3D Slicer. Ten CTPA images from the first selected subset of patients from the PERSEVERE dataset were extensively annotated with VSMOD, aiming at a complete segmentation of the pulmonary arterial network. On average, each image contained 269 identifiable vessels.

The average annotation time per image using VSMOD was 3 hours and 35 minutes. Most of this time was dedicated to interactions in the centerline module (placement of starting and directional points), while a smaller fraction was spent refining seeds and correcting segmentations in complex regions, such as arteries with a thrombi.

In contrast, a full manual annotation of the same networks with standard tools (paintbrush, manual contouring) was deemed impractical. Instead, two experts annotated a subset of representative vessels using only the native tools of 3D Slicer. The average time to annotate a single vessel was 4 minutes and 46 seconds, which extrapolates to approximately 21 hours and 22 minutes per image for a complete network of 269 vessels.

Table 4.1 summarizes these results. VSMOD reduces the total annotation time by about 80% compared to purely manual annotation, while additionally providing an explicit centerline graph and radius estimates that are not directly available with conventional tools.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Annotation method	Time per vessel	Total time per patient
Manual tools (3D Slicer)	4m46s	21h22m52s
VSMOD	48s	3h35m12s

Tab. 4.1: Average segmentation times per patient for a pulmonary vascular network with 269 vessels on average.

III.2 Plugin usage on different vascular networks

To assess the generality of our approach, we also applied VSMOD to brain vascular networks of patient 001 of the Brava dataset (See Figure 4.10). Compared with pulmonary arteries, cerebral vessels exhibit distinct geometric characteristics, including higher tortuosity and smaller calibers (see Chapter 1 I.2). Despite these differences, the centerline and segmentation modules could be applied without too much modification of the algorithmic pipeline. The access to advanced RANSAC parameters allows us to change the *Maximum inter-cylinder angle(°)* parameter. In the case of the PAT, we fixed this parameter to 60° because PAT vessels are mostly straight, whereas for brain vessels, we increased it to 90° to allow greater flexibility to fit the more tortuous brain vessels.

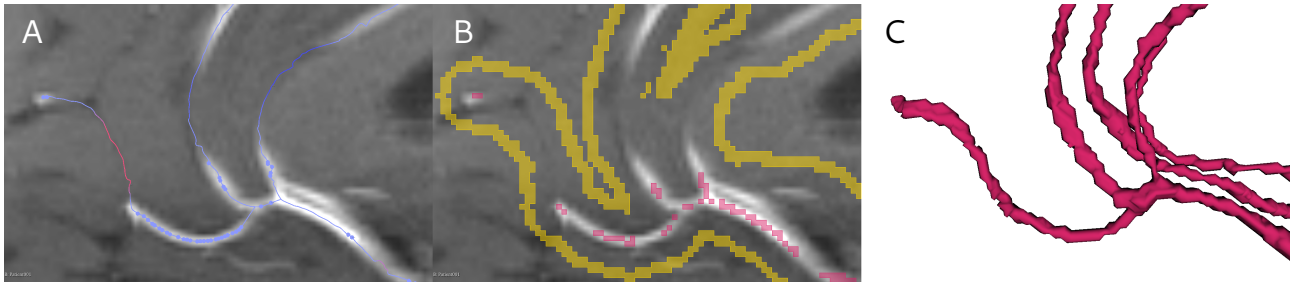


Fig. 4.10: Brain vessel annotation workflow (MRA). Left-to-right: (a) RANSAC centerlines (b) automatic region-growing seeds (c) final connected segmentation demonstrating cross-organ applicability.

Qualitative results showed that VSMOD extracted coherent centerline trees and corresponding volumetric segmentations of the brain vasculature. This experiment illustrates that the plugin is not restricted to a specific organ; it relies only on the assumption that vessels can be modeled as curved tubular structures. In practice, adapting VSMOD to a new vascular territory mainly requires adjusting advanced RANSAC parameters rather than redesigning the annotation workflow.

IV Conclusion

VSMOD provides a practical framework for generating accurate and topologically consistent vascular network annotations from volumetric medical images. By combining interactive centerline modeling with automatic region-growing segmentation, the plugin reduces annotation time by approximately 80% compared to conventional manual tools while preserving branch connec-

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

tivity and explicit graph structure. This makes VSMOD particularly well-suited to creating training and validation datasets for deep learning-based vascular segmentation.

The cross-organ experiment on brain tortuous vessels is a proof of concept demonstrating that the plugin is not limited to a specific organ. By adjusting a small set of parameters, VSMOD could be adapted to vascular networks with different geometric characteristics, such as higher tortuosity and smaller calibers, without redesigning the annotation workflow.

Several directions for future work can be envisioned. First, the RANSAC parameters and default settings could be adapted to different anatomical regions, which could be selected, providing organ-specific presets that further reduce the need for manual tuning. Second, integrating deep learning models into the segmentation module could improve robustness in challenging regions (e.g., low-contrast areas, pathology) and further reduce the need for manual seed correction.

The annotations produced by VSMOD on the PERSEVERE dataset constitute the ground truth on which the subsequent contributions of this thesis rely. Having established a reliable and efficient annotation pipeline, the next chapter 5 presents our deep learning-based approach, which leverages the annotated dataset produced by VSMOD to train models that enable us to segment the entire database.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

5 Training for vascular and thrombi segmentation

Table of contents

I Leveraging the PARSE public dataset . . . 72
I.1 Datasets and architectures . . . 72
I.2 Training configuration . . . 73
I.3 Results . . . 73
II PERSEVERE database segmentation . . . 77
II.1 Deep learning-assisted annotation strategy . . . 77
II.2 First iteration . . . 79
II.3 Second iteration . . . 82
II.4 Deployment on the full PERSEVERE database . . . 83
III Conclusion . . . 84

The previous chapter introduced the VSMOD plugin developed to facilitate vascular annotations. These annotations provide the ground truth required to train deep learning models. This chapter presents the training strategy adopted for pulmonary artery and thrombi segmentation.

During the annotation of the first sample of 10 patients from the PERSEVERE dataset with VSMOD, we tested different strategies to see whether we could leverage the PARSE public dataset. Finally, through an iterative process, we annotated, trained models, and corrected the segmentation, re-injecting it into the training models until we reached a quality sufficient for artery and thrombi mask segmentation across all 353 selected patients.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I Leveraging the PARSE public dataset

To develop a pulmonary artery segmentation model adapted to the PERSEVERE dataset, we investigated three complementary strategies: (i) supervised training on publicly available annotated data followed by transfer to PERSEVERE, (ii) semi-supervised training combining public annotated data with unannotated PERSEVERE images, and (iii) supervised training on a small set of manually annotated PERSEVERE patients. These strategies are evaluated through four models, described in the following subsections[1].

I.1 Datasets and architectures

The PARSE 2022 challenge (Y. Chu et al. 2025; Luo et al. 2024) provides a publicly available dataset of 100 CTPA scans with pulmonary artery segmentation annotations (see Chapter 2 I.2). We used 90 of these patients for model development and held out 10 as an independent test set, with the 90 patients further split into 72 for training and 18 for validation (80/20 split). In addition, all the unannotated PERSEVERE scans were used for semi-supervised training in Models 2 and 3. Finally, a first batch of 10 manually annotated PERSEVERE patients was used for Model 4, split into 6 for training, 2 for validation, and 2 for testing. The training data used for each model is summarized in Table 5.1.

To evaluate the benefit of incorporating unannotated PERSEVERE images through semi-supervised learning, we trained four models with varying levels of supervision and data sources. Model 1 serves as a fully supervised baseline, trained exclusively on the annotated PARSE data. Model 2 extends this baseline by incorporating all the unannotated PERSEVERE images through the Mean Teacher semi-supervised framework (see Chapter 3 I.1.4). Model 3 combines transfer learning and semi-supervised learning: it uses the weight of Model 1 as initialization for the Mean Teacher framework with both PARSE and unannotated PERSEVERE data. Finally, Model 4 departs from the PARSE-based strategy altogether, as it is a fully supervised TinyUNet trained exclusively on the sample of annotated PERSEVERE patients with VSMOD, without any PARSE data. The training data used for each model is summarized in Table 5.1.

Model	Architecture	PARSE (annotated)	PERSEVERE
Model 1	TinyUNet (supervised)	✓	
Model 2	Mean Teacher (semi-supervised)	✓	✓unannotated
Model 3	Mean Teacher (pretrained)	✓(pretraining + training)	✓unannotated
Model 4	TinyUNet (supervised)		✓annotated ($n = 10$)

Tab. 5.1: Training data and architecture used for each model.

[1]Preliminary experiments conducted on the CAS2023 challenge, detailed in Annex V, allowed us to establish and validate a complete training pipeline for vascular segmentation before its application to pulmonary arteries.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I.2 Training configuration

Models 1, 2, and 3 share common hyperparameters. Images were resampled to a spacing of 0.69x0.69x1.0 mm (mean voxel spacing between the median PARSE and PERSEVERE dataset) and processed with a patch size of $192 \times 192 \times 64$ voxels. The batch size was set to 2, with 250 batches per epoch, and training ran for up to 400 epochs with early stopping based on validation performance. The optimizer was Stochastic Gradient Descent (SGD) with Nesterov momentum (0.99) and weight decay of 3.0×10^{-5}, and the learning rate was fixed to 10^{-3} with a learning rate scheduler.

Model 4 had a different setting of hyperparameters. Images were resampled to the median spacing of the PERSEVERE cohort 0.72x0.72x0.5 mm. The batch size was set to 4, with 200 batches per training epoch and 50 per validation epoch, over 200 epochs. The learning rate was fixed to 10^{-3} and decayed linearly to 0 using a LinearLR scheduler, and the optimizer was Adam with no weight decay. Given the very limited training set size, extensive data augmentation was applied, including Gaussian noise ($p = 0.1$), Gaussian blur ($\sigma \in [0.5, 1.0]$, $p = 0.2$), brightness multiplication ($[0.75, 1.25]$, $p = 0.15$), contrast augmentation ($[0.75, 1.25]$, $p = 0.15$), low-resolution simulation (zoom $\in [0.5, 1.0]$, $p = 0.25$), gamma transform ($\gamma \in [0.7, 1.5]$, $p = 0.1$), and random mirroring along all axes.

I.3 Results

Quantitative results for all four models on both test sets are reported in Table 5.2. The PARSE test set is composed of 10 patients, while the PERSEVERE test set is composed of 2 patients. Qualitative results on representative patients from both datasets are presented in Figures 5.1 and 5.2.

Database	Metric	Model 1	Model 2	Model 3	Model 4
PARSE ($n = 10$ patients)	Dice (↑)	0.87 ± 0.03	0.67 ± 0.04	**0.87 ± 0.02**	0.26 ± 0.07
	clDice (↑)	0.86 ± 0.05	0.85 ± 0.04	**0.87 ± 0.04**	0.22 ± 0.04
	Precision (↑)	**0.88 ± 0.05**	0.83 ± 0.06	0.86 ± 0.04	0.24 ± 0.07
	Sensitivity (↑)	0.87 ± 0.06	0.56 ± 0.04	**0.89 ± 0.05**	0.29 ± 0.08
	β_0 (↓)	10.10 ± 4.63	15.20 ± 5.16	**8.30 ± 3.13**	530.00 ± 87.66
	β_1 (↓)	3.60 ± 2.72	6.40 ± 3.03	**3.00 ± 1.63**	226.00 ± 50.68
	β_2 (↓)	**1.00 ± 0.00**	1.90 ± 1.73	1.00 ± 0.00	17.80 ± 4.98
PERSEVERE ($n = 2$ patients)	Dice (↑)	0.75 ± 0.02	0.47 ± 0.00	0.77 ± 0.03	**0.80 ± 0.04**
	clDice (↑)	**0.73 ± 0.01**	0.66 ± 0.02	0.71 ± 0.02	0.69 ± 0.02
	Precision (↑)	0.76 ± 0.07	0.59 ± 0.11	0.72 ± 0.09	**0.84 ± 0.05**
	Sensitivity (↑)	0.80 ± 0.11	0.41 ± 0.06	**0.85 ± 0.05**	0.76 ± 0.03
	β_0 (↓)	**21.00 ± 1.41**	35.00 ± 1.41	24.50 ± 0.71	106.50 ± 3.54
	β_1 (↓)	15.50 ± 2.12	**10.50 ± 9.19**	11.50 ± 3.54	44.50 ± 14.85
	β_2 (↓)	**0.50 ± 0.71**	3.00 ± 2.83	2.50 ± 3.54	12.50 ± 0.71

Tab. 5.2: Performance comparison of all four models on the PARSE test set ($n = 10$ patients) and the PERSEVERE test set ($n = 2$ patients). Values are shown as mean ± std.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

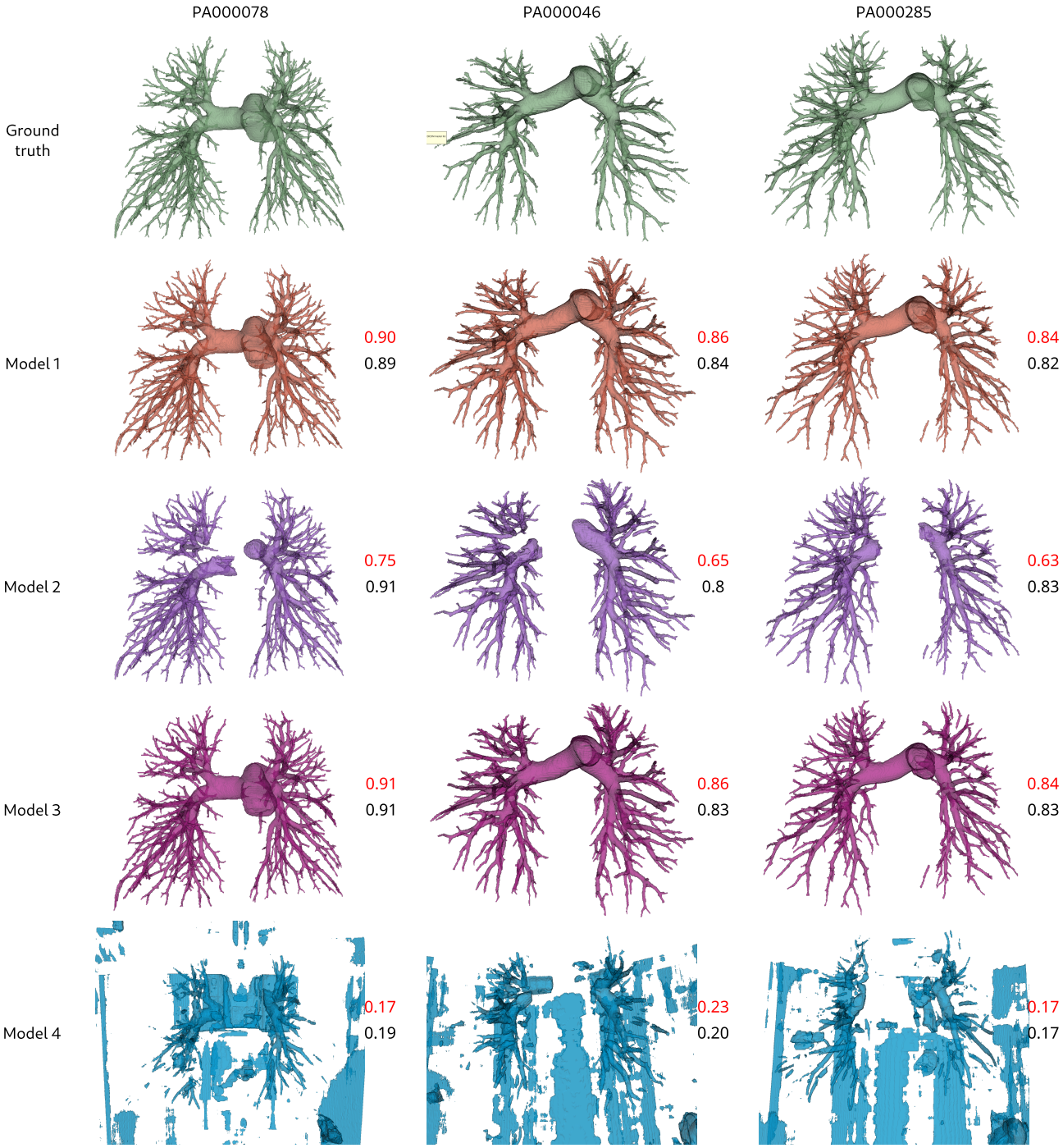


Fig. 5.1: Qualitative result of the 4 models on PARSE PA000078, PA000046, PA000285 patients. Dice values in red and clDice values in black.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

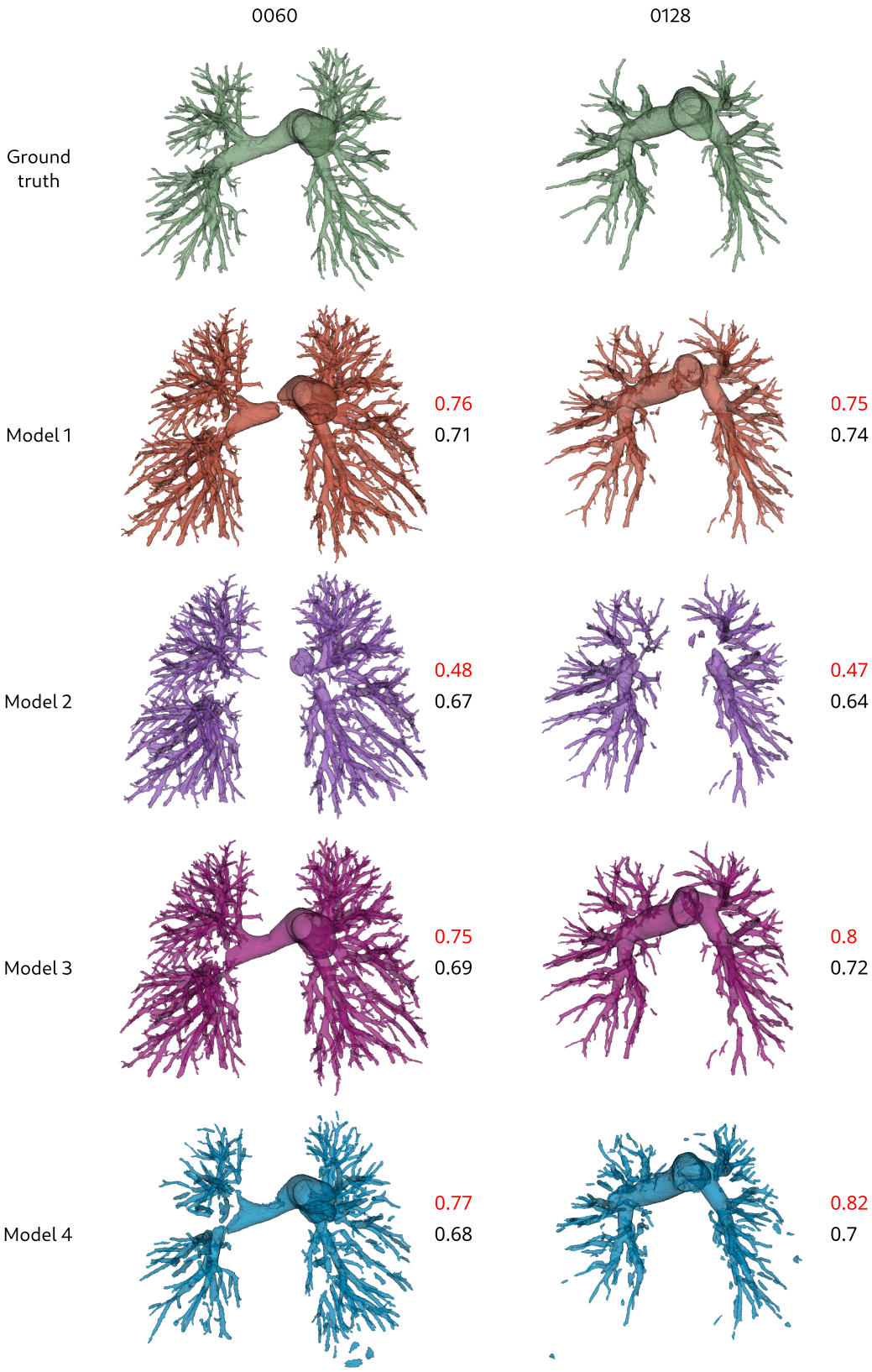


Fig. 5.2: Qualitative result of the 4 models on PERSEVERE 0060 and 0128 patients. Dice values in red and clDice values in black.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Results on the PARSE test set. On the PARSE test set, Models 1 and 3 achieve the best overlap performance, with identical Dice scores of 0.87, while Model 3 further improves topological quality, as reflected by lower β_0 (8.30 vs. 10.10) and β_1 (3.00 vs. 3.60) values. Model 2 exhibits a marked degradation across all metrics, with a Dice coefficient dropping to 0.67 and a sensitivity of only 0.55. The degradation of Model 2 is visually attributable to the partial or complete loss of the MPA in its predictions. An explanation is that the unannotated PERSEVERE images used contain many patients with pulmonary embolism, in whom thrombi partially or fully occlude the MPA, which substantially alter its appearance relative to the healthy PARSE population. The teacher network, lacking supervision over PERSEVERE images, likely produces unreliable pseudo-labels that degrade student training. Model 3 permits answering this issue by initializing the student from the supervised baseline before applying consistency regularization, thereby preserving MPA detection while still benefiting from the additional unlabeled data. Model 4 fails to generalize to PARSE, with a Dice coefficient of 0.26 and topological error counts several orders of magnitude above the other models ($\beta_0 = 530$, $\beta_1 = 226$, $\beta_2 = 17.80$). This is due to the Model 4 segmenting many structures in the images that have intensity profiles similar to those of the arteries. Because those structures are outside of the lungs, a post-processing could be applied to isolate the vessel segmentation. The failure of Model 4 on PARSE reflects the strong domain shift between the two datasets: having been trained exclusively on PERSEVERE scans from patients with pulmonary embolism, the model has adapted to their specific intensity distribution and vascular appearance, and does not generalize well to the healthy PARSE population.

Results on the PERSEVERE test set. On the PERSEVERE test set, Model 4 achieves the highest Dice coefficient (0.80) and the best precision (0.84). Models 1 and 3 reach comparable Dice values of 0.75 and 0.77, respectively, while Model 2 shows a substantial drop to 0.47, with a particularly low sensitivity of 0.41. Among Models 1 to 3, the relatively close Dice scores, with only a marginal difference of 0.02 points between Model 1 and Model 3 on PERSEVERE, suggest that the semi-supervised training did not lead to a substantial improvement in segmentation quality. The performance recovery observed in Model 3 compared to Model 2 is more likely attributable to the pre-training initialization inherited from Model 1 than to the Mean Teacher framework itself. The poor performance of Model 2 on PERSEVERE, with a sensitivity of 0.41, is consistent with the MPA detection issues observed on PARSE: the consistency regularization without a supervised initialization appears detrimental in both domains, likely because the teacher network produces unreliable pseudo-labels in regions affected by PE. The topological metrics reveal a clear separation between Model 4 and the other three models: Model 4 exhibits a β_0 value of 106.50, compared to values ranging from 21.00 to 35.00 for Models 1 to 3, and similarly elevated β_1 (44.50) and β_2 (12.50) values. Despite achieving the highest overlap score, Model 4 produces anatomically inconsistent segmentations, as shown by its topological metrics. The very limited training set of 6 patients appears insufficient for the model to learn the global topology of the pulmonary arterial tree, leading to a large number of spurious disconnected components, cycles, and enclosed cavities.

These findings motivated the design of an iterative deep learning-assisted annotation pipeline on the PERSEVERE cohort, described in the following section.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

II PERSEVERE database segmentation

The experiments conducted on the PARSE dataset demonstrated that models trained on healthy subjects do not transfer well to the pathological PERSEVERE cohort, and that a small annotated in-domain dataset is insufficient to produce topologically consistent segmentations. Obtaining a larger set of PERSEVERE-specific annotations is therefore necessary for training a reliable segmentation model. To this end, we designed a deep learning-assisted iterative annotation strategy, described hereafter.

II.1 Deep learning-assisted annotation strategy

Manually annotating even just the sample of 10 patients from the PERSEVERE cohort is verry time-consuming due to the complexity of pulmonary structure, and is not possible to apply to the full cohort in a limited timeline. To address this constraint, we adopted an active-learning-inspired strategy, illustrated in Figure 5.3, whose objective is to progressively build a large set of reliable annotations by alternating between model training and expert-guided correction. The protocol proceeds as follows. First, a small initial subset of PERSEVERE patients was semi-automatically or manually annotated, and was verified by an expert pulmonologist, providing a carefully reviewed set of ground-truth segmentations. Segmentation models are then trained on this annotated subset and applied to a broader set of unannotated patients to generate automatic predictions. These predictions are visually reviewed by the expert, and a selection of patients is chosen for manual correction based on image quality criteria. The corrected segmentations are added to the annotated pool, and the process is repeated for a second iteration. Two iterations were carried out in total.

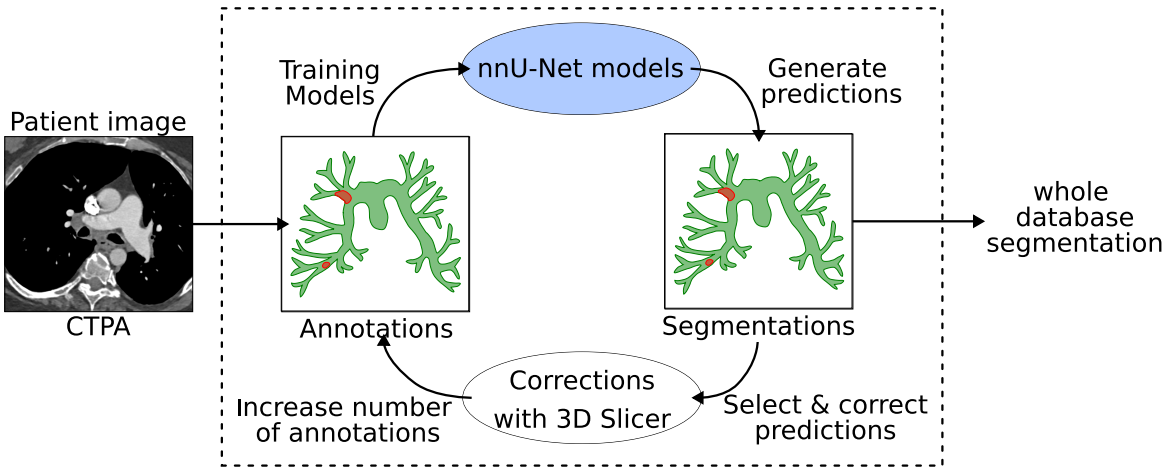


Fig. 5.3: Active-learning-inspired iterative annotation protocol used to obtain reliable pulmonary artery and thrombi segmentations across the PERSEEVERE database.

II.1.1 Annotation targets and tools

Two anatomical structures were targeted for segmentation throughout this process: the PAT and pulmonary emboli. Lung masks were also produced as additional inputs to provide more anatomical context.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Pulmonary artery annotations were produced using VSMOD, the semi-automatic 3D Slicer plugin developed in-house that combines vessel tracking and manual editing (see Chapter 4). Thrombi annotations were performed entirely manually in 3D Slicer and checked by an expert pulmonologist, who validated each annotation. Lung masks were generated automatically using `TotalSegmentator` (see Chapter 3 I.2.2).

II.1.2 Choice of segmentation framework

The preliminary experiments on the PARSE dataset relied on custom architectures such as TinyUNet and Mean Teacher. For the PERSEVERE-specific training, we switched to the nnU-Net framework (see Chapter 3 I.1.6), which automatically adapts its architecture, preprocessing, and training configuration to the input dataset. This choice was motivated by the fact that nnU-Net has demonstrated state-of-the-art performance on a wide range of medical image segmentation tasks even with a small dataset.

For each iteration, four nnU-Net models were trained, each receiving a different combination of anatomical masks as input, as illustrated in Figure 5.4:

- **Model 1**: Trained on artery masks only.
- **Model 2**: Trained on thrombi masks only.
- **Model 3**: Trained on combined artery and thrombi masks.
- **Model 4**: Trained on combined artery, thrombi, and lung masks.

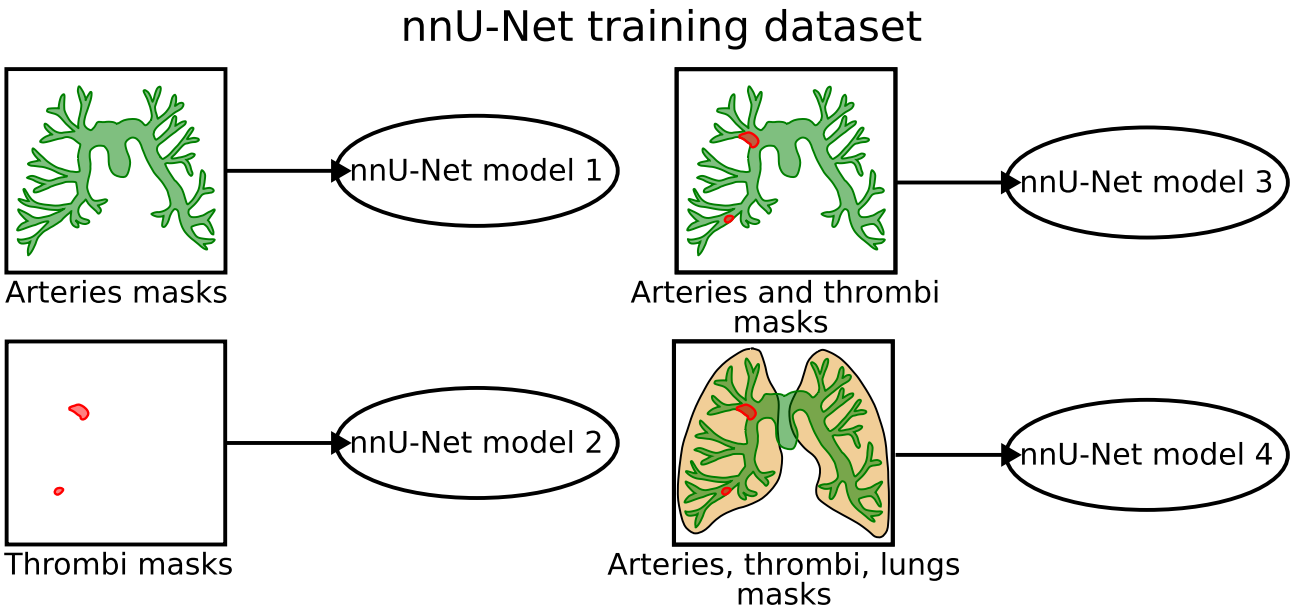


Fig. 5.4: Illustration of the four input configurations used for nnU-Net training across both iterations. Models differ in the anatomical context provided as segmentation targets.

This multi-context strategy was designed to assess whether providing additional anatomical information improves segmentation accuracy. For PE in particular, which are small and low-contrast structures, we hypothesized that providing the arterial tree and lung masks as contex-

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

tual information would help the model localize thrombi more reliably by constraining predictions to anatomically plausible regions.

For each model, nnU-Net automatically determined up to 4 configurations based on the training data characteristics: a 2D configuration, a 3D low-resolution configuration, a 3D full-resolution configuration, and a 3D cascade full-resolution configuration. The pairwise ensembles of these are also considered. In our case, only the first 3 configurations were selected based on the PERSEVERE dataset properties, as presented in Figure 5.5. The best configuration for each model was selected based on the mean foreground Dice score on the internal validation folds, which corresponds to the average Dice across all non-background classes.

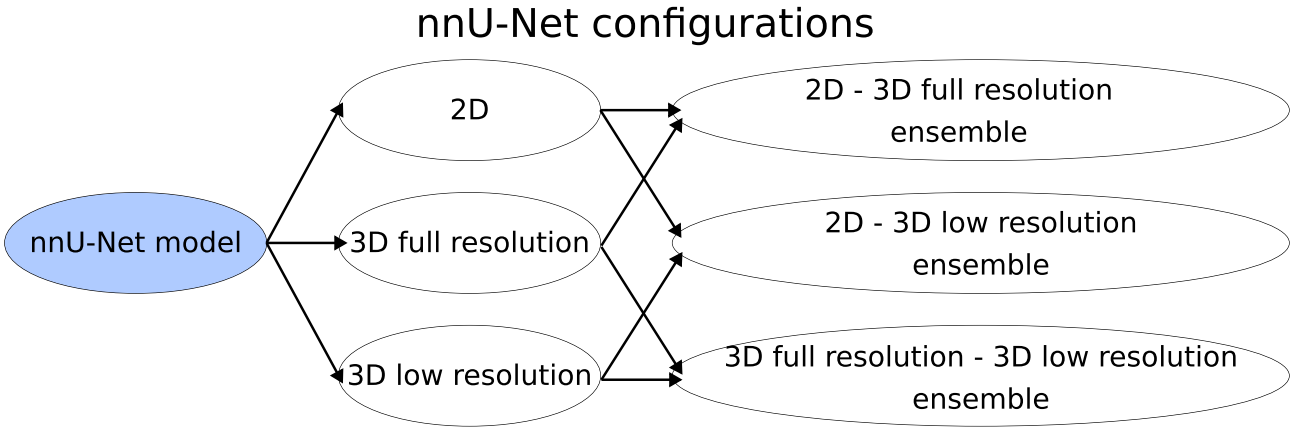


Fig. 5.5: The 3 nnU-Net configurations automatically generated and their pairwise ensemble evaluated for each model on our PERSEVERE dataset.

II.2 First iteration

II.2.1 Training and configuration selection

The first iteration relied on the 10 PERSEVERE patients' annotations: arterial trees were annotated using VSMOD, and pulmonary emboli were manually delineated in 3D Slicer, and an expert pulmonologist validated each contour. Among the 10 patients, 8 were used for training with 5-fold cross-validation and 2 were held out as an independent test set.

The mean foreground Dice scores obtained by each nnU-Net configuration on the internal validation folds are reported in Table 5.3. The 3D full-resolution configuration consistently achieved the highest validation Dice scores across all four models and was therefore selected for inference.

II.2.2 Test set results

The performance of each model on the two test patients is reported in Table 5.4. For artery segmentation, Model 1, trained on arterial masks only, achieves the highest Dice score (0.91 ± 0.05). The differences between Models 1, 3, and 4 indicate that providing additional anatomical context (thrombi and lung masks) does not improve arterial segmentation and may even slightly hinder performance.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

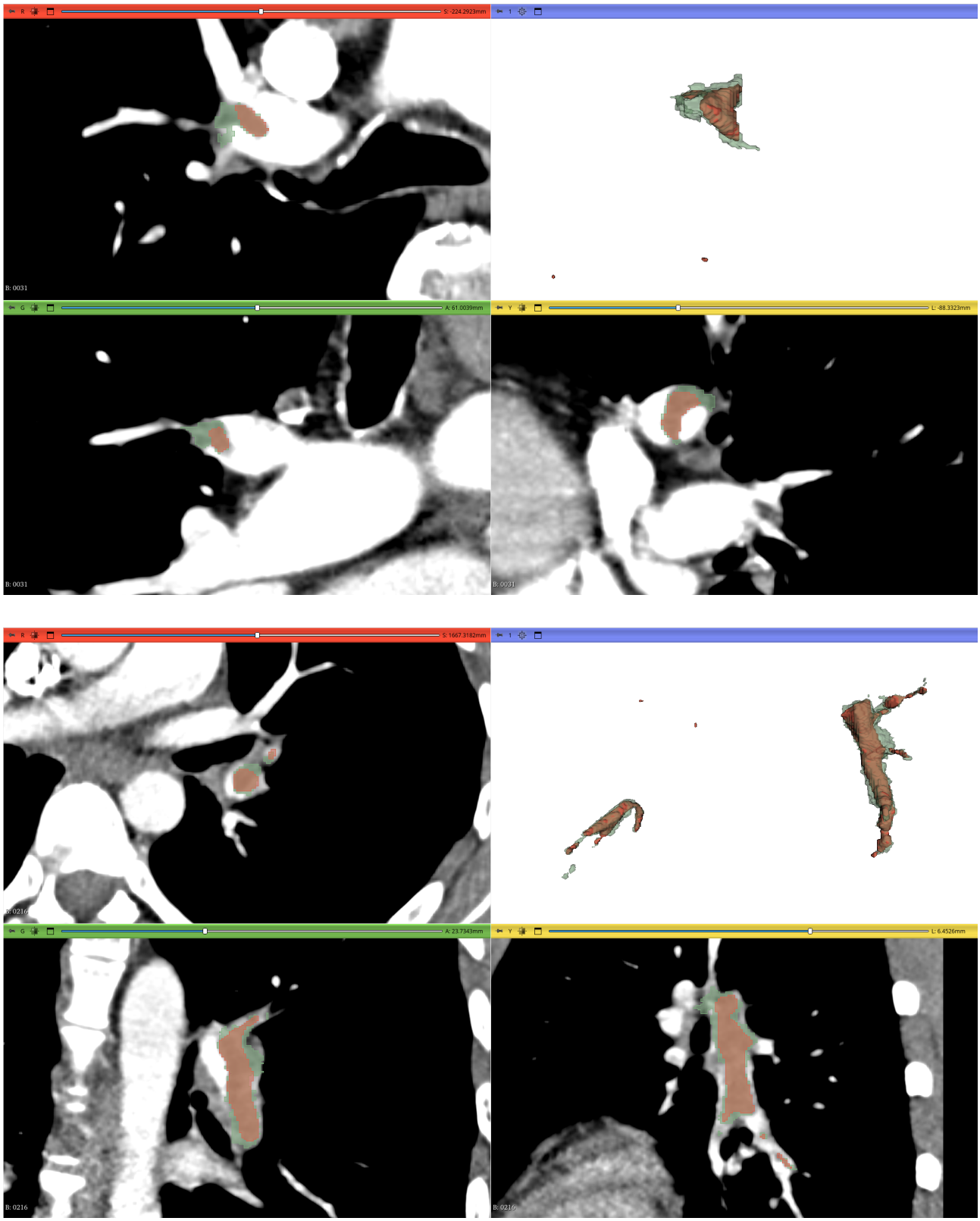


Fig. 5.6: Thrombi segmentation results on the two first-iteration PERSEVERE test patients: patient 0031 (top) and patient 0216 (bottom). Predictions are shown in red and the ground truth in green.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Model	2D	3D lowres	3D fullres	ensemble 2D 3D lowres	ensemble 2D 3D fullres	ensemble 3D lowres 3D fullres
1	0.84 ± 0.06	0.88 ± 0.05	**0.90 ± 0.02**	0.88 ± 0.04	0.89 ± 0.04	0.90 ± 0.03
2	0.34 ± 0.15	0.40 ± 0.14	**0.65 ± 0.10**	0.30 ± 0.14	0.39 ± 0.17	0.50 ± 0.12
3	0.62 ± 0.08	0.71 ± 0.07	**0.79 ± 0.04**	0.69 ± 0.07	0.74 ± 0.06	0.77 ± 0.05
4	0.75 ± 0.05	0.80 ± 0.04	**0.85 ± 0.02**	0.79 ± 0.05	0.82 ± 0.03	0.84 ± 0.03

Tab. 5.3: Mean foreground Dice scores of each nnU-Net configuration and ensemble for the first iteration (validation results). Bold values indicate the best configuration for each model.

Class	Model	Mean Dice ± Std
Arteries	**Model 1**	**0.91 ± 0.05**
	Model 3	0.89 ± 0.06
	Model 4	0.89 ± 0.06
Embolies	Model 2	0.60 ± 0.01
	Model 3	0.64 ± 0.10
	Model 4	**0.67 ± 0.13**

Tab. 5.4: Performance of artery and thrombi segmentation models: mean Dice score on test datasets (first iteration). Bold values indicate the best-performing model per class.

For thrombi segmentation, the results follow the opposite trend: Model 4, which receives artery and lung masks in addition to thrombi annotations, achieves the best performance (0.668 ± 0.126), outperforming Model 2, which relies on thrombi masks alone (0.595 ± 0.010). This improvement is consistent with our initial hypothesis: pulmonary emboli are small, low-contrast structures whose localization is tied to the arterial lumen. Providing the arterial tree and lung extent as contextual information helped the model. Qualitative results for thrombi segmentation are shown in Figure 5.6. The two patients represent contrasting embolism profiles: patient 0031 presents a predominantly central, elongated thrombus confined to the main pulmonary branches, whereas patient 0216 presents a more peripheral and distributed embolic burden. In both cases, the predicted contours closely follow the ground truth, although a larger discrepancy is observed for patient 0216, where peripheral emboli are partially missed.

II.2.3 Manual correction and patient selection

Following the first iteration inference, 20 additional patients were selected from the broader PERSEVERE cohort according to the image quality criteria defined in Chapter 2 I.1.4. The arterial segmentations produced by Model 1 were post-processed by retaining only the largest connected component, ensuring a single connected arterial tree per patient. These predictions were then reviewed, corrected when needed using VSMOD and validated by an expert pulmonologist.

In parallel, the thrombi predictions produced by Model 4 were reviewed by the expert pulmonologist, who indicated the correct location of thrombi visible on the images. However, 6 of

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

the 20 patients were ultimately excluded from thrombi annotation for the following reasons: either artifact prevented the expert from confidently distinguishing true filling defects from artefactual hypodensities, or the expert was uncertain about the diagnosis of pulmonary embolism and considered the scan too ambiguous for reliable annotation. In such cases, the risk of introducing erroneous annotations was judged to outweigh the benefit of adding these patients to the thrombi training set. The remaining 14 patients received both corrected arterial and thrombi annotations, while the 6 excluded patients contributed only as corrected arterial mask to the annotated pool. Thrombi annotations were also post-processed to ensure anatomical consistency: voxels located outside the arterial mask were removed, ensuring that all annotated emboli remained within the segmented arterial tree.

At the end of the first iteration, the annotated pool comprised 30 patients with corrected arterial segmentations and 24 patients with both arterial and thrombi annotations.

II.3 Second iteration

II.3.1 Training and configuration selection

The same four nnU-Net models were retrained following the same protocol: 5-fold cross-validation on 19 patients with 5 patients reserved for testing (Models 1, 2, 3, and 4). In addition, a variant of Model 1, referred to as Model 1bis, was trained using all 30 available arterial annotations (24 for training and 6 held as test set).

The configuration selection results are presented in Table 5.5. Compared to the first iteration, the ensemble of 3D low-resolution and 3D full-resolution configurations achieved the best validation performance for artery-only models (Models 1 and 1bis), while the 3D full-resolution configuration remained the best choice for models incorporating additional anatomical context. We used the ensemble 3D low-resolution and 3D full-resolution configuration for model 1 and 1bis, and the 3D full-resolution configuration for the other models.

Model	2D	3D lowres	3D fullres	ensemble 2D 3D lowres	ensemble 2D 3D fullres	ensemble 3D lowres 3D fullres
1	0.89 ± 0.05	0.93 ± 0.03	**0.94 ± 0.03**	0.92 ± 0.03	0.92 ± 0.03	**0.94 ± 0.03**
1bis	0.91 ± 0.03	0.94 ± 0.02	**0.95 ± 0.03**	0.93 ± 0.02	0.93 ± 0.02	**0.95 ± 0.02**
2	0.45 ± 0.22	0.46 ± 0.24	**0.61 ± 0.24**	0.39 ± 0.25	0.46 ± 0.24	0.53 ± 0.25
3	0.69 ± 0.12	0.72 ± 0.13	**0.80 ± 0.12**	0.72 ± 0.12	0.76 ± 0.13	0.77 ± 0.14
4	0.79 ± 0.08	0.82 ± 0.09	**0.85 ± 0.09**	0.82 ± 0.08	0.84 ± 0.08	0.84 ± 0.09

Tab. 5.5: Mean foreground Dice scores of each nnU-Net configuration and ensemble for the second iteration (validation results). Bold values indicate the best configuration for each model.

II.3.2 Test set results

The performance of the models on the test sets is reported in Table 5.6. Across all models and both segmentation tasks, the second iteration yields substantially higher Dice scores compared

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

to the first iteration, confirming that expanding the annotated dataset through the iterative correction process improved model performance.

Class	Model	Mean Dice ± Std
Arteries	**Model 1**	**0.94 ± 0.04**
	Model 1bis	**0.94 ± 0.04**
	Model 3	0.93 ± 0.051
	Model 4	0.94 ± 0.05
Embolies	Model 2	0.78 ± 0.14
	Model 3	0.81 ± 0.13
	Model 4	**0.82 ± 0.12**

Tab. 5.6: Performance of artery and thrombi segmentation models: mean Dice score on test datasets (second iteration). Bold values indicate the best-performing model per class.

For artery segmentation, Model 1 and Model 1bis achieve a mean Dice of 0.94 ± 0.04. Although there are no differences, Model 1bis was selected for deployment on the full cohort, as it benefits from the largest available set of arterial annotations.

For thrombi segmentation, Model 4 again achieves the best performance (0.82 ± 0.12), consistent with the first iteration findings. Model 4 was therefore selected for thrombi segmentation on the full cohort.

The conclusions from the first iteration are confirmed: adding thrombi and lung context does not improve arterial segmentation, but it does for thrombi segmentation.

II.4 Deployment on the full PERSEVERE database

Following the iterative annotation process, Model 1bis and Model 4 were applied to all 353 selected PERSEVERE patients to produce pulmonary artery and thrombi segmentation masks, respectively.

Lung and lobes masks were generated automatically for all patients using `TotalSegmentator` (see more in Annex B.1).

All segmentation masks were post-processed as during the iterative process, to enforce anatomical and topological consistency prior to graph construction and analysis. For pulmonary arteries, only the largest connected component was retained, ensuring a single connected arterial tree per patient ($\beta_0 = 1$). For pulmonary emboli, voxels located outside the arterial mask were removed, ensuring that all predicted thrombi remained strictly within the arterial lumen. The descriptive statistic of the whole cohort can be found in Annex III.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

III Conclusion

In this chapter, we presented the training strategy used to obtain reliable pulmonary artery and thrombi segmentations on the PERSEVERE dataset. Preliminary experiments evaluating transfer learning strategies using the PARSE dataset revealed a significant domain gap between healthy subjects and PE cases, motivating the need for a dataset-specific approach.

To address this limitation, we introduced a deep learning-assisted iterative annotation strategy tailored to the PERSEVERE cohort. Through successive rounds of model training, prediction, and manual correction, we progressively increased both the number and quality of annotations. This process ultimately enabled the training of robust nnU-Net models for both pulmonary artery and thrombi segmentation, achieving Dice scores of 0.94 and 0.82 , respectively. The final segmentations were applied to all 353 selected patients, and post-processing was performed to ensure anatomical consistency of the PAT. These segmentations form the foundation for the biomarker extraction pipeline described in the following chapter 6.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

6 Graph modeling and Biomarker extraction

Table of contents

I Methodology . . . 86
I.1 PAT graph modeling . . . 86
I.2 Hierarchical labeling . . . 91
I.3 PE feature extraction . . . 95
I.4 Patient level feature extraction . . . 97
II Results . . . 99
II.1 From voxel-graph to branch graph . . . 99
II.2 Graph error correction and comparison with VesselGraph pipeline . . . 101
II.3 Graph-based representation of the PERSEVERE cohort . . . 103
II.4 Hierarchical annotation of the PAT graph . . . 103
II.5 Thrombotic burden characterization through graph-derived biomarkers . . . 105
II.6 Computational time of the pipeline . . . 106
III Conclusion . . . 108

PE is a life-threatening condition whose severity is highly dependent on the location, extent, and distribution of thrombi within the PAT. As discussed in Chapter 1, current clinical severity scores rely on biomarkers that can sometimes be missing due to emergency settings. Thrombotic burden scores, such as Qanadli and Mastora, despite relying on simplified anatomical templates, are too time-consuming for manual assessment, limiting their reproducibility and clinical use. And because of their simplified anatomical templates, these scores fail to account for the patient-specific vascular architecture. The tree-shaped, acyclic, and hierarchically labeled structure makes the PAT particularly well suited to a graph-based representation, where each vessel segment corresponds to an edge, and each bifurcation or end of a vessel a node. This key property motivates the graph modeling approach developed in this Chapter. The goal of this chapter is to describe an automatic pipeline that models the PAT as a directed graph from the CTPA-derived segmentations produced in the previous Chapter 5 and extracts both local and global biomarkers relevant to PE severity assessment.

The content of this chapter has been submitted to the Journal of Biomedical and Health Informatics (JBHI) and is currently under review.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I Methodology

The pipeline goal is to model the PAT as a directed graph, enabling the extraction of novel imaging-based biomarkers. These biomarkers can be broadly divided into:

- *Local biomarkers*, which describe vascular and thrombus characteristics at the artery (edge) level (e.g. length, diameter, hierarchical position, clot volume, obstruction ratio);
- *Patient-level biomarkers*, which aggregate local information into global indicators such as the Total Embolic Volume (TEV) repartition by lobes and hierarchical levels, or automated established thrombotic burden scores (e.g., Qanadli, Mastora).

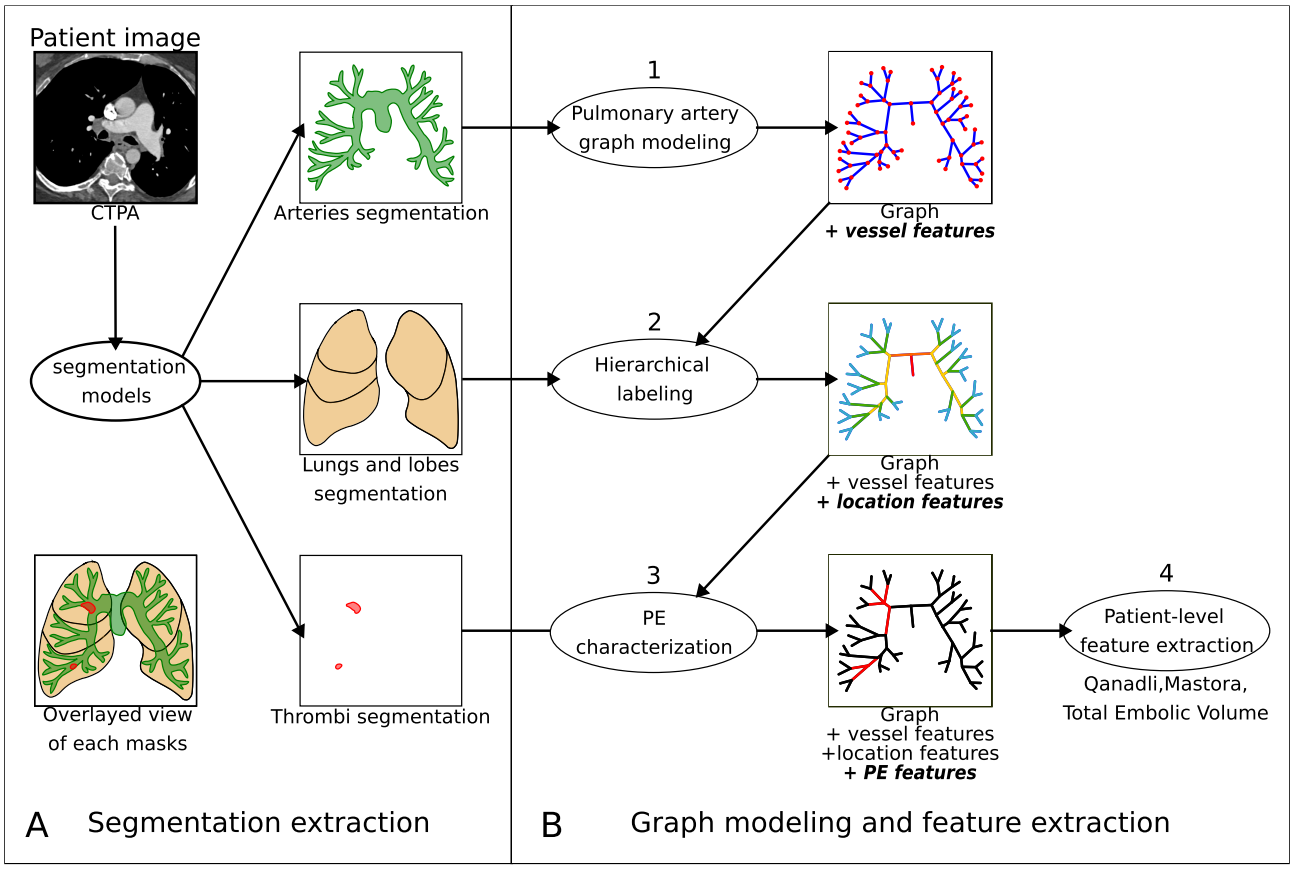


Fig. 6.1: General workflow of the proposed pipeline.
A) Segmentation of arteries, lungs, lobes and thrombi described in Chapter 5. B) Graph modeling and feature extraction pipeline. Steps (1-3) generate the graph and produce local biomarkers, which are subsequently aggregated in step (4) to produce patient-level biomarkers.

An overview of the pipeline is presented in Figure 6.1. Each step of the graph modeling and feature extraction part (B) will be described in detail in the following.

I.1 PAT graph modeling

The objective of the pulmonary artery graph modeling step (1) is to convert the binary arterial segmentation into a directed, anatomically consistent graph of the PAT enriched with geometric

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

attributes. The graph modeling process is composed of four main steps: preprocessing, graph construction, artifact correction, and graph orientation.

I.1.1 Preprocessing

The preprocessing step converts the arterial binary mask into a skeleton representation (see 3 II.1) and an associated distance map (See Figure. 6.2).

Skeletonization allows obtaining a 1-voxel-thick centerline that preserves the topology of the PAT. Skeletonization is often sensitive to noise and surface irregularities, producing spurious terminal branches. Loops may also arise from contact between different vessels in the segmentation and can create non-physiological cycles. These morphological and topological errors motivate a dedicated graph correction step after the initial graph construction (see I.1.3).

The distance map is obtained using an Euclidean distance transform on the arterial mask so that each voxel encodes the shortest distance to the vessel boundary in millimeters, using the CTPA voxel spacing. This distance map will be later used to estimate the local vessel radius along each centerline coordinate.

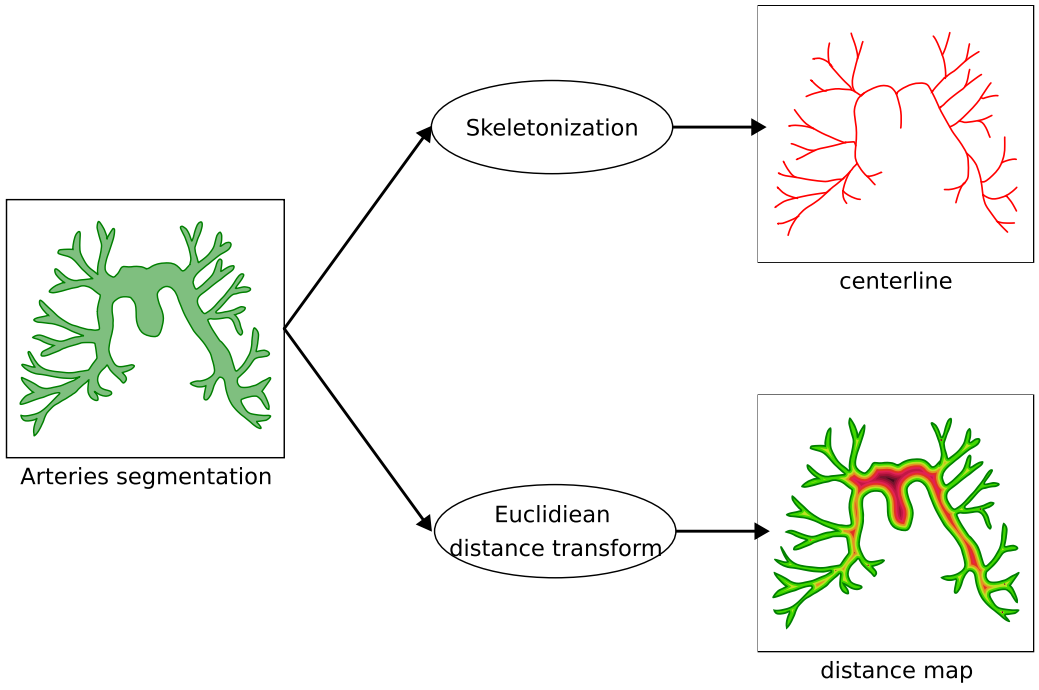


Fig. 6.2: Preprocessing of the artery mask, to extract the centerline and the distance map.

I.1.2 Graph construction

Graph construction transforms the skeleton into a branch-level graph, where edges represent individual arterial vessels. The illustration of each step is presented in Figure 6.3.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

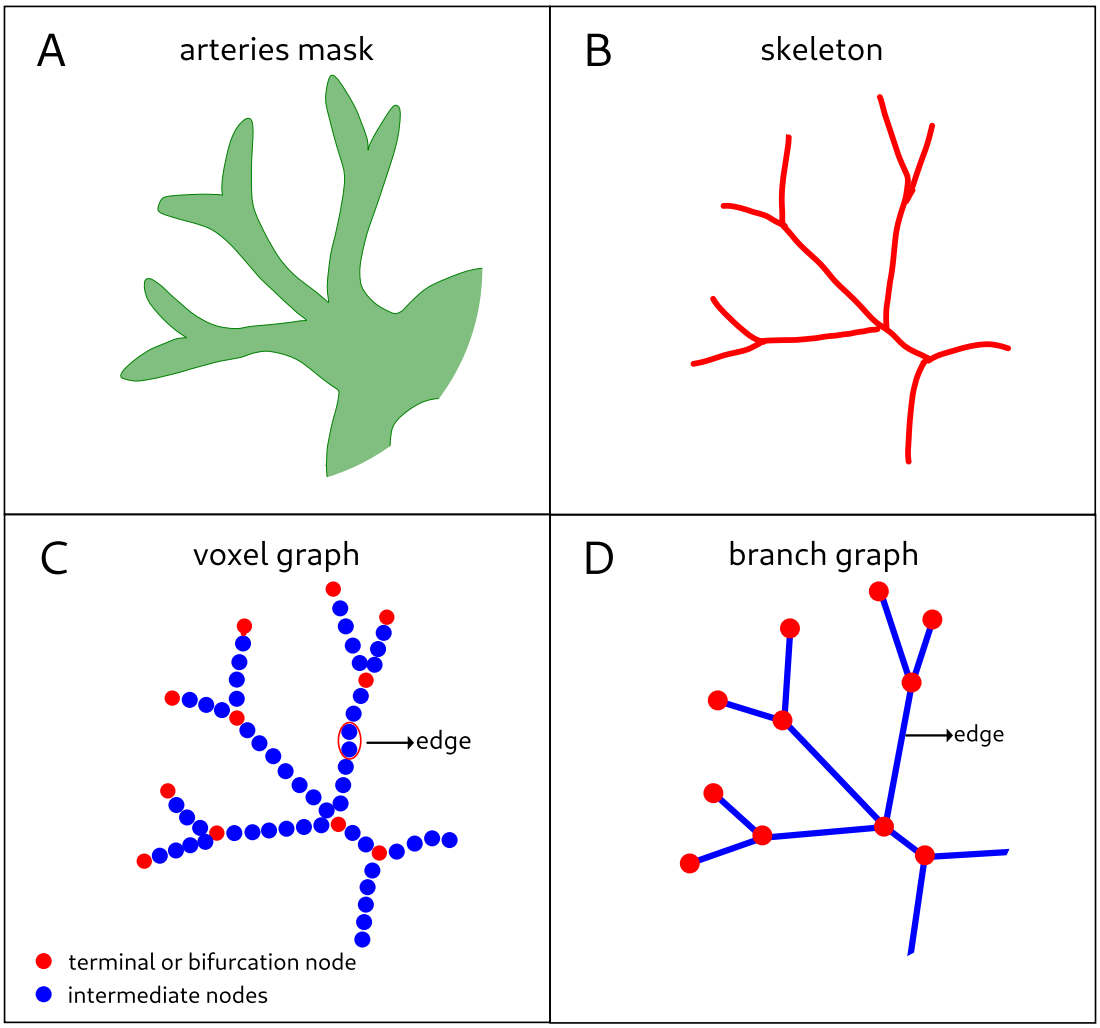


Fig. 6.3: Graph construction steps.The (A) arteries mask is skeletonized, from which we obtain the centerline of the arteries (B). The centerlines are transformed into a voxel-level graph (C) that is then converted into the branch-level graph (D).

In the previous step from the artery mask (See Figure 6.3 A) we obtained 3D centerline (See Figure. 6.3 B) which is converted into a voxel-level graph (See Figure. 6.3 C) using the `skancsr` library (Nunez-Iglesias et al. 2018) where each skeleton voxel corresponds to a node connected to its immediate spatial neighbors (3D adjacency, 26-connectivity). Each node is assigned its 3D coordinates and a local radius extracted from the distance map. Nodes are then classified according to their degree into terminal nodes (degree 1), bifurcation nodes (degree ≥ 3), and intermediate nodes (degree 2). Chains of intermediate nodes connecting two terminal or bifurcation nodes are merged into single edges, thus forming a branch-level graph (See Figure 6.3 D). Each edge of the graph now encodes the **vessels features**:

- the centerline coordinates.
- the radius profile along the centerline (and mean, min, and max radius values).
- the segment length (in mm).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

The radius profile is deduced from the distance map at each voxel position in the centerline. At this stage, the graph is undirected and still contains artifacts induced by skeletonization or non-topological cycle.

I.1.3 Graph correction

Graph correction aims to remove non-physiological cycles and the artefact from the skeletonization process, while preserving anatomically relevant arteries (See Figure. 6.4). The two errors in the graph are :

- (i) spurious terminal branches due to surface irregularities.
- (ii) cycles caused by contact between adjacent arteries in the segmentation mask.

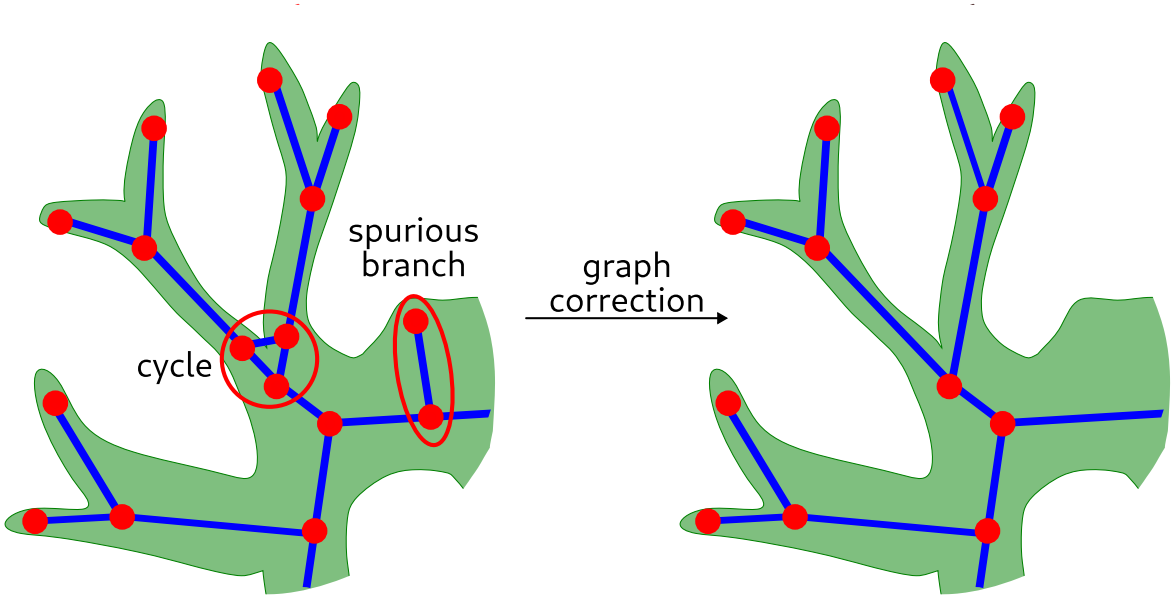


Fig. 6.4: Illustration of the cleaning process, detecting a spurious vessel caused by an irregularity on the surface of the mask, and a cycle caused by contact between two close arteries.

We will now explain our approach for correcting the two types of errors. Then we will compare our performance of correction against the Voreen framework (Meyer-Spradow et al. 2009), specifically its `VesselGraph` pipeline (Drees et al. 2021) (see Chapter 3 II.2).

Spurious branches :

Spurious terminal branches are expected to be short because they are provoked by the irregularities at the surface, so the majority of the branch is expected to be contained inside the parent artery, and the diameter at the end, the distal diameter, is expected to be small compared to the parent artery (See Figure 6.5). Let L be the branch length, d_{parent} the parent diameter, and d_{end} the terminal diameter. A length criterion :

$$L < \alpha\, d_{\text{parent}}, \qquad \alpha = 1.5, \tag{6.1}$$

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

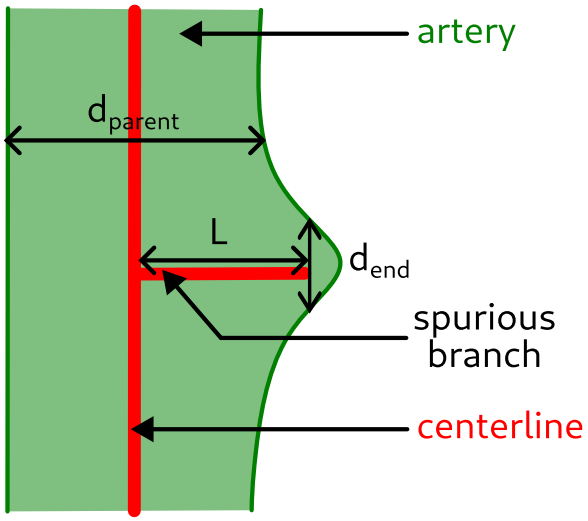


Fig. 6.5: Illustration of an artifact or a spurious vessel caused by an irregularity on the surface of the mask.

ensures that the terminal branch is mostly contained within the parent region, and a distal ratio :

$$r_{\text{end}} = \frac{d_{\text{end}}}{d_{\text{parent}}} \tag{6.2}$$

flags branches as spurious when $r_{\text{end}} < \beta$ with $\beta = 0.3$, corresponding to more than 70% diameter reduction.

Cycle :

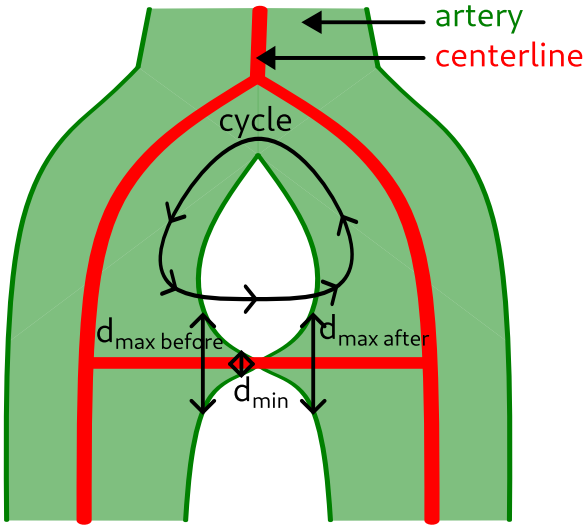


Fig. 6.6: Illustration of a cycle caused by the contact of two vessels by a few voxels, and the resulting constriction pattern.

For each edge participating in a cycle, the diameter profile along its centerline is analyzed to detect a constriction pattern. The assumption behind the constriction pattern is that a majority

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

of the cycles we encountered were caused by a contact of a few voxels between two close vessels (See Figure 6.6). Let $d_{\min}$ denote the minimum diameter along the edge, and $d_{\max}^{\text{before}}$ and $d_{\max}^{\text{after}}$ the local maxima before and after this minimum. Two ratios are defined:

$$r_{\text{dec}} = \frac{d_{\max}^{\text{before}} - d_{\min}}{d_{\max}^{\text{before}}}, \tag{6.3}$$

$$r_{\text{inc}} = \frac{d_{\max}^{\text{after}} - d_{\min}}{d_{\max}^{\text{after}}}. \tag{6.4}$$

An edge is removed if both $r_{\text{dec}} \geq \tau$ and $r_{\text{inc}} \geq \tau$ with $\tau = 0.5$, indicating a strong local constriction between two larger caliber segments. After deleting flagged edges, nodes of degree 2 are merged by concatenating centerline and radius profiles and summing lengths. To avoid the removal of major proximal arteries, the terminal branch with the largest mean diameter is protected.

I.1.4 Graph orientation

Graph orientation establishes a consistent proximal–distal direction along the PAT, which is required for hierarchical labeling and downstream feature propagation. The segmentation of the PAT starts around the pulmonary valve, and does not include any other vascular structure, so the anatomical root is defined as the terminal artery with the largest diameter, corresponding to the MPA.

Starting from this root, a breadth-first search (BFS) is performed to orient edges from proximal to distal nodes; geometric attributes (centerline ordering, radius profile) are reordered accordingly to match the assigned direction. The BFS algorithm explores the graph level-by-level, first it visits all neighbors of the root (depth 1), then visits all the neighbors of those nodes (depth 2), and continues until all nodes have been visited. The result is a cleaned, connected, and fully oriented patient-specific PAT graph that serves as the backbone for anatomical labeling and PE characterization.

I.2 Hierarchical labeling

Hierarchical labeling enriches the oriented PAT graph with anatomical context by assigning laterality, lobar affiliation, and hierarchical levels to each edge. Lung and lobe masks are used in combination with the graph structure to derive these labels. The root, or the MPA correspond to the first level in the hierarchical levels inside the PAT, then it divides into the right or left pulmonary artery (second level), it will then divide into the lobar arteries (level 3), wich divides into segmental arteries (level 4), then sub-segmental arteries and so on (level 5+) (See Figure. 6.7).

The root is automatically assigned level 1, and the laterality and lobes attributes are set to None.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

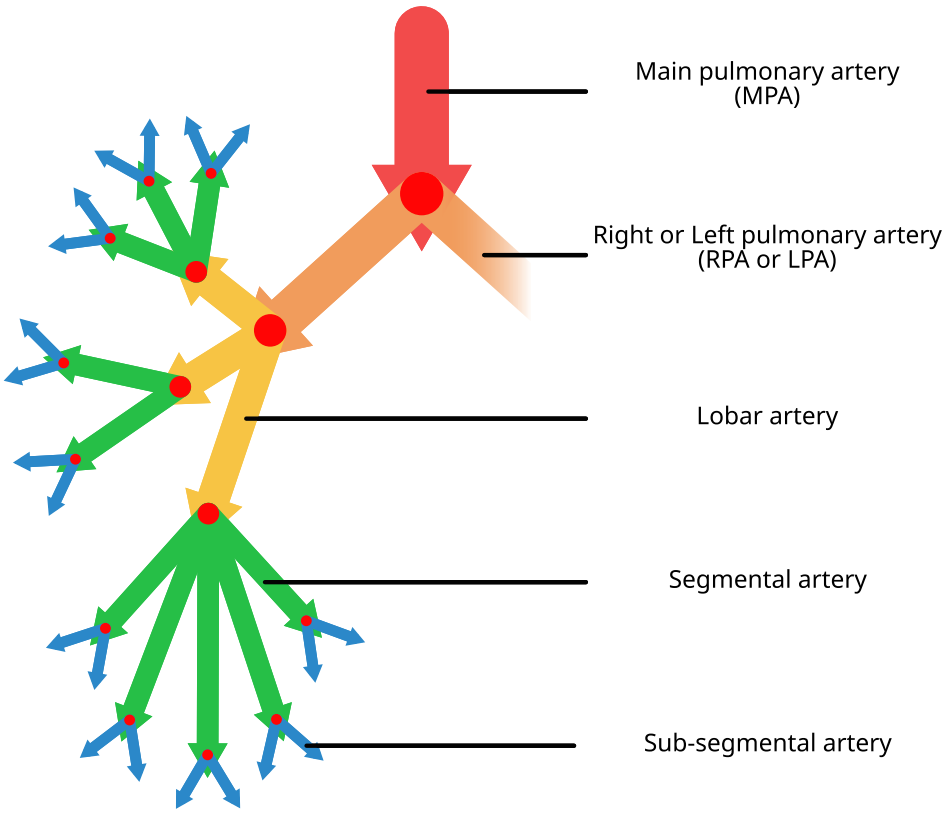


Fig. 6.7: Hierarchical levels of pulmonary arteries, each level is presented with a different color.

I.2.1 Laterality

Laterality is obtained by intersecting each edge volume with the right and left lung masks; edges are labeled right or left depending on the overlap (See 6.8). The first bifurcation distal to the MPA defines the right and left pulmonary arteries (RPA and LPA). They are directly assigned level 2; the lobe attribute will be set to None, as for the MPA.

I.2.2 Lobes affiliation

Lobar affiliation is assigned by intersecting terminal arteries with lobe masks (See Figure 6.9) and propagating the terminal label upstream along parent edges until reaching RPA or LPA (level 3). Edges receiving multiple lobe labels are considered interlobar and retain level 3. We then move on to the level assignment starting from level 3.

I.2.3 Hierarchical level

Within each lobe, hierarchical levels are refined using morphological criteria to distinguish trunks from bifurcations. Figure 6.10 illustrates an example of a trunk and the morphological properties tested to characterize whether an evaluated artery is a trunk or not.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

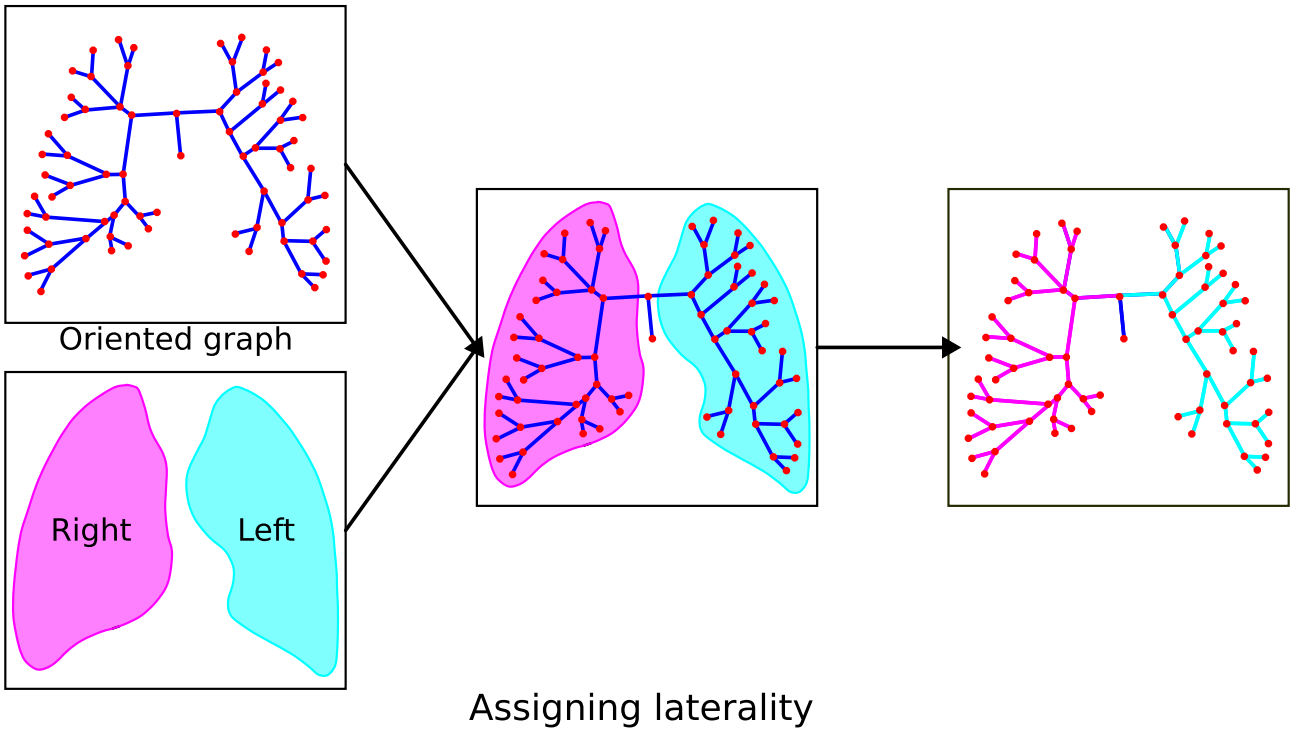


Fig. 6.8: Laterality labeling, obtained by superposition from the right and left lungs masks

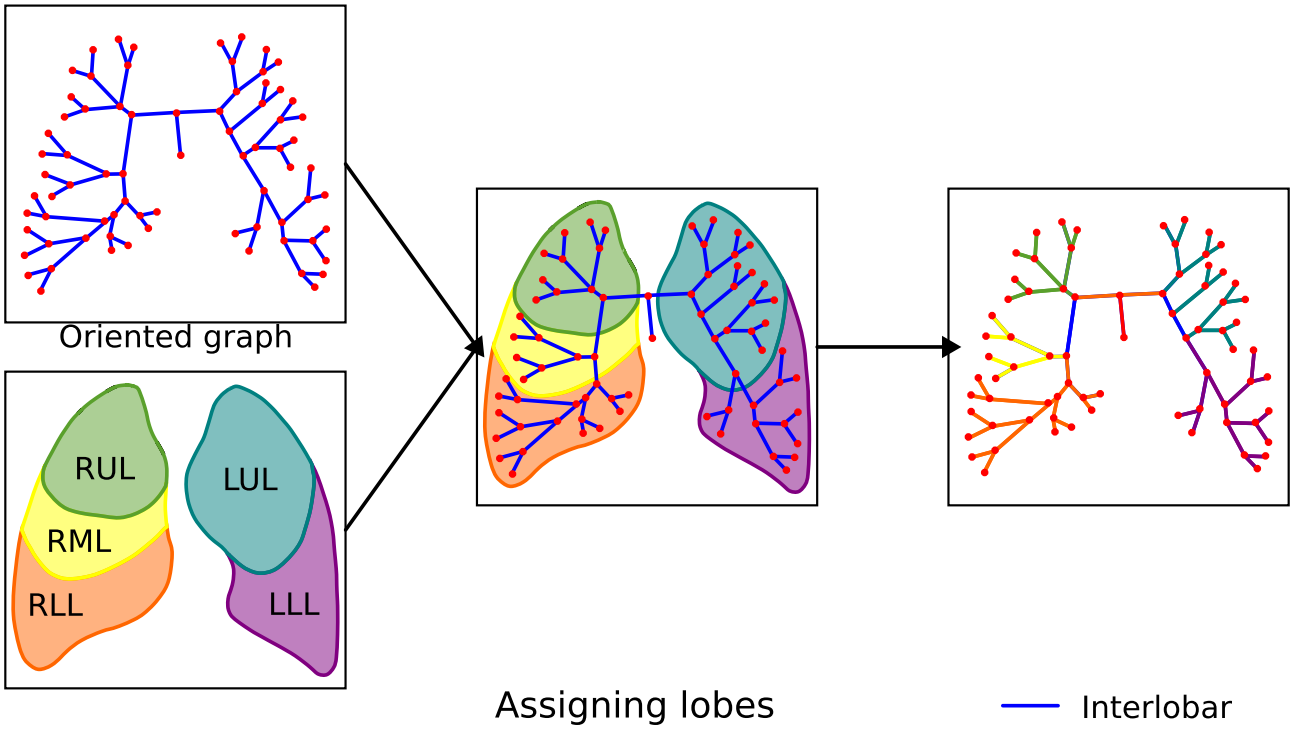


Fig. 6.9: Lobes labeling, obtained by superposition of the lobes masks with the terminal nodes and propagating the labels upwards until level 2. RUL: Right Upper Lobe; RML: Right Middle lobe; RLL: Right Lower Lobe; LUL: Left Upper Lobe; LLL: Left Lower Lobe.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

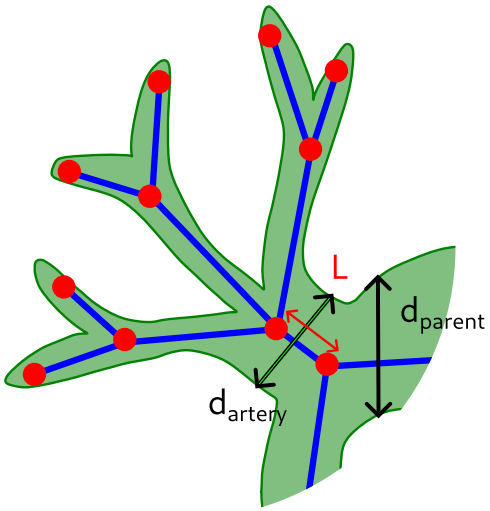


Fig. 6.10: Illustration of the morphological criteria evaluated to determine if an artery is a trunk.

Using the artery diameter d_{artery}, the parent diameter d_{parent}, and the segment length L_{artery}, two ratios are defined:

$$r_{\text{diam}} = \frac{d_{\text{artery}}}{d_{\text{parent}}}, \tag{6.5}$$

$$r_{\text{len}} = \frac{L_{\text{artery}}}{d_{\text{artery}}}. \tag{6.6}$$

An artery is classified as a trunk if $r_{\text{diam}} \geq \theta_d$ and $r_{\text{len}} \leq \theta_l$. To minimize the difference between the automatically detected segmental arteries and literature expectations (Michaud et al. 2023), various configurations of diameter thresholds (θ_d) and length thresholds (θ_l) were tested. We evaluated $\theta_d \in \{0.5, 0.6, 0.7\}$ and $\theta_l \in \{1.5, 2.0, 2.5\}$. The results of each parameter pair for each lobe on our subset of 24 patients are summarized in the result section II.4.1. Thresholds $\theta_d = 0.7$ and $\theta_l = 1.5$ were selected as they minimize the deviation from expected anatomical values across all lobes on our test sample. Trunk arteries inherit the hierarchical level of their parent artery, whereas bifurcations are assigned the parent level + 1.

At the end of the hierarchical labeling step, each edge now possesses, in addition to the **vessel features**, the **location features**:

- the laterality: left, right, or None value for MPA, RPA, and LPA.
- the lobe: RUL, RML, RLL, LUL, LLL, Interlobar or None value for MPA.
- the hierarchical level, 1 for the MPA, two for the RPA and LPA, 3 for lobar arteries, 4 for segmental ateries, and 5 or more for the rest below.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

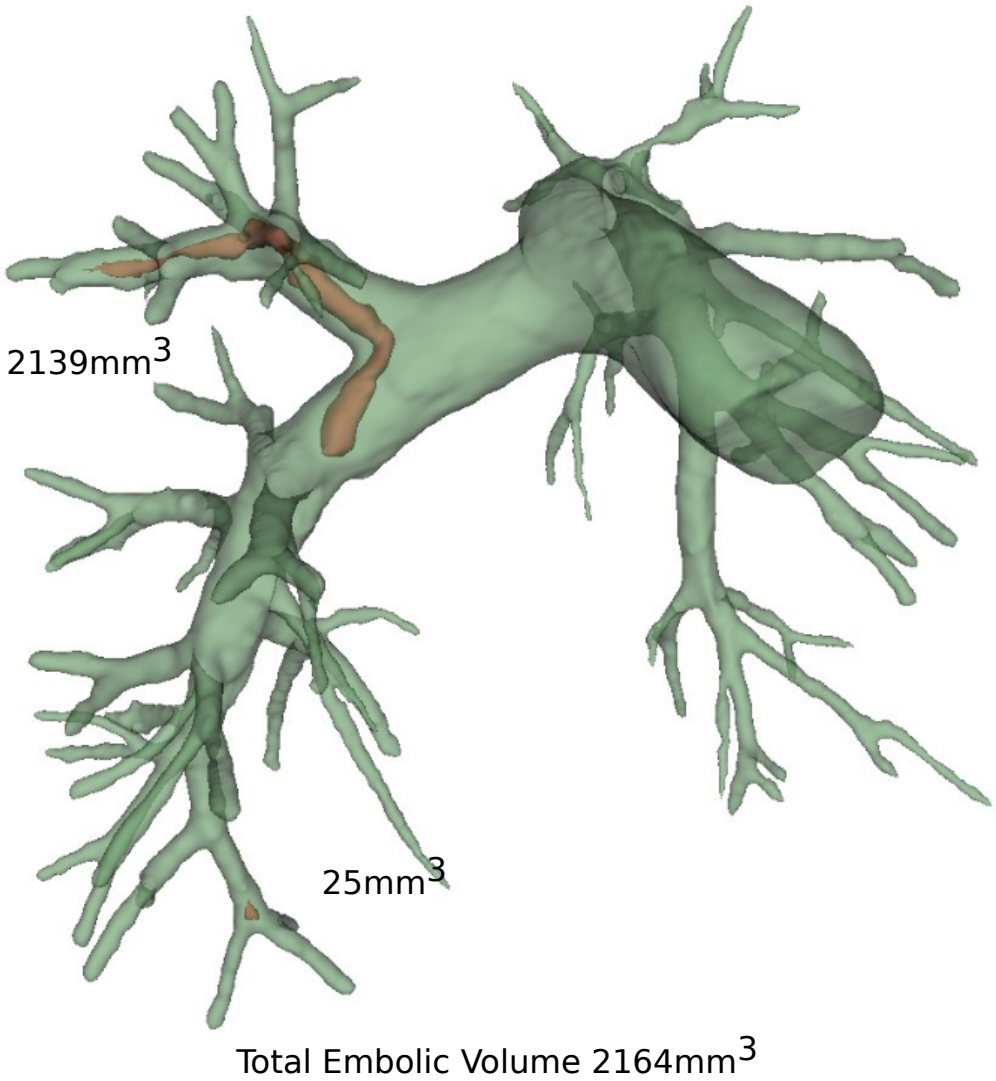


Fig. 6.11: Example of the Total Embolic Volume (TEV) computed on patient 0055.

I.3 PE feature extraction

The PE feature extraction process quantifies thrombotic obstruction from its mask, by identifying each thrombus, computing its volume, and mapping its position relative to the labeled PAT. In Figure 6.11, we can observe two clots in red inside the artery mask in green: one saddle (2139mm^3) and a small homogeneous one (25mm^3), for a TEV of 2164mm^3. For each arterial edge, the volumetric obstruction ratio is calculated as the proportion of thrombus voxels intersecting with the corresponding artery voxels ID in the 3D masks (See 6.12):

$$R_{\text{vol}} = \frac{n_{\text{thrombus}}}{n_{\text{artery}}} \tag{6.7}$$

where n_{thrombus} represents the number of voxels occupied by a thrombus within the vessel segment, and n_{artery} represents the total number of voxels from the artery segment. This ratio quantifies the proportion of artery volume occupied by the thrombus, with $R_{\text{vol}} \in [0, 1]$.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

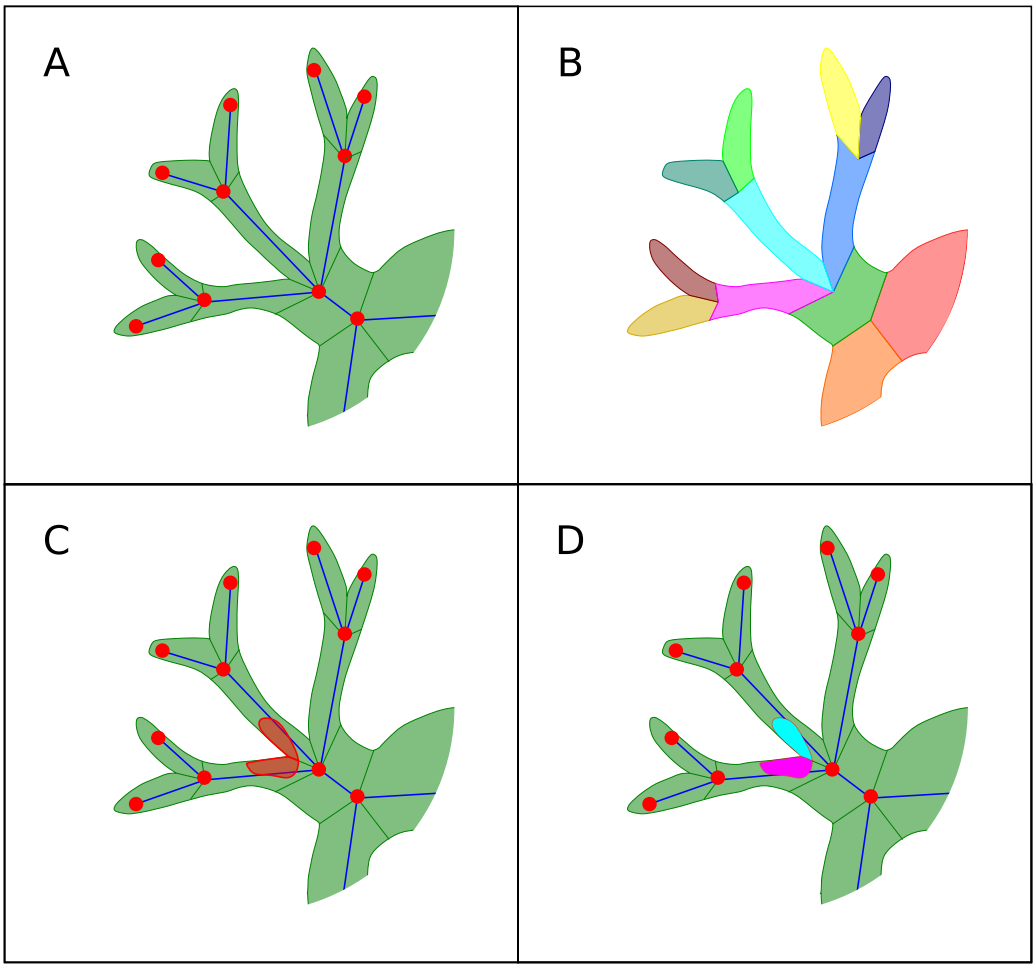


Fig. 6.12: Thrombus intersection with each edge it traverses.
(A) a representation of the Right Upper Lobe view with the artery mask (green) and the modelized graph. (B) The volume of the artery, with each artery volume labeled with the vessel ID. (C) A clot inside the artery. (D) The same clot volume intersected with each corresponding artery it traverses, allowing us to associate the values with each corresponding edge.

To capture localized luminal narrowing, which volumetric metrics alone may miss, we employ Curved Planar Reformation (CPR). This technique computationally ”straightens” the 3D artery segment along its centerline, generating a series of perpendicular cross-sections. For each cross-section, the transversal obstruction is defined as the ratio of clot area to the total arterial lumen area. The maximum transversal obstruction along the artery is a critical feature, as it reflects the most severe cross-sectional narrowing, independent of total clot volume or overall volumetric obstruction.

At the end of this step, each edge possesses, in addition to the **vessel features** and **location features**, the **PE features**:

- Thrombi volume
- The volume obstruction ratio

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

- The maximum transversal obstruction

All local features at the edge level (See Figure 6.13) can now be aggregated into meaningful values and scores in the next step.

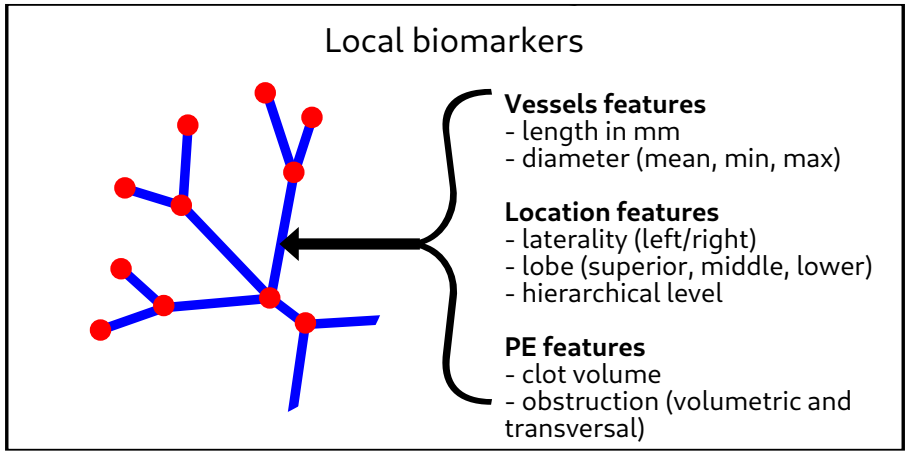


Fig. 6.13: Representation of all the local features aggregated (vessels, location, and PE) at the edge level at the end of the 3 main steps of the pipeline (1) PAT modeling, 2) Hierarchical labeling, and 3) PE characterization.

I.4 Patient level feature extraction

From the enriched directed PAT graph, local biomarkers can be aggregated into patient-level biomarkers. The main patient-level features extracted are:

- Automated Qanadli and Mastora scores derived from edge-wise maximum transversal obstruction.
- TEV and its distribution across pulmonary lobes.
- TEV distribution across hierarchical levels.

I.4.1 Automatic calculation of thrombotic burden scores

To characterize thrombotic burden, we have seen that several imaging-based scores have been proposed, such as the Qanadli and Mastora scores (Mastora et al. 2003; Qanadli et al. 2001), which quantify arterial obstruction based on clot location, number of impacted segmented arteries, or degree of obstruction (See Chapter1, III.3).

While the clinical definitions of the Qanadli and Mastora scores provide a standardized framework (see Section III.3.1 and III.3.1), their direct translation into an automated pipeline presents technical challenges. The original formulations rely on an idealized anatomical template, specifically 20 segments for Qanadli (Qanadli et al. 2001) and 31 predefined artery in mediastinal, lobar, or segmental level for Mastora (Mastora et al. 2003) (See Figure 6.14).

In practice, automated segmentation and graph extraction must deal with patient-specific anatomical variations, pathological vessel deformations, and a different approach caused by

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

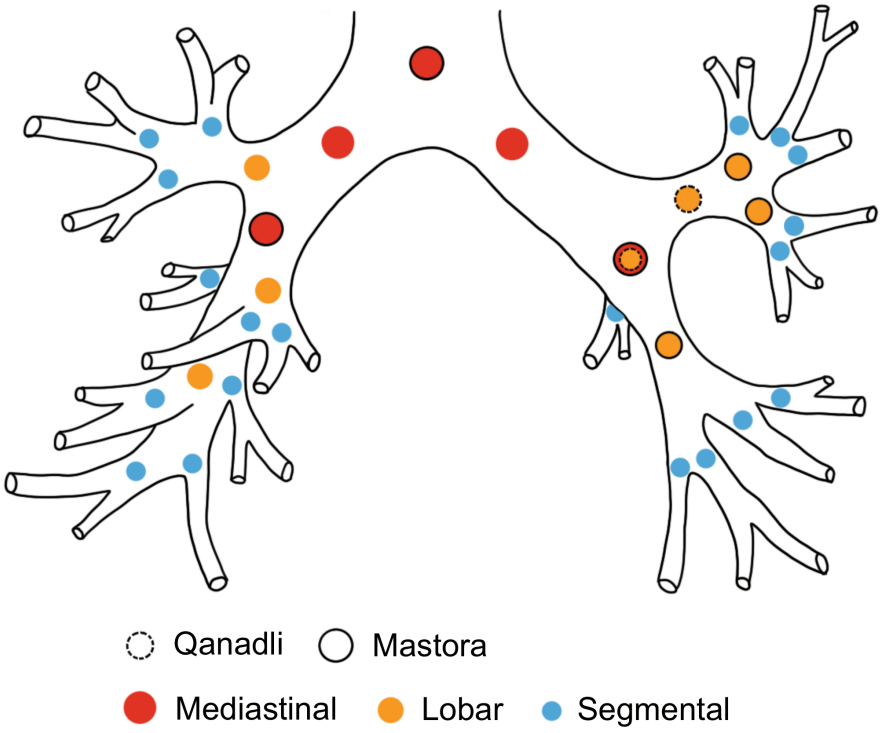


Fig. 6.14: Illustration of the anatomical positions of the arteries considered by both Qanadli and Mastora scores. Modification of a Figure from Alis and Haramati 2023

the graph structure itself (main arteries are divided into several arteries each time an artery emerges from it). For both scores, to quantify the obstruction, we utilize the maximum transversal obstruction measured via Curved Planar Reformation (CPR) along each edge of the graph. We then apply the following adaptations:

- For Qanadli, to mimic the original weighting, the degree of obstruction is defined using a threshold on obstruction value: values below 25% are considered non-obstructed, values between 25% and 75% correspond to partial obstruction (weight 1), and values above 75% to total obstruction (weight 2). The final score considers the number of distal segments emerging from the obstruction location within the patient's specific graph structure.
- For Mastora, obstruction grades follow the original definition: 1 (0–24%), 2 (25–49%), 3 (50–74%), 4 (75–99%), and 5 (100%).
- The mediastinal points are treated as follows. For the Qanadli score, it is assigned to edges with hierarchical level 2. For the Mastora score, it includes edges with levels 1 and 2, as the two mediastinal points below level 2 cannot be automatically identified with high confidence in our graph-based approach.
- For both scores, lobar arteries (traditionally considered as 2 or 3 locations per lung) are adapted to include all edges with level 3 labeling. This ensures that each branching artery is accounted for, allowing us to capture the entire lobar level based on the graph structure.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

In conclusion, for the Qanadli score, by merging hierarchical levels when necessary and utilizing the graph's connectivity to determine "distal segments" the implementation preserves the clinical logic of this score while ensuring robustness against anatomical variability. And for the Mastora score, by relying on the graph-based decomposition, our approach preserves the logic of "mediastinal, lobar, and segmental" categorization but adapts it to the elements effectively detected in the individual patient CT.

I.4.2 TEV distribution

While the TEV provides a global measure of the thrombotic presence, it lacks the spatial context necessary to characterize the repartition and distinguish between proximal or distal embolism. To address this, our pipeline utilizes the PAT graph to compute the TEV distribution across pulmonary lobes and hierarchical levels.

By intersecting the thrombi segmentation mask with the automatically generated lobe masks (RUL, RML, RLL, LUL, LLL), the pipeline quantifies the percentage of the total thrombi volume residing within each lung territory. This distribution assesses the spatial heterogeneity of the embolic involvement. The hierarchical labeling of the arterial graph (levels 1 to 6+) allows for the characterization of the embolic burden based on the depth of the obstruction within the vascular tree. In our model, levels 1–3 (MPA to Lobar arteries) are categorized as proximal, while levels 4 and bellow (starting from segmental level) are considered distal. Figure 6.15 presents the hierarchical level characterization of a thrombus in the RUL.

This spatial decomposition of embolic burden across lobes and hierarchical levels thus offers a principled basis for distinguishing clinically distinct PE profiles that would otherwise appear identical when summarized by a single volumetric measure, demonstrating the added value of the graph-based representation developed throughout this chapter.

II Results

The following section reports the quantitative evaluation of the step of the pipeline, from graph construction and correction to hierarchical labeling and biomarker extraction, on the annotated sample of 24 patients and the full PERSEVERE cohort of 353 patients.

II.1 From voxel-graph to branch graph

For both voxel-level and branch-level graphs, the following statistics are computed:

- Number of nodes and edges.
- Number of connected components (should ideally be 1).
- Number of terminal nodes ($degree = 1$, vessel endpoint).
- Number of linear nodes ($degree = 2$, Intermediate points on vessel).
- Number of bifurcation nodes ($degree \geq 3$, branching point).
- Number of cycles (indicates loops in the vascular structure).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

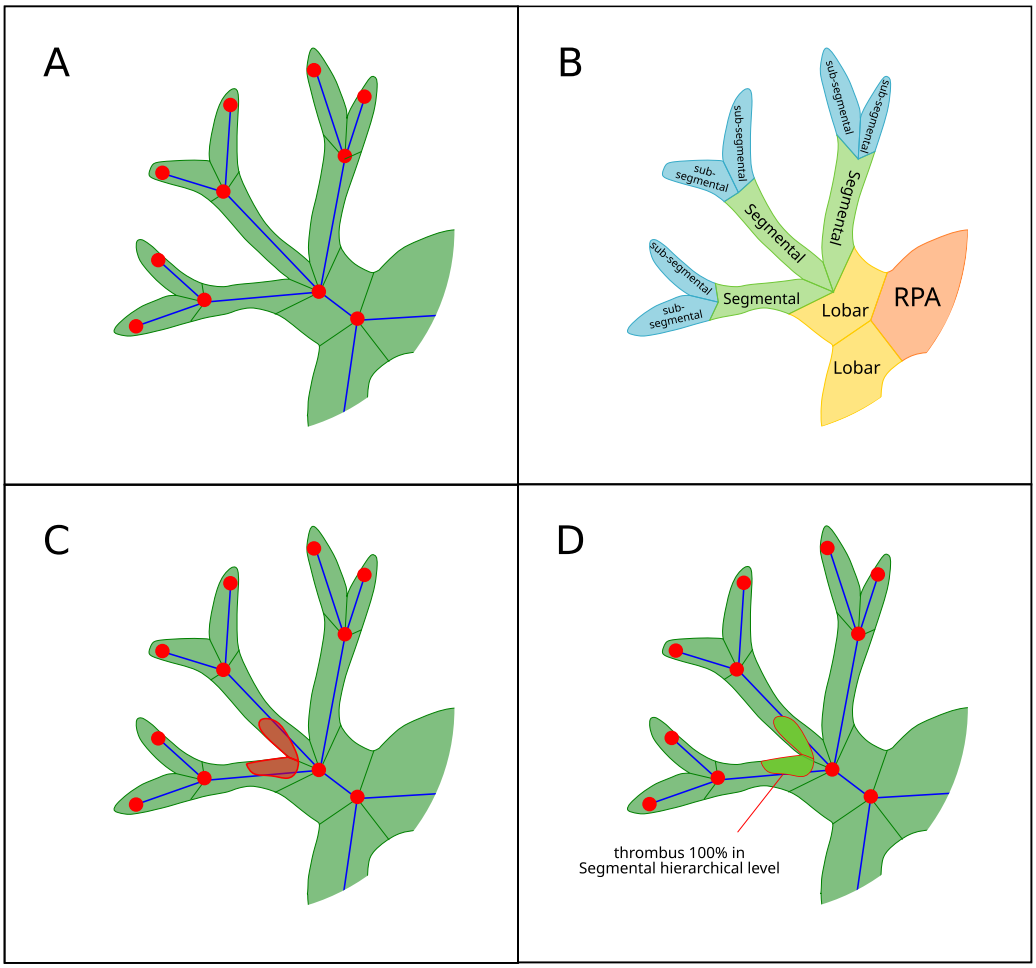


Fig. 6.15: Thrombus intersection with each hierarchical level it traverses in the RUL. (A) RUL view with the artery mask (green) and the modelized graph. (B) The volume of the artery, with each artery volume labeled with the associated hierarchical level. (C) A clot inside the artery. (D) The same clot volume intersected with each corresponding artery it traverses, allowing us to associate the hierarchical level with each corresponding edge.

- Voxel coverage percentage:

$$\text{Coverage}(\%) = \frac{|\text{Covered Voxels}|}{|\text{Original Skeleton Voxels}|} \times 100. \tag{6.8}$$

We performed the quality check for the sample on 24 patients, first testing the different steps during pipeline development, and then on the 353 PERSEVERE segmentations. Table 6.1 summarizes the statistics.

We can observe that the average number of edges is more important in our annotation sample than in the whole database segmentation ($370 \pm 130 > 243 \pm 66$). This can be explained by the extensive application during the annotation process, but we can note that there is also less variability in the segmentations. The 0.01 Standard deviation for the coverage comes from

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Graph	Nodes	Edges	Cycles	Voxel Coverage (%)
24 annotations				
Voxel	6154 ± 1818	6154 ± 1818	1.2 ± 1.5	100.0 ± 0.0
Branch	370 ± 130	370 ± 130	1.2 ± 1.5	100.0 ± 0.0
353 segmentations				
Voxel	4174 ± 1191	4173 ± 1191	0.7 ± 1.2	100.00 ± 0.00
Branch	243 ± 66	243 ± 66	0.7 ± 1.2	100.00 ± 0.01

Tab. 6.1: Graph construction main statistics from voxel-level to branch-level representation.

two patients with coverage of 99.96% for patient 0345 (missing 2 voxels from the skeleton) and 99.78% for patient 0351 (missing 11 voxels). This was corrected easily, so these two patients are included in the following results.

II.2 Graph error correction and comparison with VesselGraph pipeline

We tested Voreen and our method with a sample of 24 completely annotated patients. We only tested it on 24 patients due to the fact that `VesselGraph` graph generation requires manual interaction through the graphical interface.

Because MPA is a very large artery, it is more sensitive to the irregularities at the surface of the mask. For performances, we want to make a special focus on the MPA, since it is necessary for our orientation step, hierarchical labeling, and subsequent steps, that there is no error from spurious branches, assigning a wrong beginning of the PAT. In Figure 6.16, we can see the illustration of a patient and the resulting graphs. We can see that on the graph from our construction step, there was one spurious branch on the MPA before correction. With the `VesselGraph` result, there are more artifacts that are missed.

Table 6.2, reports graph statistics for the graph from Voreen, with the graphs from our construction, and then the correction step. It includes the number of nodes and edges, the number of spurious branches originating from the MPA (MPA artifacts), and the number of cycles detected in the resulting graphs.

Graph type	Nodes	Edges	Artifact on MPA	Cycles
VesselGraph	370 ± 108	370 ± 109	4.0 ± 4.1	1.1 ± 1.4
Constructed graph	370 ± 130	370 ± 130	1.4 ± 1.7	1.2 ± 1.5
Corrected graph	292 ± 112	291 ± 112	0.0 ± 0.0	0.1 ± 0.3

Tab. 6.2: Construction and cleaning statistics on the 24 PERSEVERE annotated sample.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

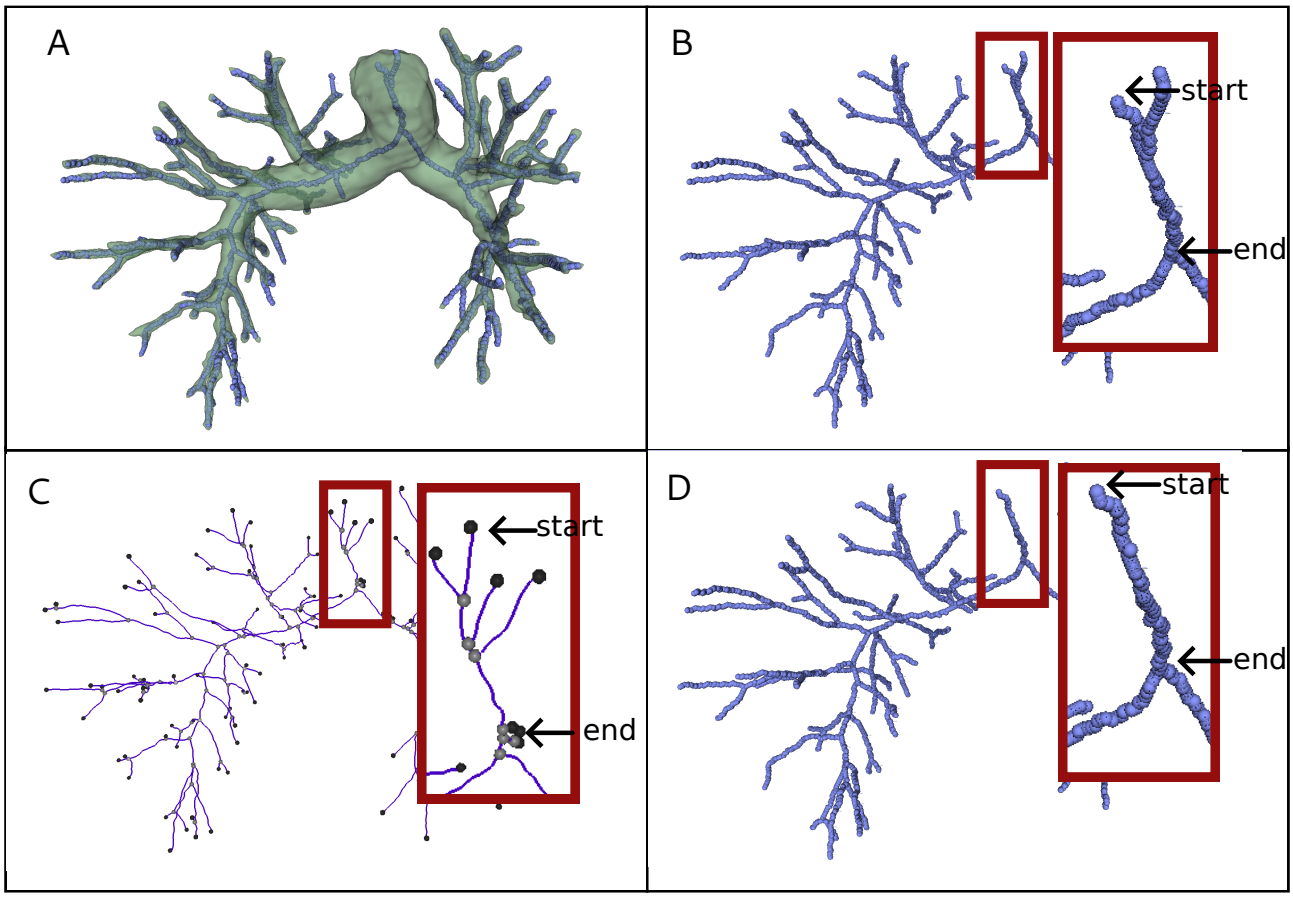


Fig. 6.16: Illustration of the artifact cleaning step on patient 0028 from the annotated PERSEVERE sample.
(A) The arteries mask (green) and the graph before cleaning, visualized with 3D Slicer. (B) Only the graph from before the cleaning process, with a highlight of the Main Pulmonary Artery (MPA) with one spurious branch. (C) The graph from `VesselGraph` is visualized from their interface. The highlight shows many remaining artifacts on the MPA. (D) Our graph, after the correction, shows no artifact remains on the MPA.

Spurious branches For the 24 annotated patients, the number of nodes and edges was reduced by $22\% \pm 5\%$, eliminating all MPA-related artifacts (from 4.0 ± 4.1 to 0.0 ± 0.0) and reducing the average number of cycles per patient from 1.2 ± 1.5 to 0.1 ± 0.3. In comparison, `VesselGraph` left an average of 4.0 ± 4.1 MPA artifacts and 1.1 ± 1.4 cycles. The maximum number of cycles present for one patient after the Voreen pipeline is 5, whereas it is 1 for our pipeline.

Cycle correction Regarding cycle details, on the sample of 24 PERSEVERE patient sample, of the 13 patients with a cycle, only two still have one remaining. The remaining cycles are ones that require manual intervention. It is important to note that since the PAT is topologically a tree structure, any cycle detected in the graph is necessarily a non-physiological artifact, arising from segmentation imperfections such as contact between adjacent arteries in the mask. The remaining cycles after automated correction are cases in which this contact extends over

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

a long segment of two adjacent branches, likely due to local tissue deformation or pathological changes in the surrounding parenchyma. These cases require manual intervention to distinguish the two vessels connected and cannot be resolved automatically without the risk of removing an anatomically valid artery.

II.3 Graph-based representation of the PERSEVERE cohort

On the 353 PERSEVERE segmentation dataset, the topology statistics before and after the corrections step is summarized in Table 6.3. Since a post-processing was applied, we indeed have the confirmation that PAT segmentations masks all only have one connected component, as attested by only $\beta_0 = 1$ values. At the graph construction step, 138 patients were flagged for cycles in the mask. This was provoked by contact between vessels; 90 patients were automatically handled by our correction step, but some cycles remained unresolved. Those unresolved residual cycles ($\beta_1 > 0$) were manually corrected in 48 patients to ensure accurate PAT topology and that this would not cause errors in the rest of the pipeline. As we explained previously II.1, only two patients had $\beta_2 > 0$, due to a few enclosed cavities within the masks, which were easily corrected, so there is no more $\beta_2 > 0$.

Metric	**Before correction**	**After correction**
Nodes	243 ± 66	193 ± 57
Edges	243 ± 66	192 ± 57
β_0 (connected components)	1.2 ± 1.5	1.0 ± 0.0
β_1 (cycles)	0.7 ± 1.2	0.5 ± 0.9
β_2 (holes)	0.0 ± 0.0	0.0 ± 0.0

Tab. 6.3: Topological statistics of the final arterial segmentations (353 patients).

II.4 Hierarchical annotation of the PAT graph

II.4.1 Selection of the morphological threshold to detect trunks.

Table 6.4 reports the mean number of segmental arteries on the 24 patients sample, automatically detected per pulmonary lobe, for each pairwise configuration of threshold $\theta_d \in \{0.5, 0.6, 0.7\}$ and $\theta_l \in \{1.5, 2.0, 2.5\}$ compared with expected anatomical values. The expected values are based on standard anatomical literature. The configuration $\theta_d = 0.7, \theta_l = 1.5$ (bold) was selected as it yields the lowest overall deviation from the expected number.

II.4.2 Application of the morphological criteria to the whole PERSEVERE dataset

To ensure the reliability of the hierarchical labeling, a quality check was performed on the cohort. This process identified two primary categories of errors. First, the "unknown edges" (edges failing to receive a hierarchical or lobar label). These failures were primarily caused by morphological irregularities in the proximal arterial mask or by segmentation irregularities/discontinuities that affected the MPA centerline. Where possible, simple mask corrections

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Configuration		RUL	RML	RLL	LUL	LLL
θ_d	θ_l	(Exp: 3)	(Exp: 2)	(Exp: 5)	(Exp: 5)	(Exp: 4)
0.7	**1.5**	**4.21±1.80**	**2.58±2.52**	**5.29±1.81**	**5.17±1.46**	**5.46±2.10**
0.7	2.0	4.83±2.32	2.62±2.58	6.71±2.34	5.38±1.65	5.96±2.24
0.7	2.5	5.50±2.80	**2.58±2.52**	8.08±2.10	5.38±1.65	6.83±2.19
0.6	1.5	5.29±1.88	**2.58±2.52**	6.00±1.96	5.79±1.76	5.83±2.11
0.6	2.0	7.04±2.88	2.79±2.74	7.96±2.59	7.04±2.26	6.46±2.31
0.6	2.5	8.42±3.63	3.08±2.80	10.00±2.60	7.42±2.48	8.17±2.23
0.5	1.5	5.46±1.89	**2.58±2.52**	6.21±2.04	6.67±3.25	6.08±2.16
0.5	2.0	7.67±2.87	3.00±2.92	8.29±2.72	8.54±4.02	6.92±2.45
0.5	2.5	10.12±3.57	3.38±2.75	10.54±2.68	10.25±4.62	9.12±2.49

Tab. 6.4: Comparison of mean detected segmental arteries (± Std) for different morphological thresholds.

were applied to resolve these issues. Patients for whom mask degradation was too severe to correct were excluded from the analysis. Second, the "missing lobes" (lobes for which no arterial branches were detected). In a number of these cases, the absence of detected arteries was clinically justified by collapsed or severely damaged lobes or by vessels not detectable in the image. When the absence of vessels in a lobe was anatomically correct, and the rest of the PAT segmentation quality was correct, these cases were retained in the dataset to preserve anatomical fidelity. A detailed per-patient listing of identified errors and applied corrections is provided in the Appendix (VII).

The final result side to side of chosen morphological criteria and set $\theta_d 7$ and θ_l are summarized in Table 6.5. On the full cohort, the mean number of detected segmental arteries per lobe was close to expected anatomical values, with deviations generally below one segment per lobe, confirming the plausibility of the hierarchical labeling strategy despite underlying anatomical variability and pathologies.

Lobe	Expected number of segmental arteries	Mean number of segmental arteries automatically detected	
		24 annotations	353 segmentations
RUL	3	4.2 ± 1.8	3.4 ± 1.5
RML	2	2.6 ± 2.5	1.7 ± 1.0
RLL	5	5.3 ± 1.8	4.9 ± 2.2
LUL	5	5.2 ± 1.5	4.9 ± 1.2
LLL	4	5.5 ± 2.1	5.3 ± 2.2
Total	19	22.7 ± 4.9	20.2 ± 4.3

Tab. 6.5: Mean number of segmental arteries for 24 annotations and to the 353 segmentations

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

II.5 Thrombotic burden characterization through graph-derived biomarkers

This section reports the evaluation of the biomarkers derived from the graph-based pipeline. We first assess the agreement between automatically computed thrombotic burden scores and their manual value calculated by an expert, before examining the spatial distribution of embolic burden across lobes and hierarchical levels.

II.5.1 Comparison of the manual VS automatic scores

The agreement between automatic and manual thrombotic burden scores was evaluated using Bland-Altman analysis (see Figure 6.17 for the Qanadli score and Figure 6.18 for Mastora score) and Spearman correlation (Table 6.6).

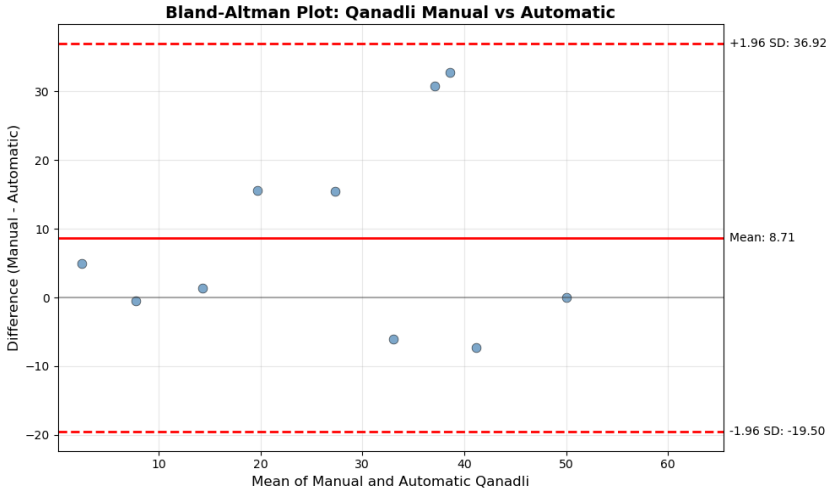


Fig. 6.17: Bland-Altman for Qanadli manual VS automatic scores.

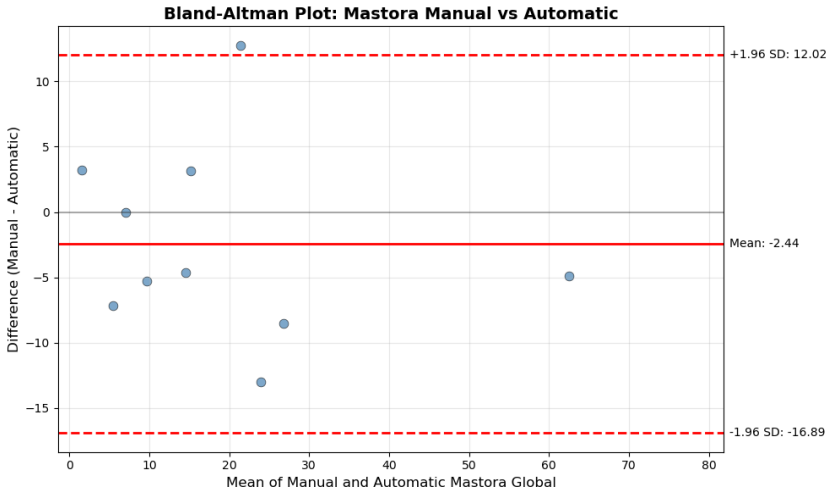


Fig. 6.18: Bland-Altman for Mastora manual VS automatic scores.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Scores	Mean diff ± Std	Lower LoA	Upper LoA	Spearman correlation ρ (p-value)
Qanadli	4.70 ± 10.45	-15.78	25.18	0.83 (0.0029)
Mastora	-2.45 ± 7.39	-16.93	12.02	0.86 (0.0015)

Tab. 6.6: Bland-Altman and Spearman correlation

The Bland-Altman analysis revealed better agreement for the Mastora score compared to the Qanadli score. Specifically, the Mastora score exhibited a smaller systematic bias (mean difference: -2.45% vs. 4.70%) and narrower limits of agreement (range: 28.95% vs. 40.96%), indicating more consistent alignment between manual and automatic measurements. The mean differences suggest a slight underestimation for the Mastora score and a slight overestimation for the Qanadli score.

Spearman correlation coefficients confirmed a strong monotonic association between manual and automatic scores, with $\rho = 0.83$ ($p = 0.0029$) for the Qanadli score and $\rho = 0.86$ ($p = 0.0015$) for the Mastora score. These results demonstrate statistically significant agreement between the two measurement approaches.

II.5.2 TEV distribution to characterize distal or proximal thrombotic load

The TEV distribution across lobes and hierarchical level provides a spatial description of embolic burden, enabling a clear distinction between proximal and distal PE patterns that might otherwise present with identical global clot volumes. In Figure 6.19, two patients can present with the exact same total embolic volume (e.g., 2cm^3) but exhibit vastly different obstruction patterns:

- **Patient A (Proximal):** The majority of the TEV is concentrated in the right lung (100%), more specifically in the RUL (66%). The TEV has a majority of presence in level 3 (72%). This signifies a proximal obstruction.
- **Patient B (Distal):** The TEV is distributed in the right lung (75%), and 52% of the TEV is located in the RUL. The TEV is distributed across level 4, 5, and 6+ (subsegmental) for a total of 98%. Despite having the same volume as Patient A, the clinical management and prognostic implications may differ significantly due to the peripheral nature of the clots.

II.6 Computational time of the pipeline

Table 6.7 reports the per-patient execution time for each step of the pipeline, measured across the full PERSEVERE cohort of 353 patients.

On average, processing a single patient requires approximately one minute for pulmonary artery graph modeling, four minutes for hierarchical labeling, and four minutes for PE characterization, for a total of roughly nine minutes per patient. The computation of patient-level features is

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

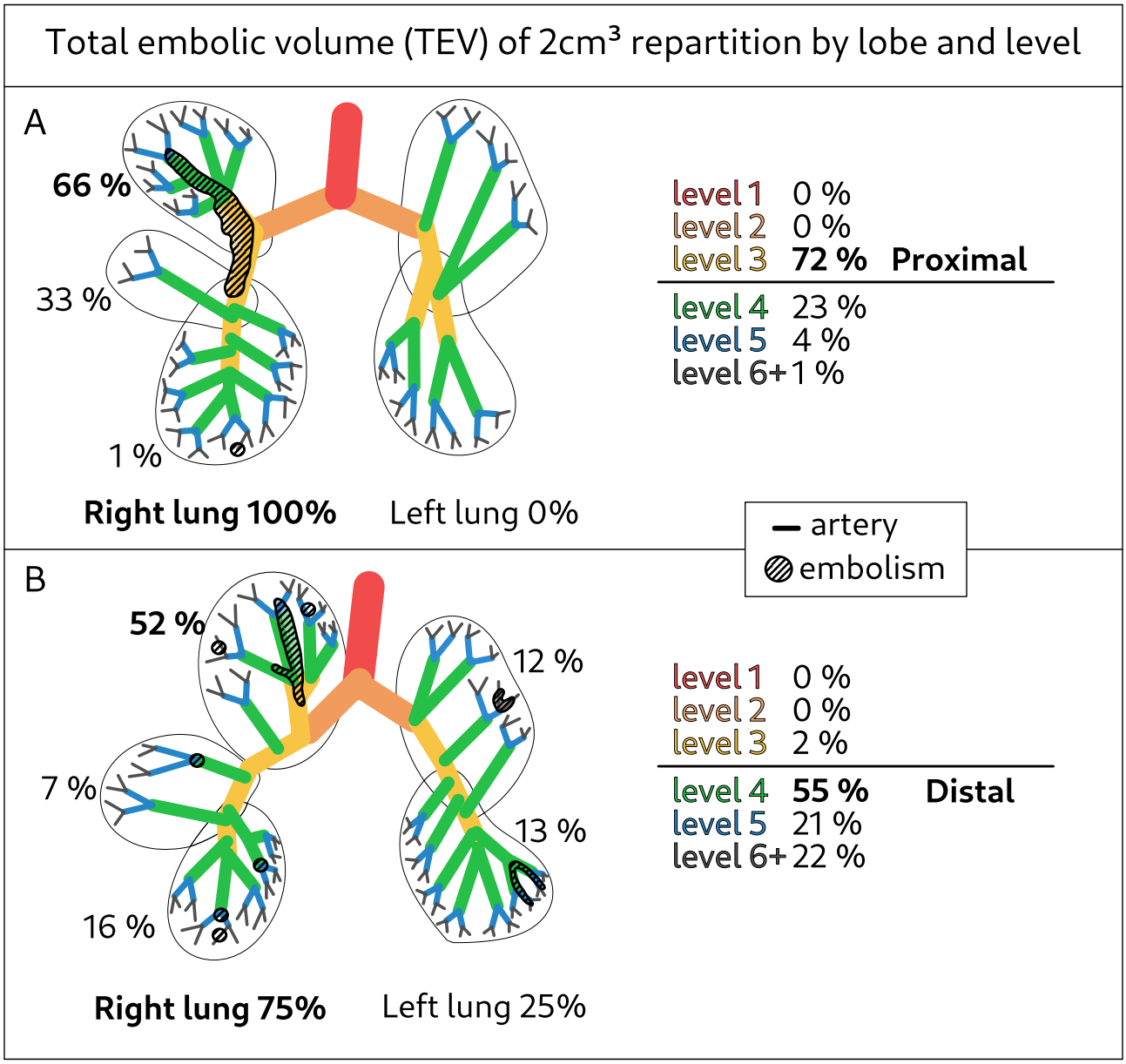


Fig. 6.19: TEV repartition by lobes and hierarchical levels provides a spatial description of embolic burden that can distinguish proximal from distal PE patterns with identical global clot volume.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Step	Execution time
(1) Pulmonary Artery Graph Modeling	
(a) Preprocessing	15.5 ± 4.0 s
(b) Graph construction	16.9 ± 4.4 s
(c) Graph correction	17.3 ± 5.3 s
(d) Graph orientation	12.1 ± 3.3 s
Total (a–d)	61.8 ± 15.8 s (1 min 01 s ± 0 min 15 s)
(2) Hierarchical labeling	238.4 ± 92.7 s (3 min 58 s ± 1 min 32 s)
(3) PE characterization	228.8 ± 178.5 s (3 min 48 s ± 2 min 58 s)
(4) Patient-level feature extraction	0.0 ± 0.2 s
Total (1–4)	528.4 ± 201.9 s (8 min 48 s ± 3 min 21 s)

Tab. 6.7: Per-patient execution time across the PERSEVERE dataset (353 patients).

negligible once the enriched directed graph, with all local features assigned, has been produced. The largest source of variability across patients is observed in the PE characterization step (±2 min 58 s), which reflects differences in the number and volume of thrombi to characterize across patients.

III Conclusion

In this chapter, we presented a graph-based framework for modeling the pulmonary arterial tree and extracting embolism-related biomarkers from CTPA segmentations. Starting from the PAT, thrombi, lung, and lobe masks, the proposed pipeline constructs an anatomically oriented, directed graph of the PAT, enabling the extraction of vessel, location, and PE features at the edge level. The graph correction step demonstrated clear improvements over existing tools such as `VesselGraph`, particularly in eliminating spurious artifacts on the MPA and reducing non-physiological cycles. The hierarchical labeling strategy yielded segmental artery counts consistent with anatomical expectations across all lobes, despite the inherent variability of patient-specific vascular architectures. Local edge-level features were subsequently aggregated into patient-level biomarkers, including automated adaptations of the Qanadli and Mastora thrombotic burden scores, as well as spatial distributions of thrombotic burden. The strong Spearman correlations obtained between automatic and manual scores (ρ=0.83 and ρ=0.86, respectively) confirm that the pipeline produces clinically meaningful measurements. Taken together, these results establish graph modeling as a patient-specific representation of the pulmonary vasculature, capable of accommodating anatomical variability while supporting the extraction of a rich set of biomarkers. However, extracting biomarkers is only one part of the clinical challenge. The ultimate goal is to leverage these features to predict patient outcomes and support clinical decision-making.

More broadly, the graph-based pipeline developed in this chapter is not limited to pulmonary

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

arterial imaging. The core components, including centerline extraction and vessel-level feature computation, are applicable to any tubular vascular structure that can be segmented. The following Chapter 7 explores precisely such a transfer, applying these steps of this pipeline to a different clinical context: neurovascular imaging. In that setting, a conventional imaging modality, limited by low spatial resolution and blooming artifacts from metallic implants, which complicate post-operative follow-up, will be compared with a recently introduced modality that offers ultra-high resolution and material decomposition. By reusing the graph modeling framework presented here, we perform a quantitative comparison of the vessel segmentation quality across both modalities in selected regions of interest (ROIs), providing an illustration of the broader applicability of the tools developed throughout this chapter.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

7 Quantification of SPCCT improvement for brain vessel visualization

Table of contents

I Study context . . . 112
I.1 Acquisition and reconstruction protocol . . . 113
I.2 Preprocessing . . . 113
II Methodology . . . 115
II.1 Vessel segmentation and ROI corrections . . . 115
II.2 Graph-based metric extraction . . . 115
III Results . . . 116
III.1 Subjective image quality analysis . . . 116
III.2 The circle of Willis . . . 116
III.3 Lenticulostriate arteries . . . 118
III.4 Coil surroundings . . . 118
IV Conclusion . . . 121

The graph-based pipeline developed in the previous chapter (see Chapter 6) was designed and validated in the context of pulmonary arterial imaging, where vessels follow a relatively well-defined hierarchical tree structure. To assess whether the core components of this pipeline, namely centerline extraction and vessel-level feature computation, generalize to a different anatomical setting, we applied it to cerebral arterial imaging, where vessels exhibit distinct geometric and morphological properties, including smaller diameters, higher tortuosity, and a more complex spatial organization. This work was carried out in collaboration with Dr.Salim Si-Mohamed, whose research focuses on SPCCT and its clinical applications in neuroradiology.

Dual-energy CT (DECT) is currently the standard modality for cerebrovascular imaging in post-operative follow-up. However, it suffers from limited spatial resolution and significant blooming artifacts in the presence of metallic implants such as coils or stents, which substantially impact the visualization of small vessels in the region of the implant. Spectral photon-counting CT (SPCCT) is a recently introduced modality that offers both ultra-high spatial resolution and material decomposition capabilities, and could therefore address these limitations. The objective of this study is to quantify the contribution of SPCCT compared to DECT for the visualization of cerebral arteries. To this end, we exploit SPCCT and DECT acquisitions of

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

the same patient. We apply our graph modeling pipeline to identify and precisely characterize selected cerebral arteries of interest within three regions of interest (ROIs). The resulting vessel-level measurements are then compared across both modalities to evaluate the quantitative gain offered by SPCCT.

This work was recently accepted to be published in the Journal of Neuroradiology.

I Study context

Neurovascular imaging requires submillimeter resolution to accurately delineate small-caliber intracranial vessels. Conventional CT angiography is constrained by beam-hardening artifacts introduced by high-density materials, such as aneurysm coils, thereby limiting its diagnostic performance. DECT improves contrast and reduces metal artifacts, yet remains suboptimal for visualizing submillimeter structures (Greffier, Viry, et al. 2025). SPCCT is an emerging modality that combines ultra-high spatial resolution with advanced material decomposition capabilities, enabling the visualization of vascular details that were previously undetectable on both conventional and dual-energy CT (Greffier, Villani, et al. 2023).

Both modalities were evaluated using two types of image reconstructions: polychromatic images (abbreviated -p) and virtual monoenergetic images at 40 keV (abbreviated -40keV). Polychromatic images are reconstructed from the full X-ray spectrum and correspond to the standard clinical output. Virtual monoenergetic images, in contrast, are synthetically reconstructed to simulate the attenuation that would be observed at a single photon energy level, here 40 keV. At this low energy level, iodine contrast agents exhibit a substantially higher attenuation coefficient, resulting in an increased contrast-to-noise ratio in vascular structures. Both reconstructions are compared in this study because they represent complementary diagnostic perspectives: polychromatic images reflect standard clinical conditions, while 40 keV monoenergetic images are expected to improve the visibility of small and distal vessels by enhancing iodine-based contrast. Accordingly, the main hypothesis of this study is that SPCCT will outperform DECT across both reconstruction types, with the 40 keV monoenergetic images expected to provide the best vessel conspicuity in both modalities, owing to the combined effect of higher iodine contrast and, in the case of SPCCT, improved spatial resolution.

To focus the comparison on clinically relevant anatomical regions, the neuroradiology experts defined three ROIs of interest, selected for their diagnostic importance in post-operative follow-up and for the challenges they pose to conventional DECT imaging, particularly in the presence of metallic implants. The three ROIs that were selected to be segmented are:

- **The Circle of Willis:** This region provides an overview of the major communicating arteries, as well as a portion of the basilar trunk. It was selected because it contains small perforating arteries that are challenging to visualize with conventional imaging modalities, making it a relevant area to assess the resolution gain offered by SPCCT.
- **Coil Region:** This region corresponds to the area surrounding a cerebral aneurysm coil implant. It was selected because metallic artifacts generated by the coil typically obscure the adjacent arterial lumen on conventional CT, severely limiting post-operative follow-

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

up. Evaluating this region allows assessment of the ability of each modality to reduce metal-induced blooming artifacts and to delineate the arterial lumen near the implant.

- **Left and right Lenticulostriate Arteries (LAs):** These fine perforating vessels, with diameters below 1 mm, are among the smallest clinically relevant intracranial arteries. Similar to the perforating arteries of the basilar trunk, they represent a challenging target for conventional imaging, and were selected to evaluate whether SPCCT permit their reliable segmentation where DECT cannot.

I.1 Acquisition and reconstruction protocol

A 68-year-old male patient with multiple cerebral coils underwent cerebral CT angiography (CTA) on both a clinical DECT (CT7500, Philips Healthcare) and a prototype SPCCT, three days apart in February 2023. Both used helical acquisition at 120 kVp with Iomeprol-400 injection (4 mL/s), bolus tracking at 110 HU in the common carotid artery, and saline flush. Radiation dose was lower for SPCCT (100 mAs) vs. DECT (184 mAs). Reconstructions included polychromatic (DECT-p, SPCCT-p) and 40 keV monoenergetic images (DECT-40keV, SPCCT-40keV). SPCCT used ultra-high-resolution parameters: 1024 matrix, 0.25 mm slice thickness, voxel size $0.25 \times 0.23 \times 0.23$ mm, detailed-2 kernel (cut-off 11.7 lp/cm). DECT used 512 matrix, 0.80 mm slices, voxel size $0.48 \times 0.48 \times 0.80$ mm^3, B kernel (cut-off 9.1 lp/cm).

I.2 Preprocessing

The DECT and SPCCT images differ in spatial resolution and field of view. A resampling and registration step was therefore necessary to align the DECT images and their corresponding segmentation masks to the SPCCT space, ensuring that the selected regions of interest (ROIs) could be meaningfully compared across modalities (See Figure 7.1).

A rigid registration was applied using the ANTsPy library (Avants, Tustison, and Song 2008). Nearest-neighbor interpolation was employed during the registration of the label images to preserve the discreteness of the label indices, whereas linear interpolation was applied to the intensity images to maintain image quality. The result of the preprocessing step is summarized in Table 7.1.

Dataset	**Matrix Size**	**Voxel Spacing (mm)**	**Total Voxels**
Original DECT			
DECT Polychromatic	$512 \times 512 \times 429$	$0.488 \times 0.488 \times 0.400$	112.5×10^6
DECT 40 keV	$512 \times 512 \times 215$	$0.488 \times 0.488 \times 0.800$	56.4×10^6
Resampled & Registered			
DECT Polychromatic	$1024 \times 1024 \times 640$	$0.195 \times 0.195 \times 0.250$	671.1×10^6
DECT 40 keV	$1024 \times 1024 \times 640$	$0.195 \times 0.195 \times 0.250$	671.1×10^6
SPCCT Polychromatic	$1024 \times 1024 \times 640$	$0.195 \times 0.195 \times 0.250$	671.1×10^6
SPCCT 40 keV	$1024 \times 1024 \times 640$	$0.195 \times 0.195 \times 0.250$	671.1×10^6

Tab. 7.1: Technical specifications of the images used in the study.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

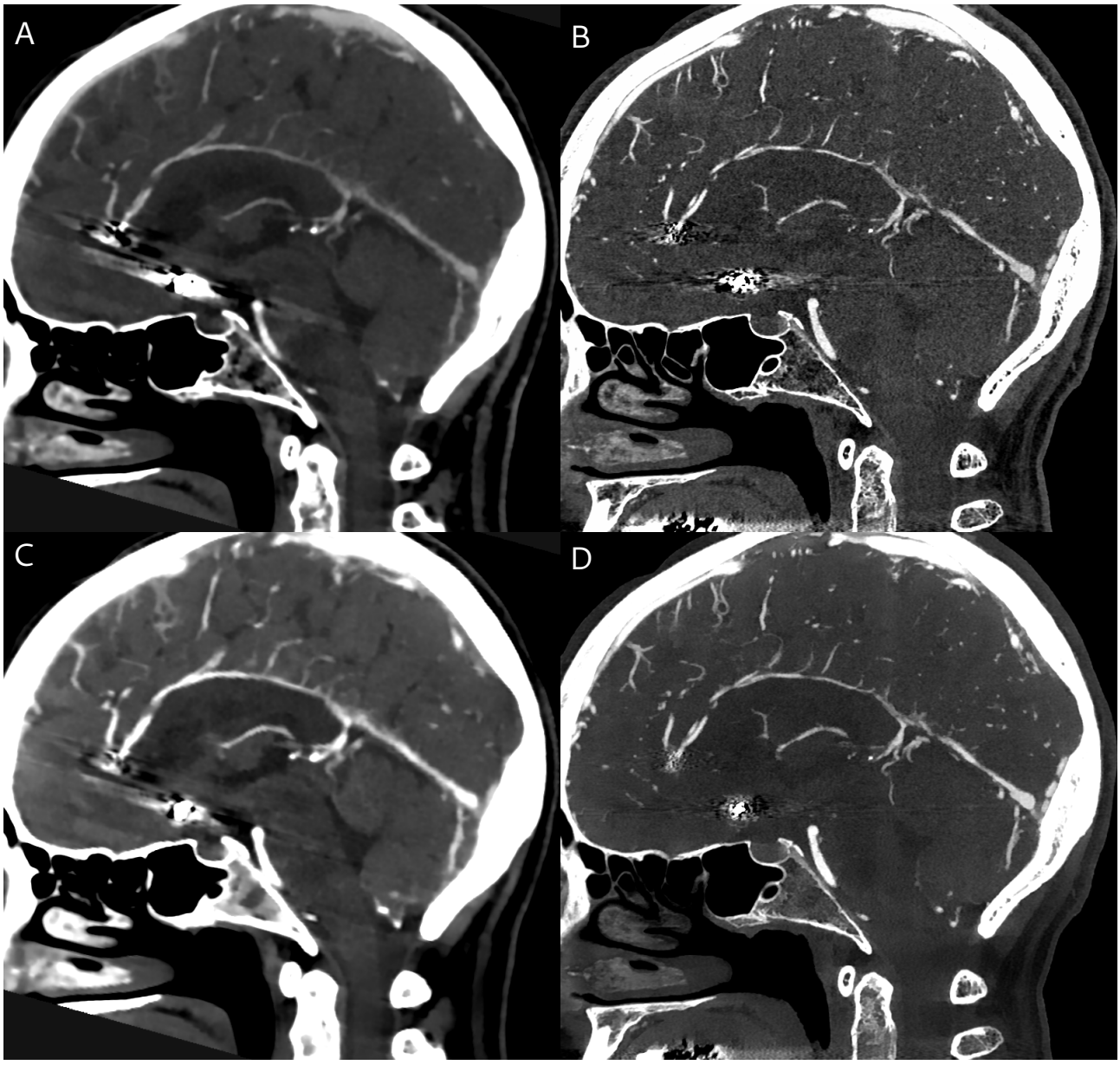


Fig. 7.1: Overview of the 4 image modalities. (A) Dual Energy polychromatique (B) SPCCT polychromatique (C) Dual energy 40 keV (D) SPCCT 40 keV

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

II Methodology

II.1 Vessel segmentation and ROI corrections

Initial segmentations were produced by Pierre Rougé, using the topology-aware U-Net (Rougé, Passat, and Merveille 2024), trained for cerebrovascular segmentation. Raw segmentation models' outputs contained errors: over-segmentation of veins as arteries, under-segmentation near coils, and discontinuities in small vessels. In Figure 7.2, we can observe the segmentation on the DECT-p and the SPCCT-p and the three ROIs that were selected to be corrected.

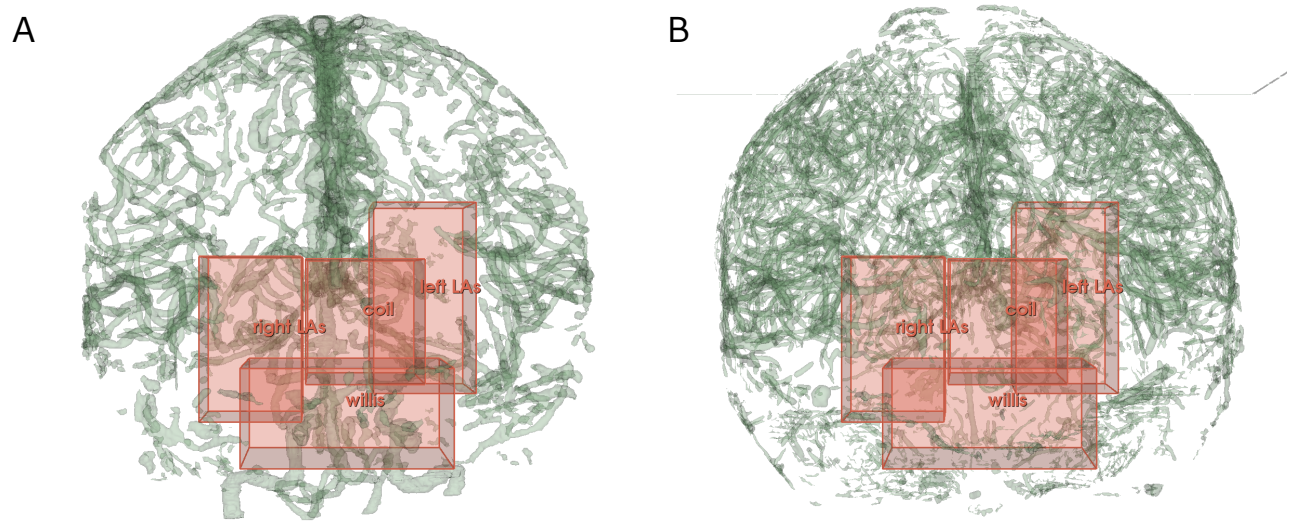


Fig. 7.2: Segmentations in green of (A) Dual Energy polychromatique and (B) SPCCT polychromatique with the four ROIs in highlighted in red.

The presence of these automatic segmentation errors made manual corrections necessary to ensure clinical validity. I corrected each of the ROIs, and corrections were performed using 3D Slicer (v5.8). All corrections were reviewed and validated by a senior neuro-radiologist.

II.2 Graph-based metric extraction

Following validation, the corrected masks were processed through a custom Python pipeline that reuses the graph modeling component described in Chapter 6 I.1. Briefly, the vascular segmentation mask is skeletonized and converted into a graph representation, where each edge represents a vessel segment, and each node represents a bifurcation or endpoint. This representation is particularly well suited to the present comparison as it permits the measurement of per-vessel properties and allows counting the number of distinct arterial segments detected in each modality. The extracted vessel metrics are chosen to capture different aspects of the improvement brought by SPCCT :

- **Total annotated arterial volume** reflects the overall extent of the segmented vasculature.
- **Number of vessels**, defined as the number of distinct edges in the vascular graph.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

- **Diameter per vessel**, extracted via local distance transform along the vessel centerlines, characterizes the detectable vessel size and the caliber distribution of the segmented network.

III Qualitative and Quantitative Results

In this section, after the subjective image quality analysis, we will present the corrected segmentation for each ROI, and the statistics tables of the graph-based brain arteries.

III.1 Subjective image quality analysis

The image interpretation was performed in consensus by two senior radiologists using a five-point Likert scale (1 = unacceptable, 2 = suboptimal, 3 = acceptable, 4 = above average, 5 = excellent). Values less than 3 were considered unsatisfactory for clinical use (see Figure 7.2).

	Overall image quality	**Diagnostic confidence**	**Distal lumen conspicuity**	**Distal lumen sharpness**	**Coil delineation**	**Intra-stent lumen visualization**	**Metallic artefact**
DECT-p	3	2	1	2	1	1	1
DECT-40keV	3	2	1	2	1	1	1
SPCCT-p	5	5	3	5	3	5	4
SPCCT-40keV	5	5	4	5	4	5	4

Tab. 7.2: Subjective analysis of image quality in a 68-year-old man who underwent cerebral CT angiography using dual-energy CT (DECT) and a clinical prototype spectral photon-counting CT (SPCCT).

SPCCT provides superior image quality, diagnostic confidence, and artifact reduction compared to DECT, making it a promising modality for neurovascular imaging.

III.2 The circle of Willis

In figure 7.3, we can observe a slice of the axial view in all the modalities with volumetric information and diameter information of the main arteries detected in the segmentation mask. The quantitative analysis of the Circle of Willis (Table 7.3) reveals a significant gap in performance between the two modalities. SPCCT demonstrated a marked improvement in the detection of small-caliber arteries. Specifically, SPCCT identified more than double the number of arterial segments compared to DECT (53 for SPCCT-p and 47 SPCCT-40keV vs. 26 DECT-P and 47 DECT-40keV).

This increased detection is reflected in the total annotated volume, which was approximately 20% higher in SPCCT images. Most importantly, the minimum detectable diameter reached 0.62 mm with SPCCT, whereas DECT was limited to vessels larger than 1.04 mm. These results suggest that SPCCT successfully captures distal branches and perforating arteries that remain below the detection threshold of conventional energy-integrating detectors. It should be noted that these differences reflect both the intrinsic resolution advantage of SPCCT and the manual corrections applied during the annotation process: vessels that were genuinely invisible

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

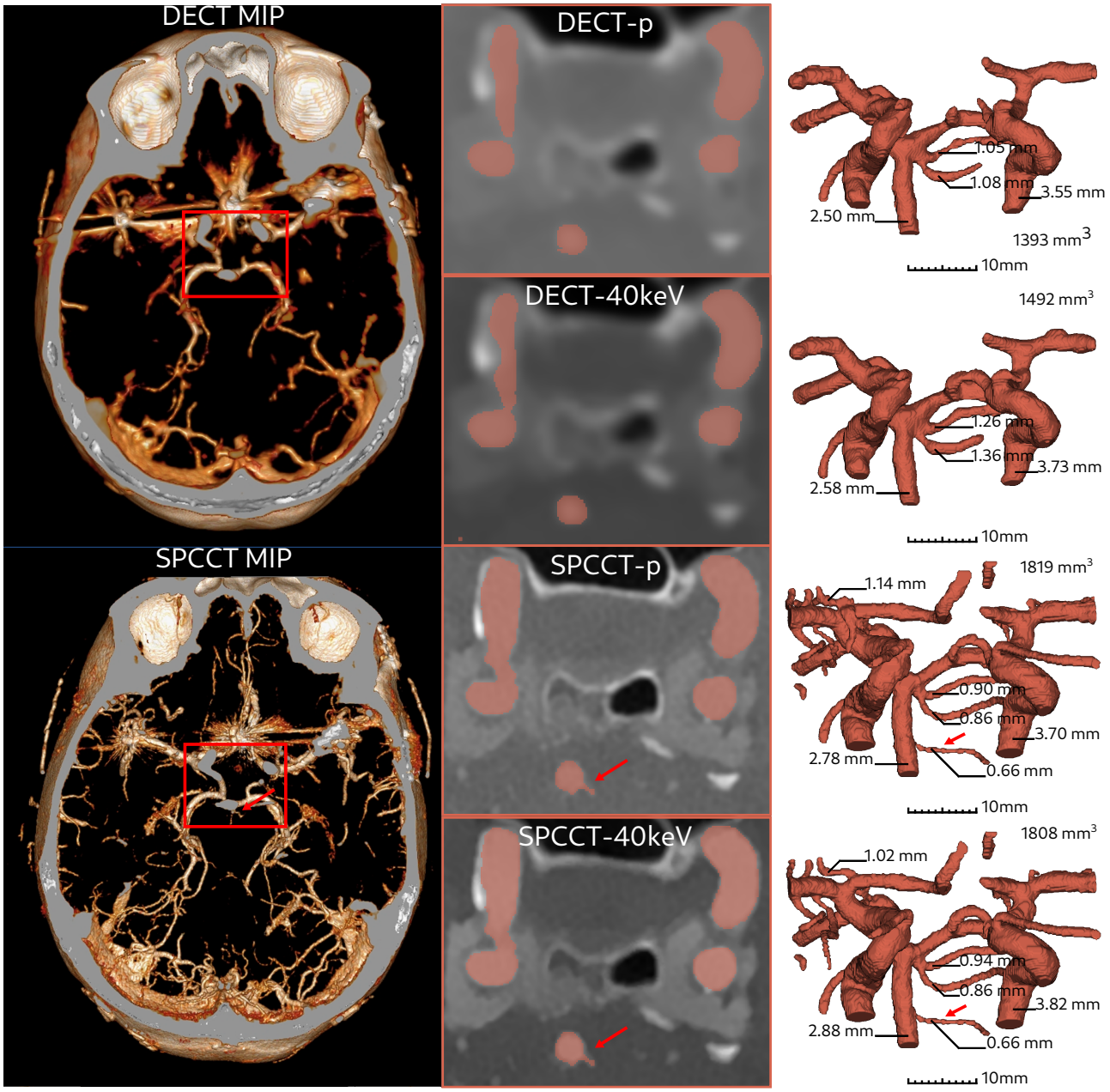


Fig. 7.3: Axial Maximum-Intensity Projections (MIP) with 15 mm slab thickness, and zoomed slices for DECT-p, DECT-40keV, SPCCT-p and SPCCT-40keV with corresponding segmentations. Red arrow indicated a perforating artery of the basilar trunk.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Image type	Volume (mm^3)	Number of artery	Mean diameter (mm)	Min diameter (mm)	Max diameter (mm)
DECT-p	1393	26	2.04	1.04	3.55
DECT-40keV	1492	23	2.23	1.26	3.58
SPCCT-p	1819	53	1.75	0.62	3.71
SPCCT-40keV	1808	47	1.89	0.67	3.79

Tab. 7.3: Circle of Willis: Post-correction metrics. With SPCCT modality we detected 2× more vessels, down to 0.62 mm.

on DECT could not be annotated, regardless of the correction effort, whereas vessels that were detectable but missed by the segmentation model were recoverable through correction in both modalities. Furthermore, the higher spatial resolution of SPCCT produces sharper vessel wall delineation, whereas the lower resolution of DECT causes the vessel lumen to appear blurred, which leads to an overestimation of vessel diameter in DECT, or rather a better definition of the vessels in SPCCT.

III.3 Lenticulostriate arteries

The left and right LAs were evidence of SPCCT's resolution capabilities (see Figure 7.4). These fine vessels were exclusively segmentable on the SPCCT datasets; they were undetectable for reliable annotation on both DECT-p and DECT-40keV.

Image type	Volume L/R (mm^3)	Number of artery L/R	Mean diameter L/R (mm)	Min diameter L/R (mm)	Max diameter L/R (mm)
SPCCT-p	249/362	10/30	0.74/0.86	0.57/0.65	1.03/1.41
SPCCT-40keV	240/317	9/22	0.94/0.81	0.58/0.59	2.49/1.12

Tab. 7.4: Lenticulostriate arteries (SPCCT only; undetectable on DECT).

As shown in Table 7.4, 40 keV imaging improved the contrast-to-noise ratio, resulting in a slightly lower vessel count than polychromatic images (e.g., 22 vs. 30 on the right side). The lower number of vessels can be explained by improved distinction, allowing the removal of false positives (veins) that were segmented (an example is highlighted in Figure 7.4 with the blue arrow). The smaller diameters detected are similar in the LAs (0.58mm) and in the left LAs (0.59mm).

III.4 Coil surroundings

The region surrounding the cerebral coil provided a direct measure of the SPCCT's ability to handle high-density artifacts (See Figure 7.5).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

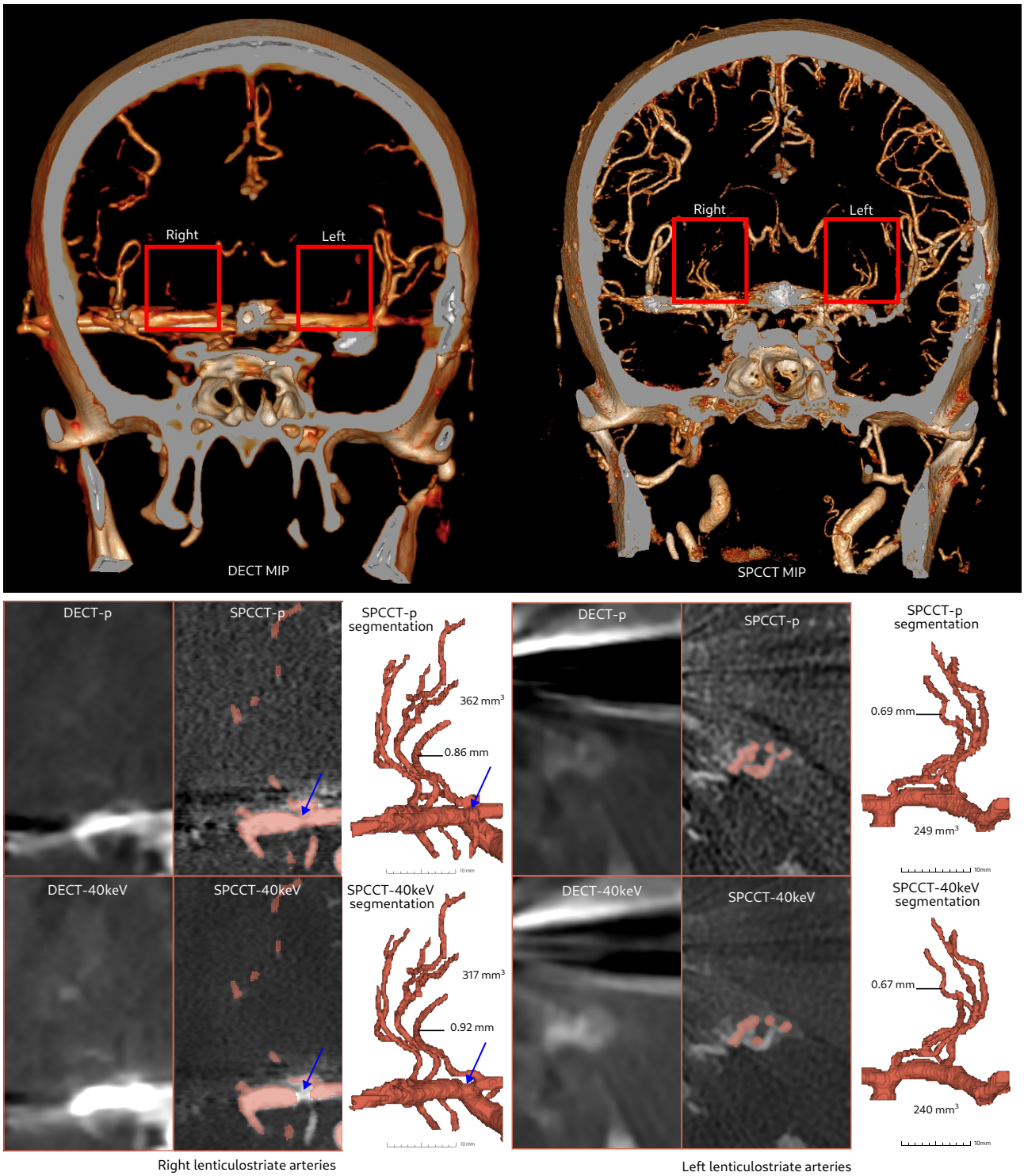


Fig. 7.4: Coronal maximum-intensity projections (MIP, 15mm slab thickness), and zoomed slices for DECT-p, DECT-40keV, SPCCT-p and SPCCT-40keV with corresponding segmentations. Arrows in blue highlights the possibility of better discrimination between arteries and veins with 40kev.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

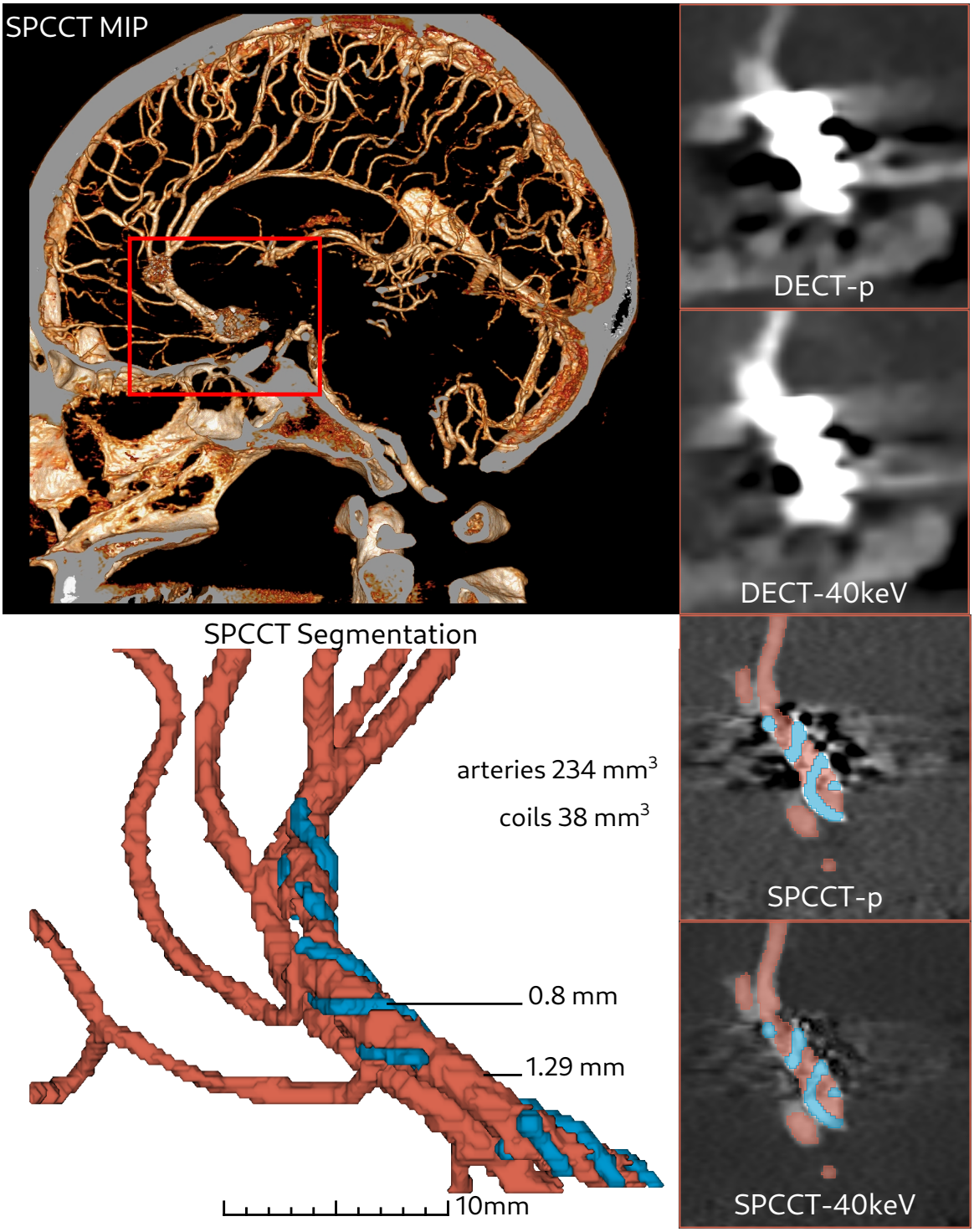


Fig. 7.5: Sagittal oblique maximum-intensity projections reconstruction (15 mm slab thickness) and zoomed slices for arterial lumen visualization with a coil in DECT-p, DECT-40keV, SPCCT-p, and SPCCT-40keV images with corresponding segmentations. Arteries are in red and coils in blue.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Image type	Volume (mm^3)	Number of artery	Mean diameter (mm)	Min diameter (mm)	Max diameter (mm)
Artery	234	19	1.10	0.74	1.46
Coil	38	9	0.78	0.65	0.86

Tab. 7.5: Artery and coil near implant.

In DECT, blooming artifacts typically render the arterial lumen adjacent to the metal indistinguishable, making post-operative evaluation difficult. In this study, SPCCT enabled a precise delineation of the arterial lumen (234 mm^3) in the immediate area of the coil (Table 7.5).

IV Conclusion

The graph-based modeling pipeline, originally designed for pulmonary arteries, was successfully applied to cerebral artery masks, confirming its adaptability to more tortuous and geometrically complex vascular structures. The quantitative analysis of all detected vessels was made possible through the vessel features extracted from the graph, providing a structured and reproducible basis for comparing DECT and SPCCT.

The results of this pilot study demonstrate that SPCCT's submillimeter spatial resolution (0.25 mm) significantly outperforms state-of-the-art DECT (0.8 mm) in neurovascular imaging. By effectively doubling the number of detected arterial segments in the Circle of Willis and allowing detection of LAs down to approximately 0.6 mm, as well as a basilar trunk perforating artery, SPCCT allows for better distinction of small arteries.

A key finding was the modality-specific effect of 40 keV reconstructions. On DECT, 40 keV increased the apparent vessel volume, likely due to increased blooming and partial-volume effects that cause vessels to merge. On SPCCT, however, 40 keV improved specificity and distal lumen conspicuity. This suggests that combining spectral binning with high spatial resolution enables high-contrast imaging without the penalty of artificial vessel enlargement.

The ability of SPCCT to reduce metal-induced blooming artifacts has immediate clinical relevance for the follow-up of patients treated with coils or stents. While DECT images were rated as unsatisfactory for post-treatment assessment, SPCCT provided clear lumen delineation near metallic implants. This suggests that SPCCT could become the preferred non-invasive modality for monitoring aneurysm occlusion and stent patency, potentially reducing the need for repeat invasive arteriography.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Conclusion

This thesis was motivated by the clinical need for automated, quantitative tools to support risk stratification of PE. The core objective was to design a complete pipeline, from CTPA acquisitions through image segmentation and graph-based biomarker extraction, capable of characterizing the PAT and the thrombotic burden. This required addressing three interconnected challenges: producing reliable ground-truth annotations, training robust deep-learning segmentation models, and building an anatomically consistent graph representation of the PAT.

The contributions of this thesis can be summarized as follows.

First, we addressed the lack of ground-truth annotations in the PERSEVERE dataset by developing VSMOD, an open-source plugin for semi-automated vascular segmentation (Chapter 4). VSMOD reduces annotation time by 80% compared to fully manual approaches, produces topologically connected segmentation masks, and directly generates a graph representation of the vascular network. Building on these properties, a sample of high-quality PAT annotations was produced from the PERSEVERE cohort, serving as the foundation for all subsequent contributions. This work resulted in a publication in the Journal of Open Source Software (JOSS).

Second, we developed and evaluated deep learning models for automated pulmonary artery and thrombi segmentation, trained on the high-quality annotations generated with VSMOD. These models demonstrated strong performance on the PERSEVERE cohort and enabled the large-scale, automated extraction of vascular and thrombotic information (Chapter 5).

Third, we proposed a complete graph-based pipeline for modeling the PAT, including centerline extraction, graph construction, hierarchical labeling, and biomarker computation (Chapter 6). This pipeline provided a structured and quantitative representation of the vascular tree and thrombotic burden. This work was submitted to the Journal of Biomedical and Health Informatics (JBHI).

Fourth, we demonstrated the transferability of this graph modeling framework to a different anatomical territory, namely, neurovascular imaging. The pipeline was applied to cerebral CT angiography to compare segmentation quality between DECT and SPCCT acquisitions in the presence of metallic implants, illustrating its adaptability beyond the pulmonary context. This work resulted in a publication in the Journal of Neurobiology.

Despite these contributions, several limitations must be acknowledged.

Regarding annotations, VSMOD was primarily designed for PAT segmentation in CTPA. While a preliminary transfer to cerebrovascular imaging demonstrated that the graph-based region-growing strategy is not inherently limited to the pulmonary territory, the current parameterization remains optimized for pulmonary artery geometry, characterized by relatively large calibers and low tortuosity. Manual tuning is still required for vascular structures with different prop-

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

erties. In addition, the automatic seed generation in the segmentation module relies on local radius values estimated by the RANSAC algorithm. In regions of low contrast or pathological deformation, these radius estimates can be inaccurate, leading to erroneous seed initialization that requires manual correction by the user.

Regarding segmentation, the models trained on the PERSEVERE cohort showed limited transferability to external datasets such as PARSE (Chapter 5). This limitation arises from both acquisition-related differences, including scanner manufacturer and reconstruction protocols, and pathology-related discrepancies, as the PERSEVERE cohort consists exclusively of PE patients with altered vascular morphology. As a result, the models may learn representations tied to specific pathological patterns and acquisition conditions, limiting their generalization. Furthermore, publicly available annotated datasets either lack pathological cases or remain limited in size, and attempts to leverage semi-supervised strategies such as Mean Teacher did not yield satisfactory improvements. For thrombi segmentation, the variability in the appearance of hypodense filling defects remains a challenge. Although datasets such as FUMPE (Saadatmand-Tarzjan 2018) and CAD-PE (González et al. 2020) provide annotated thrombi, their limited size restricts their impact on model generalization.

Regarding graph modeling, the pipeline still encounters difficulties in the presence of severe pathological deformations. These difficulties can originate from two distinct sources: upstream segmentation errors, such as merged vessel walls or spurious connections introduced by the deep learning model, and limitations of the graph construction and correction algorithms themselves, which may fail to correctly resolve ambiguous topological configurations in heavily deformed vascular trees. Both sources contribute to residual cycles or ambiguous edges that are not fully resolved by the current correction procedure.

The neurovascular application remains a proof of concept based on a single patient, which demonstrates the feasibility of transferring the graph-based pipeline to a different anatomical territory but does not constitute a validation of its generalizability to broader neurovascular cohorts. Furthermore, the current pipeline is designed for vascular networks with a tree-like topology, treating all cycles as artifacts. This assumption holds for the PAT but limits the framework's applicability to vascular networks with different topologies, such as the cerebrovascular network, which contains the anatomically valid cycle of the circle of Willis. Extending the modeling framework to such structures would require dedicated methods for distinguishing true anatomical cycles from segmentation artifacts.

These limitations open several directions for future work.

From an annotation perspective, two complementary directions can be envisioned. First, VSMOD should be evaluated on vascular structures beyond the pulmonary territory, in order to assess its generalizability and establish organ-specific parameter sets adapted to the geometric properties of each structure. Second, the region-growing segmentation module could be replaced or augmented by a deep learning model, which would improve robustness in pathological cases and low-quality images where radius estimation and seed generation are currently unreliable.

From a segmentation standpoint, improving generalization across acquisition protocols and

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

patient populations will require both access to multi-center datasets with diverse acquisition protocols and the development of training strategies that explicitly address domain and concept shifts. For pulmonary artery segmentation specifically, a promising direction is a two-stage approach that combines pre-training on large healthy datasets such as AirRC (J. Liu et al. 2025), which provides diverse pulmonary vascular annotations across multiple centers, with fine-tuning on pathological cases from the PERSEVERE cohort. Recent general-purpose segmentation tools, such as the updated version of TotalSegmentator, also offer a promising basis for transfer and fine-tuning in this context. Looking further ahead, vascular foundation models (Wang, Meng, and Z. Xu 2025) trained across multiple anatomical territories and imaging modalities could reduce the need for large annotated datasets.

From a modeling perspective, moving toward learning-based, topology-aware approaches represents a promising direction. Recent work, such as the Implicit Point-Graph Network (Xie, Yang, et al. 2025; Xie, Y. Zhu, et al. 2026) proposes to jointly model segmentation and topology, enabling more robust handling of disconnected branches and anatomically inconsistent structures. Such approaches could improve graph construction and enable finer anatomical labeling at the pulmonary segment name level. Extending the pipeline to vascular networks with cyclic structures will also require the development of methods capable of distinguishing anatomically valid cycles from artifacts, for instance, by combining anatomical priors with geometric criteria.

From a clinical perspective, the automated graph-based representation of the PAT opens the way for large-scale prognostic studies. The extracted biomarkers, including thrombotic burden scores and embolic volume spatial distributions, provide a rich description of PE patient PAT that could improve risk stratification beyond existing clinical scores. In particular, the spatial organization of clots may help differentiate patients with similar total embolic volumes but different clinical outcomes.

More general perspectives arise from the joint availability of graph-based representations and clinical data.

The availability of clinical biomarkers such as BNP and troponin in the PERSEVERE dataset further enables the investigation of correlations between graph-derived features and patient outcomes. Establishing whether these biomarkers improve the prediction of RV dysfunction, clinical deterioration, or treatment response represents an important next step.

Finally, the graph-based representation is naturally adapted to the use of Graph Neural Networks (GNNs) for outcome prediction. The cohort of 353 patient graphs constitutes a structured dataset in which edges encode vessel, location and PE features. GNNs can exploit this structure to capture interactions across the vascular tree and learn predictive patterns that go beyond aggregated scalar biomarkers. Preliminary work in related domains (Paetzold et al. 2021; Vaida and Z. Huang 2026) suggests the potential of this approach.

Overall, this thesis establishes a complete and automated framework for graph-based analysis of the PAT in PE, from annotation to clinical biomarker extraction. By providing a structured and quantitative representation of thrombotic burden, it lays the groundwork for future developments in both methodological advances and clinical decision support, and opens the way

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

toward more personalized and topology-aware approaches to vascular disease analysis.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Publication list

Journal articles

Morgane des Ligneris, Nathan Painchaud, Allan Serva, Laurent Bertoletti, Pierre Croisille, Carole Frindel, Odyssée Merveille. A Patient-Specific Pulmonary Arterial Tree Digital Twin to Extract Pulmonary Embolism Biomarkers. IEEE Journal of Biomedical and Health Informatics. Special Issue on "Integrating Imaging with Personalized Cardiovascular Medicine: Current Advances and Future Directions". *Submitted* (Chapter 6)

Morgane des Ligneris, Odyssée Merveille, Alexandre Bani-Sadr, Pierre Rougé, Carole Frindel, Salim Si-Mohamed. Cerebral Spectral Photon-counting CT angiography, an alternative to invasive cerebral angiography. Journal of Neuroradiology. *Submitted* (Chapter 7)

Morgane des Ligneris, Gabriel Husak, Azéline Aillet, Guillaume Dollé, Carole Frindel, Odyssée Merveille. VSMOD: A Vessel Segmentation and MODelization plugin for 3D Slicer. Journal of Open Source Software, 2025, 10 (114), pp.8563. 10.21105/joss.08563 . hal-05312036 (Chapter 4)

Méghane Decroocq, **Morgane Des Ligneris**, Titouan Poquillon, Maxime Vincent, Manon Aubert, Timothée Jacquesson, Carole Frindel. Automation of Cranial Nerve Tractography by Filtering Tractograms for Skull Base Surgery. Frontiers in Neuroimaging, 2022, 1, pp.838483. 10.3389/fnimg.2022.838483 hal-03771217

International Conference

Morgane des Ligneris, Nathan Painchaud, Allan Serva, Laurent Bertoletti, Pierre Croisille, Carole Frindel, Odyssée Merveille. VPH 2026 : International Conference on the Virtual Physiological Human, Septembre 2026. *Submitted* (Chapter 6)

Morgane Des Ligneris, Axel Bonnet, Yohan Chatelain, Tristan Glatard, Michaël Sdika, Gaël Vila, Valentine Wargnier-Dauchelle, Sorina Camarasu-Pop, Carole Frindel. Reproducibility of Tumor Segmentation Outcomes with a Deep Learning Model. International Symposium on Biomedical Imaging (ISBI), Apr 2023, Cartagena de Indias, Colombia. 10.1109/ISBI53787.2023.10230482 . hal-04006057 *Poster*

Méghane Decroocq, **Morgane Des Ligneris**, Timothée Jacquesson, Carole Frindel. Viewpoint Selection for Fibrous Structures in a Pre-operative Context: Application to Cranial Nerves Surrounding Skull Base Tumors. International Conference on Systems, Signals and Image Processing IWSSIP 2021, Jun 2021, Bratislava, Slovakia. pp.53-64, 10.1007/978-3-030-96878-6_5 . hal-03772158 *Conference paper*

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

National conference

Morgane des Ligneris, Nathan Painchaud, Allan Serva, Laurent Bertoletti, Pierre Croisille, Carole Frindel, Odyssée Merveille. From automated CTPA arterial segmentation to graph-based pulmonary arterial tree digital twin for pulmonary embolism biomarker extraction. IABM 2026 : Colloque Français d'Intelligence Artificielle en Imagerie Biomédicale, Mars 2026. *Poster* (Chapter 6)

Morgane des Ligneris, Gabriel Jacquinot, Azéline Aillet, Carole Frindel and Odyssée Merveille. PulmonaryArteriesSegmentor : Plugin 3D Slicer d'annotation et de segmentation semi-automatique des artères pulmonaires. IABM 2024 : Colloque Français d'Intelligence Artificielle en Imagerie Biomédicale, Mars 2024. *Poster* (Chapter 4)

Communication

Morgane des Ligneris, Carole Frindel, Odyssée Merveille. Pipeline d'extraction automatique de biomarqueurs pour la caractérisation des embolies pulmonaires. GDR IASIS : Journée thématique Imagerie Vasculaire, 12 mai 2025. *Oral* (Chapter 6)

Gaël Vila, Axel Bonnet, Carole Frindel, Helene Ratiney, Frédéric Cervenansky, **Morgane Des Ligneris**, C. Mouton, Emmanuel Medernach, Jerome Pansanel, Tristan Glatard, Sorina Camarasu-Pop. Projet ReproVIP. Journée Reproductibilité Labex PRIMES, Dec 2022, Lyon, France. hal-04006460 . *Oral*

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Softwares and libraries

`Chest Imaging Platform` 3D Slicer extension, a toolkit for lung specific image-based algorithms. https://www.chestimagingplatform.org/.

`Draw Tube` Tool to draw a vessel by interactively placing tubes from the extension SlicerSegmentEditorExtraEffects available for 3D Slicer.
https://github.com/lassoan/SlicerSegmentEditorExtraEffects.

`Grow from Seeds` Multi-label region growing (L. Zhu et al. 2014) available in 3D Slicer.

`Lungmask` Deep learning model that segments the lungs and lobes (Hofmanninger et al. 2020). https://github.com/Slicer/SlicerLungCTAnalyzer/. Available on 3D Slicer.

`RVesselX` (Lamy et al. 2022) proposes a semi-automatic framework to help annotate liver vessels, available in 3D Slicer. https://github.com/R-Vessel-X/SlicerRVXLiverSegmentation.

`TotalSegmentator` Deep learning model that segments more than one hundred anatomical structures on CT scans (Wasserthal et al. 2023). Available on 3D Slicer.

`VesselGraph` (Drees et al. 2021) is graph modeling pipeline from the Voreen (Meyer-Spradow et al. 2009) Software. https://www.uni-muenster.de/Voreen/index.html.

`skancsr` Submodule of `scan`, a Python library to analyze skeleton images with a Compressed Sparse Row (CSR) format used for graph representation (Nunez-Iglesias et al. 2018). https://scikit-image.org/docs/stable/api/skimage.morphology.html.

`skimage.morphology` Submodule of `scikit-image`, a Python library, that provides morphological image processing operations.
https://scikit-image.org/docs/stable/api/skimage.morphology.html.

3D Slicer Open source software to visualize and work on 3D images. Version 5.8.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Bibliography

Alis, Jonathan and Linda B. Haramati (2023). "CT in Pulmonary Embolism". In: *PERT Consortium Handbook of Pulmonary Embolism : Research, Care, and Management*. Ed. by Charles J. Grodzin et al. Cham: Springer International Publishing, pp. 1–22. ISBN: 978-3-030-70904-4. DOI: 10.1007/978-3-030-70904-4_64-1. URL: https://doi.org/10.1007/978-3-030-70904-4_64-1 (cited on pp. 18, 19, 98).

An, Junho et al. (Feb. 2022). "Acute Pulmonary Embolism and Chronic Thromboembolic Pulmonary Hypertension: Clinical and Serial CT Pulmonary Angiographic Features". In: *Journal of Korean Medical Science* 37.10. ISSN: 1598-6357. DOI: 10.3346/jkms.2022.37.e76 (cited on pp. 22, 142).

Antiga, Luca et al. (Nov. 2008). "An Image-Based Modeling Framework for Patient-Specific Computational Hemodynamics". In: *Medical & Biological Engineering & Computing* 46.11, pp. 1097–1112. ISSN: 1741-0444. DOI: 10.1007/s11517-008-0420-1 (cited on p. 52).

Aujesky, Drahomir et al. (Oct. 2005). "Derivation and Validation of a Prognostic Model for Pulmonary Embolism". In: *American Journal of Respiratory and Critical Care Medicine* 172.8, pp. 1041–1046. ISSN: 1073-449X, 1535-4970. DOI: 10.1164/rccm.200506-862OC (cited on pp. 15, 141).

Avants, Brian, Nick Tustison, and Gang Song (Nov. 2008). "Advanced normalization tools (ANTS)". In: *Insight J* 1–35. DOI: 10.54294/uvnhin (cited on p. 113).

Bankier, Alexander A. et al. (Apr. 1997). "Severity Assessment of Acute Pulmonary Embolism with Spiral CT: Evaluation of Two Modified Angiographic Scores and Comparison with Clinical Data". In: *Journal of Thoracic Imaging* 12.2, pp. 150–158. ISSN: 0883-5993. DOI: 10.1097/00005382-199704000-00012 (cited on pp. 17, 142).

Carneiro-Esteves, Sophie, Antoine Vacavant, and Odyssée Merveille (2025). "Restoring Connectivity in Vascular Segmentations Using a Learned Post-processing Model". In: *Topology- and Graph-Informed Imaging Informatics*. Ed. by Chao Chen, Yash Singh, and Xiaoling Hu. Cham: Springer Nature Switzerland, pp. 55–65. ISBN: 978-3-031-73967-5. DOI: 10.1007/978-3-031-73967-5_6 (cited on pp. 50, 58, 144).

Casey, Scott D. et al. (May 2023). "Letter to the Editor: "Clinical Controversies in the Management of Acute Pulmonary Embolism"". In: *Expert Review of Respiratory Medicine* 17.6, pp. 425–426. ISSN: 1747-6348. DOI: 10.1080/17476348.2023.2214733 (cited on pp. 20, 21).

Chu, Yuetan et al. (Mar. 2025). "Deep Learning-Driven Pulmonary Artery and Vein Segmentation Reveals Demography-Associated Vasculature Anatomical Differences". In: *Nature Communications* 16.1, p. 2262. ISSN: 2041-1723. DOI: 10.1038/s41467-025-56505-6 (cited on p. 72).

Çiçek, Özgün et al. (2016). "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. Ed. by Sebastien Ourselin et al. Cham: Springer International Publishing, pp. 424–

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

432. ISBN: 978-3-319-46723-8. DOI: 10.1007/978-3-319-46723-8_49 (cited on pp. 43, 144).

Cui, Hejie, Xinglong Liu, and Ning Huang (2019). "Pulmonary Vessel Segmentation Based on Orthogonal Fused U-Net++ of Chest CT Images". In: vol. 11769, pp. 293–300. DOI: 10.1007/978-3-030-32226-7_33. arXiv: 2107.01502 [cs, eess] (cited on p. 48).

Decroocq, Méghane et al. (Oct. 2023). "Modeling and Hexahedral Meshing of Cerebral Arterial Networks from Centerlines". In: *Medical Image Analysis* 89, p. 102912. ISSN: 1361-8415. DOI: 10.1016/j.media.2023.102912 (cited on pp. 54, 145).

Des Ligneris, Morgane et al. (Oct. 10, 2025). "VSMOD: A Vessel Segmentation and MODelization Plugin for 3D Slicer". In: *Journal of Open Source Software* 10.114, p. 8563. ISSN: 2475-9066. DOI: 10.21105/joss.08563. URL: https://joss.theoj.org/papers/10.21105/joss.08563 (cited on pp. 57, 146).

Drees, Dominik et al. (June 2021). "Scalable Robust Graph and Feature Extraction for Arbitrary Vessel Networks in Large Volumetric Datasets". In: *BMC Bioinformatics* 22.1, p. 346. ISSN: 1471-2105. DOI: 10.1186/s12859-021-04262-w (cited on pp. 53, 89, 129).

Fourdrain, Alex et al. (Jan. 2018). "Three-dimensional CT angiography of anatomic variations in the pulmonary arterial tree". In: *Surgical and Radiologic Anatomy* 40. DOI: 10.1007/s00276-017-1914-z (cited on pp. 7, 8, 141).

Furlan, Alessandro et al. (Mar. 2011). "Accuracy and Reproducibility of Blood Clot Burden Quantification With Pulmonary CT Angiography". In: *American Journal of Roentgenology* 196.3, pp. 516–523. ISSN: 0361-803X. DOI: 10.2214/AJR.10.4603 (cited on pp. 20, 142).

González, Germán et al. (Mar. 2020). *Computer Aided Detection for Pulmonary Embolism Challenge (CAD-PE)*. DOI: 10.48550/arXiv.2003.13440. arXiv: 2003.13440 [cs, eess] (cited on p. 124).

Greffier, Joël, Nicolas Villani, et al. (Apr. 2023). "Spectral CT Imaging: Technical Principles of Dual-Energy CT and Multi-Energy Photon-Counting CT". In: *Diagnostic and Interventional Imaging* 104.4, pp. 167–177. ISSN: 2211-5684. DOI: 10.1016/j.diii.2022.11.003 (cited on pp. 112, 151).

Greffier, Joël, Anaïs Viry, et al. (Feb. 2025). "Photon-Counting CT Systems: A Technical Review of Current Clinical Possibilities". In: *Diagnostic and Interventional Imaging* 106.2, pp. 53–59. ISSN: 2211-5684. DOI: 10.1016/j.diii.2024.09.002 (cited on p. 112).

Hariharan, Praveen et al. (Nov. 2016). "Relation Among Clot Burden, Right-Sided Heart Strain, and Adverse Events After Acute Pulmonary Embolism". In: *The American Journal of Cardiology* 118.10, pp. 1568–1573. ISSN: 1879-1913. DOI: 10.1016/j.amjcard.2016.08.025 (cited on p. 21).

Hofmanninger, Johannes et al. (Aug. 2020). "Automatic Lung Segmentation in Routine Imaging Is Primarily a Data Diversity Problem, Not a Methodology Problem". In: *European Radiology Experimental* 4.1, p. 50. ISSN: 2509-9280. DOI: 10.1186/s41747-020-00173-2 (cited on pp. 47, 129, 144).

Hu, Xiaoling et al. (June 2019). *Topology-Preserving Deep Image Segmentation*. DOI: 10.48550/arXiv.1906.05404. arXiv: 1906.05404 [cs] (cited on pp. 50, 144).

Huang, Wei-Ming et al. (May 2022). "Quantitative Volumetric Computed Tomography Embolic Analysis, the Qanadli Score, Biomarkers, and Clinical Prognosis in Patients with Acute Pulmonary Embolism". In: *Scientific Reports* 12.1, p. 7620. ISSN: 2045-2322. DOI: 10.1038/s41598-022-11812-6 (cited on pp. 20, 21, 52, 142, 145).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Isensee, Fabian et al. (Feb. 2021). "nnU-Net: A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation". In: *Nature Methods* 18.2, pp. 203–211. ISSN: 1548-7105. DOI: 10.1038/s41592-020-01008-z (cited on pp. 45, 46, 144).

Jiménez, David, Drahomir Aujesky, et al. (Aug. 2010). "Simplification of the Pulmonary Embolism Severity Index for Prognostication in Patients with Acute Symptomatic Pulmonary Embolism". In: *Archives of Internal Medicine* 170.15, pp. 1383–1389. ISSN: 1538-3679. DOI: 10.1001/archinternmed.2010.199 (cited on pp. 15, 141).

Jiménez, David, Javier de Miguel-Díez, et al. (Jan. 19, 2016). "Trends in the Management and Outcomes of Acute Pulmonary Embolism: Analysis From the RIETE Registry". In: *Journal of the American College of Cardiology* 67.2, pp. 162–170. ISSN: 0735-1097. DOI: 10.1016/j.jacc.2015.10.060. URL: https://www.sciencedirect.com/science/article/pii/S0735109715073489 (cited on pp. 13, 141).

Keshwani, Deepak et al. (2020). "TopNet: Topology Preserving Metric Learning for Vessel Tree Reconstruction and Labelling". In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. Ed. by Anne L. Martel et al. Cham: Springer International Publishing, pp. 14–23. ISBN: 978-3-030-59725-2. DOI: 10.1007/978-3-030-59725-2_2 (cited on pp. 49, 144).

Konstantinides, Stavros V et al. (Jan. 2020). "2019 ESC Guidelines for the Diagnosis and Management of Acute Pulmonary Embolism Developed in Collaboration with the European Respiratory Society (ERS)". In: *European Heart Journal* 41.4, pp. 543–603. ISSN: 0195-668X, 1522-9645. DOI: 10.1093/eurheartj/ehz405 (cited on pp. 12, 14, 31, 141).

Lamy, Jonas et al. (May 2022). "The 3D Slicer RVXLiverSegmentation Plug-in for Interactive Liver Anatomy Reconstruction from Medical Images". In: *Journal of Open Source Software* 7.73, p. 3920. ISSN: 2475-9066. DOI: 10.21105/joss.03920 (cited on pp. 54, 64, 129, 145, 146).

Lee, T. C., R. L. Kashyap, and C. N. Chu (Nov. 1, 1994). "Building Skeleton Models via 3-D Medial Surface Axis Thinning Algorithms". In: *CVGIP: Graphical Models and Image Processing* 56.6, pp. 462–478. ISSN: 1049-9652. DOI: 10.1006/cgip.1994.1042. URL: https://www.sciencedirect.com/science/article/pii/S104996528471042X (cited on p. 52).

Lekadir, Karim et al. (Feb. 2025). "FUTURE-AI: International Consensus Guideline for Trustworthy and Deployable Artificial Intelligence in Healthcare". In: *BMJ* 388, e081554. ISSN: 1756-1833. DOI: 10.1136/bmj-2024-081554 (cited on pp. 49, 144).

Li, Qingya et al. (2026). "Graph-PAVNet: A Graph-Based Learning Framework for Pulmonary Artery and Vein Separation Using Multimodal Feature Sampling". In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2025*. Ed. by James C. Gee et al. Cham: Springer Nature Switzerland, pp. 373–382. ISBN: 978-3-032-05162-2. DOI: 10.1007/978-3-032-05162-2_36 (cited on pp. 54, 145).

Liu, Jian et al. (Nov. 2025). "A Custom Annotated Dataset for Segmentation of Pulmonary Veins, Arteries, and Airways". In: *Scientific Data* 12.1, p. 1806. ISSN: 2052-4463. DOI: 10.1038/s41597-025-06074-6 (cited on pp. 125, 155).

Liu, Weifang et al. (June 2020). "Evaluation of Acute Pulmonary Embolism and Clot Burden on CTPA with Deep Learning". In: *European Radiology* 30.6, pp. 3567–3575. ISSN: 1432-1084. DOI: 10.1007/s00330-020-06699-8 (cited on pp. 52, 145).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Long, Kun et al. (Jan. 2021). "Probability-Based Mask R-CNN for Pulmonary Embolism Detection". In: *Neurocomputing* 422, pp. 345–353. ISSN: 0925-2312. DOI: 10.1016/j.neucom.2020.10.022 (cited on pp. 51, 144).

Luijten, Dieuwke et al. (Mar. 2023a). "Clinical Controversies in the Management of Acute Pulmonary Embolism: Evaluation of Four Important but Controversial Aspects of Acute Pulmonary Embolism Management That Are Still Subject of Debate and Research". In: *Expert Review of Respiratory Medicine* 17.3, pp. 181–189. ISSN: 1747-6348. DOI: 10.1080/17476348.2023.2190888 (cited on p. 20).

— (June 2023b). "Response to Letter to the Editor: "Clinical Controversies in the Management of Acute Pulmonary Embolism: Evaluation of Four Important but Controversial Aspects of Acute Pulmonary Embolism Management That Are Still Subject of Debate and Research"". In: *Expert Review of Respiratory Medicine* 17.6, pp. 427–428. ISSN: 1747-6348. DOI: 10.1080/17476348.2023.2215015 (cited on p. 21).

Luo, Gongning et al. (Aug. 2024). *Efficient Automatic Segmentation for Multi-Level Pulmonary Arteries: The PARSE Challenge*. DOI: 10.48550/arXiv.2304.03708. arXiv: 2304.03708 [eess] (cited on pp. 33, 72, 143).

Ma, Xiaochao et al. (Apr. 2023). "Analysis of Pulmonary Artery Variation Based on 3D Reconstruction of CT Angiography". In: *Frontiers in Physiology* 14. ISSN: 1664-042X. DOI: 10.3389/fphys.2023.1156513 (cited on pp. 7, 8, 141).

Mastora, Ioana et al. (Jan. 2003). "Severity of Acute Pulmonary Embolism: Evaluation of a New Spiral CT Angiographic Score in Correlation with Echocardiographic Data". In: *European Radiology* 13.1, pp. 29–35. ISSN: 1432-1084. DOI: 10.1007/s00330-002-1515-y (cited on pp. 17, 97, 142).

Mehdipoor, Ghazaleh et al. (May 2020). "Patient-Level, Institutional, and Temporal Variations in Use of Imaging Modalities to Confirm Pulmonary Embolism". In: *Circulation: Cardiovascular Imaging* 13.5, e010651. DOI: 10.1161/CIRCIMAGING.120.010651. URL: https://www.ahajournals.org/doi/10.1161/CIRCIMAGING.120.010651 (cited on pp. 13, 141).

Meinel, Felix G. et al. (July 2015). "Predictive Value of Computed Tomography in Acute Pulmonary Embolism: Systematic Review and Meta-analysis". In: *The American Journal of Medicine* 128.7, 747–759.e2. ISSN: 0002-9343, 1555-7162. DOI: 10.1016/j.amjmed.2015.01.023 (cited on pp. 13, 21).

Meyer-Spradow, J. et al. (2009). "Voreen: A Rapid-Prototyping Environment for Ray-Casting-Based Volume Visualizations". In: *IEEE computer graphics and applications* 29.6, pp. 6–13. ISSN: 0272-1716. DOI: 10.1109/MCG.2009.130 (cited on pp. 53, 89, 129).

Michaud, Erinleigh et al. (Sept. 2023). "Anatomical Variations in Pulmonary Arterial Branches in Patients Undergoing Evaluation for Chronic Thromboembolic Pulmonary Hypertension". In: *Journal of the Society for Cardiovascular Angiography & Interventions* 2.5. ISSN: 2772-9303. DOI: 10.1016/j.jscai.2023.101108 (cited on pp. 7–9, 94, 141).

Miller, G. A. H. et al. (June 1971). "Comparison of Streptokinase and Heparin in Treatment of Isolated Acute Massive Pulmonary Embolism". In: *British Medical Journal* 2.5763, pp. 681–684. ISSN: 0007-1447 (cited on pp. 16, 141).

Miller, Megan et al. (Oct. 2023). "A Topological Data Analysis Study on Murine Pulmonary Arterial Trees with Pulmonary Hypertension". In: *Mathematical Biosciences* 364, p. 109056. ISSN: 0025-5564. DOI: 10.1016/j.mbs.2023.109056 (cited on p. 50).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Mishra, Suraj et al. (Nov. 2021). "VTG-Net: A CNN Based Vessel Topology Graph Network for Retinal Artery/Vein Classification". In: *Frontiers in Medicine* 8. ISSN: 2296-858X. DOI: 10.3389/fmed.2021.750396 (cited on pp. 54, 145).

Nakada, Kei et al. (Jan. 2010). "Relation between Pulmonary Embolus Volume Quantified by Multidetector Computed Tomography and Clinical Status and Outcome for Patients with Acute Pulmonary Embolism". In: *Japanese Journal of Radiology* 28.1, pp. 34–42. ISSN: 1862-5274. DOI: 10.1007/s11604-009-0380-x (cited on pp. 20, 142).

Nam, Ju Gang et al. (Dec. 2021). "Automatic Pulmonary Vessel Segmentation on Noncontrast Chest CT: Deep Learning Algorithm Developed Using Spatiotemporally Matched Virtual Noncontrast Images and Low-keV Contrast-Enhanced Vessel Maps". In: *European Radiology* 31.12, pp. 9012–9021. ISSN: 1432-1084. DOI: 10.1007/s00330-021-08036-z (cited on p. 48).

Nunez-Iglesias, Juan et al. (Feb. 15, 2018). "A New Python Library to Analyse Skeleton Images Confirms Malaria Parasite Remodelling of the Red Blood Cell Membrane Skeleton". In: *PeerJ* 6, e4312. ISSN: 2167-8359. DOI: 10.7717/peerj.4312. PMID: 29472997. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC5816961/ (cited on pp. 53, 88, 129).

Paetzold, Johannes C. et al. (Aug. 2021). *Whole Brain Vessel Graphs: A Dataset and Benchmark for Graph Learning and Neuroscience (VesselGraph)*. DOI: 10.5281/zenodo.5301621. arXiv: 2108.13233 [cs] (cited on pp. 125, 156).

Peters, Ralph et al. (Mar. 2007). "Labeling the Pulmonary Arterial Tree in CT Images for Automatic Quantification of Pulmonary Embolism". In: *Proceedings of SPIE, 2007 vol. 6514* 6514. DOI: 10.1117/12.705840 (cited on pp. 51, 145).

Pu, Jiantao et al. (Oct. 2023). "Automated Detection and Segmentation of Pulmonary Embolisms on Computed Tomography Pulmonary Angiography (CTPA) Using Deep Learning but without Manual Outlining". In: *Medical Image Analysis* 89, p. 102882. ISSN: 1361-8415. DOI: 10.1016/j.media.2023.102882 (cited on pp. 51, 144).

Qanadli, Salah D. et al. (June 2001). "New CT Index to Quantify Arterial Obstruction in Pulmonary Embolism". In: *American Journal of Roentgenology* 176.6, pp. 1415–1420. ISSN: 0361-803X. DOI: 10.2214/ajr.176.6.1761415 (cited on pp. 17, 97, 142).

Rai, Ansaar T. et al. (July 2013). "Cerebrovascular Geometry in the Anterior Circulation: An Analysis of Diameter, Length and the Vessel Taper". In: *Journal of Neurointerventional Surgery* 5.4, pp. 371–375. ISSN: 1759-8486. DOI: 10.1136/neurintsurg-2012-010314 (cited on pp. 11, 141).

Raju, Sreenivasa Narayana et al. (Apr. 2021). "Pulmonary Arterial Dilatation: Imaging Evaluation Using Multidetector Computed Tomography". In: *The Indian Journal of Radiology & Imaging* 31.2, pp. 409–420. ISSN: 0971-3026. DOI: 10.1055/s-0041-1734225 (cited on p. 9).

Reichold, Johannes et al. (Aug. 2009). "Vascular Graph Model to Simulate the Cerebral Blood Flow in Realistic Vascular Networks". In: *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism* 29.8, pp. 1429–1443. ISSN: 1559-7016. DOI: 10.1038/jcbfm.2009.58 (cited on pp. 54, 145).

Reymond, Philippe et al. (July 2009). "Validation of a One-Dimensional Model of the Systemic Arterial Tree". In: *American Journal of Physiology. Heart and Circulatory Physiology* 297.1, H208–222. ISSN: 1522-1539. DOI: 10.1152/ajpheart.00037.2009 (cited on pp. 54, 145).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Ronneberger, Olaf, Philipp Fischer, and Thomas Brox (May 2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. DOI: 10.48550/arXiv.1505.04597. arXiv: 1505.04597 [cs] (cited on pp. 42, 144).

Rotzinger, David C. et al. (Aug. 2020). "CT Pulmonary Angiography for Risk Stratification of Patients with Nonmassive Acute Pulmonary Embolism". In: *Radiology: Cardiothoracic Imaging* 2.4, e190188. ISSN: 2638-6135. DOI: 10.1148/ryct.2020190188 (cited on pp. 21, 22, 142).

Rougé, Pierre, Nicolas Passat, and Odyssée Merveille (Dec. 2024). "Topology aware multitask cascaded U-Net for cerebrovascular segmentation". In: *PLOS ONE* 19.12, pp. 1–20. DOI: 10.1371/journal.pone.0311439. URL: https://doi.org/10.1371/journal.pone.0311439 (cited on pp. 43, 45, 58, 115, 144, 152).

Saadatmand-Tarzjan, Mahdi (May 2018). "FUMPE". In: DOI: 10.6084/m9.figshare.c.4107803.v1 (cited on p. 124).

Shi, Yujie et al. (2025). "Application and Challenges of Deep Learning in Pulmonary Vessels Segmentation of CTPA Images". In: *IEEE Journal of Biomedical and Health Informatics*, pp. 1–13. ISSN: 2168-2208. DOI: 10.1109/JBHI.2025.3638808 (cited on pp. 42, 48).

Shit, Suprosanna et al. (June 2021). "clDice – A Novel Topology-Preserving Loss Function for Tubular Structure Segmentation". In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16555–16564. DOI: 10.1109/CVPR46437.2021.01629. arXiv: 2003.07311 [cs, eess] (cited on pp. 37, 58, 143, 144).

Smith, Stephen M. et al. (Jan. 2004). "Advances in Functional and Structural MR Image Analysis and Implementation as FSL". In: *NeuroImage*. Mathematics in Brain Imaging 23, S208–S219. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2004.07.051 (cited on p. 164).

Sofi, Shakir Showkat and Nadezhda Alsahanova (Sept. 2023). "Image Segmentation with Topological Priors". In: *2023 IEEE High Performance Extreme Computing Conference (HPEC)*, pp. 1–6. DOI: 10.1109/HPEC58863.2023.10363528. arXiv: 2205.06197 [cs] (cited on pp. 50, 144).

Stenseke, Jakob (2021). "Persistent homology and the shape of evolutionary games". In: *Journal of Theoretical Biology* 531, p. 110903. ISSN: 0022-5193. DOI: https://doi.org/10.1016/j.jtbi.2021.110903. URL: https://www.sciencedirect.com/science/article/pii/S0022519321003222 (cited on p. 38).

Tajeri, Taraneh et al. (Feb. 2024). "The Correlation between the CT Angiographic Pulmonary Artery Obstructive Index and Clinical Data in Patients with Acute Pulmonary Thromboembolism". In: *Emergency Radiology* 31.1, pp. 45–51. ISSN: 1438-1435. DOI: 10.1007/s10140-023-02187-w (cited on pp. 21, 22, 142).

Tarvainen, Antti and Harri Valpola (Apr. 2018). *Mean Teachers Are Better Role Models: Weight-averaged Consistency Targets Improve Semi-Supervised Deep Learning Results*. arXiv: 1703.01780 [cs, stat] (cited on pp. 44, 144).

Tromeur, Cécile et al. (Jan. 2018). "Risk Factors for Recurrent Venous Thromboembolism after Unprovoked Pulmonary Embolism: The PADIS-PE Randomised Trial". In: *European Respiratory Journal* 51.1. ISSN: 0903-1936, 1399-3003. DOI: 10.1183/13993003.01202-2017 (cited on pp. 22, 142).

Truong, Quynh A. et al. (Jan. 2012). "Reference Values for Normal Pulmonary Artery Dimensions by Noncontrast Cardiac Computed Tomography". In: *Circulation: Cardiovascular Imaging* 5.1, pp. 147–154. DOI: 10.1161/CIRCIMAGING.111.968610 (cited on p. 6).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Vaida, Maria and Ziyuan Huang (Jan. 2026). "Multimodal Graph Neural Networks in Healthcare: A Review of Fusion Strategies across Biomedical Domains". In: *Frontiers in Artificial Intelligence* 8. ISSN: 2624-8212. DOI: 10.3389/frai.2025.1716706 (cited on pp. 125, 156).

Van Rikxoort, Eva M and Bram Van Ginneken (Sept. 2013). "Automated Segmentation of Pulmonary Structures in Thoracic Computed Tomography Scans: A Review". In: *Physics in Medicine and Biology* 58.17, R187–R220. ISSN: 0031-9155, 1361-6560. DOI: 10.1088/0031-9155/58/17/R187 (cited on pp. 42, 60).

Vedovati, M.C. et al. (Dec. 2013). "Prognostic Role of Embolic Burden Assessed at Computed Tomography Angiography in Patients with Acute Pulmonary Embolism: Systematic Review and Meta-analysis". In: *Journal of Thrombosis and Haemostasis* 11.12, pp. 2092–2102. ISSN: 15387836. DOI: 10.1111/jth.12429 (cited on pp. 20, 142).

Walsh, Peter N. et al. (Apr. 1973). "An Angiographic Severity Index for Pulmonary Embolism". In: *Circulation* 47.4s2, pp. II–101. DOI: 10.1161/01.CIR.47.4S2.II-101 (cited on pp. 17, 142).

Wang, Yinuo, Cai Meng, and Zhe Xu (2025). "Collect vascular specimens in one cabinet: A hierarchical prompt-guided universal model for 3D vascular segmentation". In: *Computerized Medical Imaging and Graphics* 125, p. 102650. ISSN: 0895-6111. DOI: https://doi.org/10.1016/j.compmedimag.2025.102650. URL: https://www.sciencedirect.com/science/article/pii/S0895611125001594 (cited on pp. 125, 156).

Wasserthal, Jakob et al. (Sept. 2023). "TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images". In: *Radiology: Artificial Intelligence* 5.5, e230024. DOI: 10.1148/ryai.230024. URL: https://pubs.rsna.org/doi/10.1148/ryai.230024 (cited on pp. 48, 129, 143, 144).

Wendelboe, Aaron M. and Gary E. Raskob (Apr. 29, 2016). "Global Burden of Thrombosis". In: *Circulation Research*. DOI: 10.1161/CIRCRESAHA.115.306841. URL: https://www.ahajournals.org/doi/10.1161/CIRCRESAHA.115.306841 (cited on pp. 12, 141).

Wright, Susan N. et al. (Nov. 2013). "Digital Reconstruction and Morphometric Analysis of Human Brain Arterial Vasculature from Magnetic Resonance Angiography". In: *NeuroImage* 82, pp. 170–181. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2013.05.089 (cited on p. 11).

Xie, Kangxian, Jiancheng Yang, et al. (Jan. 2025). "Efficient Anatomical Labeling of Pulmonary Tree Structures via Deep Point-Graph Representation-Based Implicit Fields". In: *Medical Image Analysis* 99, p. 103367. ISSN: 1361-8415. DOI: 10.1016/j.media.2024.103367 (cited on pp. 125, 156).

Xie, Kangxian, Yufei Zhu, et al. (Mar. 2026). "Template-Guided Reconstruction of Pulmonary Segments with Neural Implicit Functions". In: *Medical Image Analysis* 109, p. 103916. ISSN: 1361-8415. DOI: 10.1016/j.media.2025.103916 (cited on pp. 125, 156).

Xu, Yajun et al. (2018). "Pulmonary Vessel Segmentation via Stage-Wise Convolutional Networks With Orientation-Based Region Growing Optimization". In: *IEEE Access* 6, pp. 71296–71305. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2018.2867859 (cited on p. 48).

Yureidini, Ahmed, Erwan Kerrien, and Stéphane Cotin (Feb. 2012). "Robust RANSAC-based Blood Vessel Segmentation". In: *Medical Imaging 2012: Image Processing*. Vol. 8314. SPIE, pp. 474–481. DOI: 10.1117/12.911670 (cited on p. 65).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Zhang, Minghui et al. (Sept. 2025). *TopoSculpt: Betti-Steered Topological Sculpting of 3D Fine-grained Tubular Shapes*. DOI: 10.48550/arXiv.2509.03938. arXiv: 2509.03938 [cs] (cited on pp. 50, 144).

Zhao, Fengjun et al. (Apr. 2019). "Segmentation of Blood Vessels Using Rule-Based and Machine-Learning-Based Methods: A Review". In: *Multimedia Systems* 25.2, pp. 109–118. ISSN: 1432-1882. DOI: 10.1007/s00530-017-0580-7 (cited on p. 42).

Zhu, Linagjia et al. (2014). "An Effective Interactive Medical Image Segmentation Method Using Fast GrowCut". In: *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Interactive Medical Image Computing Workshop* (cited on pp. 60, 129).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Synthèse en français

Introduction . . . 139
Chapitre 1 : Contexte médical . . . 141
Chapitre 2 : Jeu de données et métriques . . . 142
Chapitre 3 : État de l'art . . . 144
Chapitre 4: Développement d'un plugin semi-automatique d'annotation des structures vasculaires . . . 145
Chapitre 5 : Entraînement des modèles de segmentation vasculaire et des thrombus . 147
Chapitre 6 : Modélisation sous forme de graphe et extraction des biomarqueurs . . . 149
Chapitre 7 : Quantification de l'apport du SPCCT pour la visualisation des vaisseaux cérébraux . . . 151
Conclusion . . . 154

Introduction

L'embolie pulmonaire (EP) est une pathologie cardiovasculaire potentiellement mortelle causée par l'obstruction des artères pulmonaires par des caillots de sang, également appelés thrombus. Sa sévérité est fortement dépendante de la localisation, de l'étendue et de la distribution des caillots au sein de l'arbre artériel pulmonaire (AAP) et peut conduire à un dysfonctionnement du ventricule droit, voire à un choc obstructif et au décès. L'EP est le troisième syndrome cardiovasculaire aigu le plus fréquent après l'infarctus du myocarde, avec une incidence annuelle en hausse. Une évaluation précise de la sévérité est essentielle pour guider le traitement, mais reste difficile en pratique clinique. Le diagnostic repose principalement sur l'angioscanner pulmonaire (CTPA), qui offre une visualisation détaillée de la vascularisation pulmonaire, et les patients sont stratifiés en trois catégories de risque déterminant la stratégie thérapeutique.

Les recommandations actuelles évaluent en premier lieu l'instabilité hémodynamique. Chez les patients stables, le dysfonctionnement du ventricule droit est évalué à l'aide de marqueurs morphologiques et fonctionnels, notamment le rapport des diamètres ventriculaires droit et gauche (VD/VG) mesuré sur le CTPA, ainsi que les taux sanguins de troponine et de BNP. Cependant, cette approche présente plusieurs limites : les biomarqueurs sanguins ne sont pas toujours disponibles en situation d'urgence, ils reflètent les conséquences de l'EP plutôt que les thrombus eux-mêmes, et le CTPA n'étant pas synchronisé avec le cycle cardiaque, des artefacts de mouvement peuvent altérer l'estimation du rapport VD/VG. Pour mieux caractériser la charge thrombotique, plusieurs scores ont été proposés dans la littérature, tels que les scores de Qanadli et de Mastora. Toutefois, ces scores reposent sur des modèles vasculaires simplifiés, sont trop chronophages pour être calculés manuellement en routine, et ne tiennent pas compte de la variabilité anatomique interindividuelle, limitant ainsi leur adoption en pratique clinique.

Dans ce contexte, cette thèse s'inscrit dans le projet ANR PERSEVERE, qui vise à développer des outils computationnels pour l'évaluation automatisée de la sévérité de l'EP à partir d'images

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

CTPA. Le projet dispose d'une cohorte clinique de 431 angioscanners de patients atteints d'EP, mais sans annotations. L'objectif de cette thèse est de concevoir et d'évaluer un pipeline de bout en bout combinant segmentation par apprentissage profond, modélisation par graphe de l'AAP, et extraction automatisée de biomarqueurs locaux et globaux cliniquement pertinents, directement à partir des images CTPA. Plusieurs défis interconnectés doivent être relevés pour atteindre cet objectif : produire des annotations de référence fiables sur une cohorte clinique qui en est dépourvue, développer des modèles de segmentation robustes malgré un décalage de domaine entre données saines et pathologiques, construire une représentation par graphe topologiquement cohérente de l'AAP, et évaluer la généralisabilité du framework à d'autres territoires vasculaires.

La première contribution est le développement de VSMOD, un plugin open source pour 3D Slicer permettant la segmentation et la modélisation semi-automatisées de structures vasculaires à partir d'images médicales volumétriques. En combinant la modélisation interactive de lignes centrales avec une segmentation par croissance de région, cet outil réduit le temps d'annotation d'environ 80% par rapport aux approches manuelles, tout en préservant la connectivité vasculaire grâce à une représentation par graphe. Bien que conçu initialement pour annoter l'AAP dans des angioscanners, en modifiant les paramétrages, il peut être utilisé sur d'autres structures vasculaires, notamment les cérébrovasculaires. Ce travail a donné lieu à une publication dans le *Journal of Open Source Software* (JOSS), et l'outil est disponible publiquement sur Github[1]. Les annotations générées avec VSMOD sur un échantillon de la cohorte PERSEVERE servent de vérité terrain pour le reste des contributions.

La deuxième contribution porte sur l'entraînement de modèles robustes pour la segmentation des artères pulmonaires et des thrombus sur la cohorte PERSEVERE annotée. Des expériences initiales de transfert d'apprentissage depuis le jeu de données du challenge PARSE ont mis en évidence un décalage de domaine entre données saines et pathologiques, conduisant à une approche spécifique à la cohorte. Une stratégie itérative combinant entraînement du modèle, prédiction automatique et correction manuelle a permis d'améliorer progressivement la qualité et la quantité des annotations, aboutissant à l'entraînement de modèles nnU-Net atteignant des scores de Dice de 0,943 pour les artères pulmonaires et 0,817 pour les thrombus.

La troisième contribution est la conception d'un pipeline automatique modélisant l'AAP sous la forme d'un graphe orienté dérivé des segmentations, et extrayant des biomarqueurs locaux et globaux cliniquement pertinents pour l'évaluation de la sévérité de l'EP. Le pipeline comprend la construction et la correction du graphe, l'étiquetage hiérarchique des vaisseaux, et l'agrégation de caractéristiques au niveau des vaisseaux en métriques au niveau du patient. Il permet notamment le calcul automatisé des scores adaptés de Qanadli et Mastora ainsi que la distribution du volume embolic total (VET). De fortes corrélations de Spearman avec les scores manuels ($\rho = 0{,}83$ et $\rho = 0{,}86$) confirment la pertinence clinique des biomarqueurs produits. Ce travail a été soumis au *IEEE Journal of Biomedical and Health Informatics* (JBHI).

La quatrième contribution illustre la généralisabilité de l'approche de modélisation par graphe à un autre territoire vasculaire. Le pipeline est appliqué à des segmentations cérébrovasculaires

[1] https://github.com/omerveille/VSMOD

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

issues d'acquisitions DECT et SPCCT, permettant une comparaison quantitative de la qualité de segmentation entre modalités dans des régions d'intérêt sélectionnées. Cette expérience montre que l'approche par graphe proposée se généralise à des structures vasculaires présentant des géométries et des topologies différentes. Ce travail a été soumis et accepté pour publication dans le *Journal of Neurobiology*.

Chapitre 1 : Contexte médical

Dans ce chapitre, nous présentons le contexte anatomique et clinique nécessaire aux analyses d'imagerie médicale menées dans le cadre de cette thèse.

Nous décrivons d'abord l'organisation hiérarchique de l'AAP, depuis l'artère pulmonaire principale jusqu'aux artères sous-segmentaires, en soulignant les variations interindividuelles documentées dans la littérature, tant dans le nombre de segments (Ma et al. 2023) que dans les motifs de ramifications (Fourdrain et al. 2018; Michaud et al. 2023). Le modèle anatomique classique dénombre 19 artères segmentaires réparties entre les cinq lobes pulmonaires (Michaud et al. 2023), bien que certains scores cliniques simplifient cette représentation à 20 artères segmentaires au total. Nous montrons que la nature arborescente et acyclique de l'AAP le rend particulièrement adapté à une représentation par graphes, où chaque segment vasculaire correspond à une arête et chaque bifurcation ou fin de vaisseau à un nœud. Nous présentons également les artères cérébrales, visualisées par ARM-TOF, dont la topologie cyclique via le polygone de Willis et les trajectoires tortueuses contrastent avec celles de l'AAP (Rai et al. 2013), et qui constituent une structure d'intérêt pour tester la généralisabilité des méthodes développées.

Nous présentons ensuite l'embolie pulmonaire (EP), troisième syndrome cardiovasculaire aigu le plus fréquent après l'infarctus du myocarde et l'accident vasculaire cérébral (Wendelboe and Raskob 2016), diagnostiquée par angioscanner pulmonaire (Jiménez, de Miguel-Díez, et al. 2016; Mehdipoor et al. 2020). Sur le plan physiopathologique, l'obstruction artérielle provoque une surcharge du ventricule droit, pouvant évoluer jusqu'au choc obstructif, mettant en jeu le pronostic vital à court terme. Nous décrivons sa physiopathologie, depuis la surcharge du ventricule droit jusqu'au choc obstructif, ainsi que le rôle aggravant des comorbidités principales que sont l'insuffisance cardiaque gauche, l'insuffisance respiratoire et le cancer.

Nous exposons ensuite les outils d'évaluation de la sévérité de l'EP. Nous décrivons d'abord la stratification du risque définie par les recommandations de la Société européenne de cardiologie (Konstantinides et al. 2020), reposant sur le rapport VD/VG, les taux de BNP et de troponine, qui permet de classer les patients en trois niveaux de risque : élevé, intermédiaire et faible.

Nous présentons ensuite les scores cliniques pronostiques, dont le PESI (Aujesky et al. 2005) et sa version simplifiée, le sPESI (Jiménez, Aujesky, et al. 2010), qui permettent d'évaluer le pronostic des patients en intégrant simultanément l'impact cardiaque et les comorbidités principales que sont l'insuffisance cardiaque gauche, l'insuffisance respiratoire et le cancer.

Nous présentons enfin les scores de charge thrombotique, depuis les scores de G. A. H. Miller

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

et al. 1971 et Walsh et al. 1973, adaptés au scanner par Bankier et al. 1997, jusqu'aux scores de Qanadli et al. 2001 et Mastora et al. 2003, plus précis mais trop chronophages pour la routine clinique, ainsi que le volume embolique total (Furlan et al. 2011; W.-M. Huang et al. 2022; Nakada et al. 2010). Ces scores, et notamment ceux de Qanadli et Mastora, servent de référence pour l'évaluation automatisée développée au chapitre 6. Nous discutons enfin de la valeur pronostique de la charge thrombotique, qui, si elle ne prédit pas isolément la mortalité à court terme (Vedovati et al. 2013), présente des corrélations significatives avec plusieurs paramètres cliniques (Rotzinger et al. 2020; Tajeri et al. 2024) et le risque de complications à long terme (An et al. 2022; Tromeur et al. 2018).

Face à ces limitations, le projet Pulmonary Embolism Risk Stratification basEd on Vascular nEtwoRk modEling (PERSEVERE), vise à automatiser l'extraction de biomarqueurs locaux et globaux directement depuis l'angioscanner, selon trois étapes successives : la segmentation de l'AAP et des thrombus, la modélisation par graphes, et l'extraction des biomarqueurs. Ces contributions constituent le cœur des travaux développés dans cette thèse (voir Figure A.1).

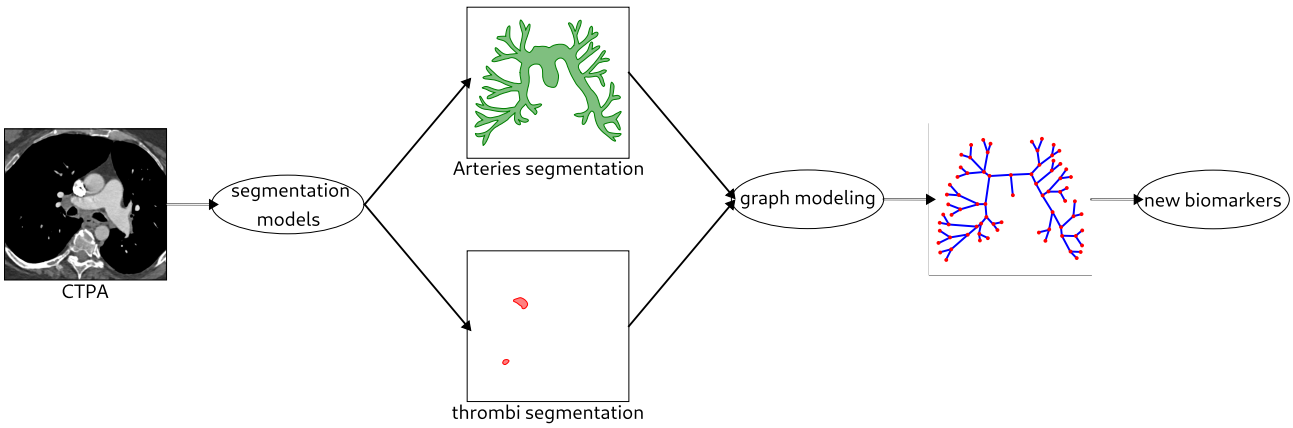


Fig. A.1: Visualisation globale des objectifs du projet PERSEVERE

Chapitre 2 : Jeu de données et métriques

Dans ce chapitre, nous présentons les jeux de données et le cadre d'évaluation utilisés dans cette thèse.

Nous commençons par décrire le jeu de données interne PERSEVERE composé de 431 patients diagnostiqués d'une EP aiguë au CHU de Saint-Étienne entre avril 2014 et juin 2020. Les patients présentant une instabilité hémodynamique n'ont pas été inclus, excluant d'emblée les cas à haut risque. Après application de trois critères d'exclusion portant sur l'homogénéité du fabricant, la complétude des biomarqueurs sanguins (BNP et troponine) et la qualité des images, une cohorte finale de 353 patients a été retenue. Nous montrons notamment qu'aucune normalisation d'intensité supplémentaire n'est nécessaire entre les deux modèles de scanner retenus (différence moyenne de 22.5 HU, $t = -1.343$, $p = 0.180$). La cohorte retenue est dominée par les patients à risque intermédiaire, qui représentent environ 86% de la population totale (risque intermédiaire-bas : 45.0%, risque intermédiaire-haut : 41.1%). Les patients à faible risque ne

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

représentent que 13.9% de la cohorte. Les images ont été prétraitées par rééchantillonnage à l'espacement médian et un recadrage est effectué autour du masque pulmonaire obtenu avec TotalSegmentator (Wasserthal et al. 2023). Un sous-ensemble de 10 patients, répartis sur les trois catégories de sévérité, a été sélectionné pour l'annotation manuelle (voir Figure A.2).

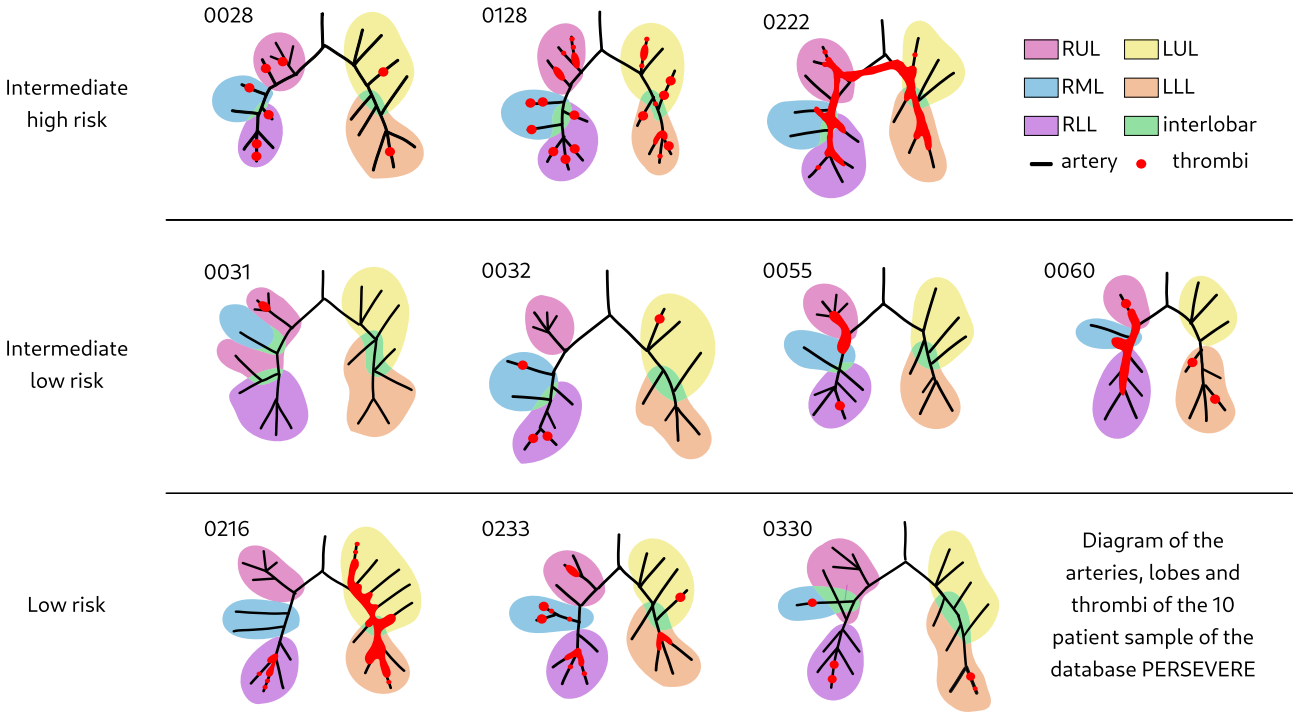


Fig. A.2: Illustration des 10 patients sélectionnés pour l'échantillon avec une vue d'ensemble de la localisation des thrombus. Chaque lobe est représenté par une couleur, et la portion interlobaire correspond à la zone où une artère traverse d'un lobe à un autre. Les artères périphériques correspondent au niveau segmentaire.

Nous présentons ensuite le jeu de données public PARSE 2022 (Luo et al. 2024), composé de 100 volumes d'angioscanner annotés pour la segmentation des artères pulmonaires chez des patients sains. Ce jeu de données a permis de tester le transfert de modèles entraînés sur des patients sains vers des données de patients pathologiques (voir chapitre 5).

Nous décrivons enfin les métriques utilisées, qui combinent des métriques de recouvrement et des métriques topologiques. Le Dice et le clDice (Shit et al. 2021) quantifient respectivement l'accord volumétrique global et la connectivité des structures tubulaires, le second s'appuyant sur les squelettes des masques pour accorder une importance égale aux petits et aux grands vaisseaux. Les métriques topologiques reposent sur les nombres de Betti, dont les trois premiers nombres (β_0, β_1, β_2) permettent de détecter respectivement les déconnexions, les cycles non anatomiques et les cavités pour des objets en trois dimensions. Pour l'AAP, dont la topologie attendue est celle d'un arbre connexe et acyclique, une segmentation idéale satisfait $\beta_0 = 1$, $\beta_1 = 0$ et $\beta_2 = 0$, et tout écart constitue un indicateur d'erreur topologique que les métriques de recouvrement seules ne permettraient pas de détecter.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Chapitre 3 : État de l'art de la segmentation des structures pulmonaires et de la modélisation de l'AAP

Nous présentons dans ce chapitre un état de l'art partant de la segmentation des structures anatomiques jusqu'à la modélisation par graphes et l'extraction de biomarqueurs cliniques à partir d'images CTPA.

Nous introduisons d'abord les principales architectures d'apprentissage profond utilisées dans cette thèse. Nous décrivons l'architecture U-Net (Ronneberger, Fischer, and Brox 2015), qui a établi le paradigme dominant pour la segmentation médicale, puis son extension volumétrique 3D U-Net (Çiçek et al. 2016), qui donne accès au contexte spatial inter-coupes au prix d'une empreinte mémoire plus élevée. Nous présentons ensuite des variantes allégées comme le TinyU-Net (Rougé, Passat, and Merveille 2024), des approches semi-supervisées telles que le cadre Mean Teacher (Tarvainen and Valpola 2018), des architectures en cascade intégrant la perte clDice pour encourager la connectivité vasculaire, ainsi que le pipeline nnU-Net (Isensee et al. 2021), qui automatise la configuration des hyperparamètres et constitue aujourd'hui une référence pour de nombreuses tâches de segmentation biomédicale.

Nous présentons ensuite les méthodes de segmentation des poumons et des lobes, qui constituent une étape préliminaire indispensable à l'analyse vasculaire. Nous décrivons en particulier `Lungmask` (Hofmanninger et al. 2020), reposant sur des ensembles de modèles 2D U-Net, et TotalSegmentator (Wasserthal et al. 2023), un outil de segmentation multi-organes construit sur nnU-Net couvrant plus de cent quatre-vingts structures anatomiques. Nous soulignons leur capacité à traiter des données hétérogènes issues de scanners cliniques courants, ce qui les rend particulièrement adaptés aux pipelines d'analyse automatisée.

Nous abordons ensuite la segmentation de l'arbre artériel pulmonaire, pour laquelle les défis sont plus importants. Nous montrons que la qualité des annotations constitue un facteur critique (Lekadir et al. 2025), les vaisseaux distaux étant fréquemment sous-annotés en raison de leur faible contraste. Nous discutons de la cohérence topologique des segmentations, en soulignant que les fonctions de coût volumétriques standard telles que le Dice sont insensibles aux erreurs de connectivité. Nous présentons les approches proposées pour pallier cette limitation : la fonction de coût clDice (Shit et al. 2021), des méthodes de post-traitement correctif (Carneiro-Esteves, Vacavant, and Merveille 2025; Keshwani et al. 2020) et des méthodes incluant les nombres de Betti (Hu et al. 2019; Sofi and Alsahanova 2023; Zhang et al. 2025).

Nous présentons ensuite les méthodes de segmentation des thrombus, dont les défis sont de nature différente. Nous décrivons une première famille de méthodes reformulant le problème en termes de localisation par boîte englobante (Long et al. 2021), et une seconde cherchant à réduire la dépendance aux annotations manuelles, notamment la supervision faible par seuillage adaptatif proposée par Pu et al. 2023. Nous montrons que la distinction entre défauts de remplissage partiels et artéfacts de mouvement reste un problème ouvert.

Nous présentons également les approches de quantification de la charge thrombotique, organisées en deux catégories : les méthodes basées sur les scores cliniques de Qanadli et de Mastora

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

(Peters et al. 2007), et les méthodes fondées sur le Volume Embolique Total (W.-M. Huang et al. 2022; W. Liu et al. 2020). Nous montrons que ces deux familles de mesures sont complémentaires plutôt que redondantes, et qu'aucune des méthodes recensées n'intègre la localisation des embolies dans un modèle structuré de l'AAP.

Enfin, nous présentons la modélisation par graphes de l'arbre artériel pulmonaire. Nous décrivons l'extraction de la ligne centrale et les deux types de graphes qui peuvent en être construits : le graphe voxelique, offrant une représentation localement précise, et le graphe de branches, plus compact et directement interprétable cliniquement (voir Figure A.3). Nous illustrons les applications de ces modèles : simulation hémodynamique (Reichold et al. 2009; Reymond et al. 2009), génération de maillages structurés (Decroocq et al. 2023), quantification morphologique (Mishra et al. 2021), et construction de jeux de données annotés (Lamy et al. 2022; Li et al. 2026). Nous soulignons que la construction de ces graphes à partir de segmentations imparfaites reste un défi, les erreurs topologiques se propageant aux analyses en aval, ce qui motive les choix méthodologiques développés dans la suite de cette thèse.

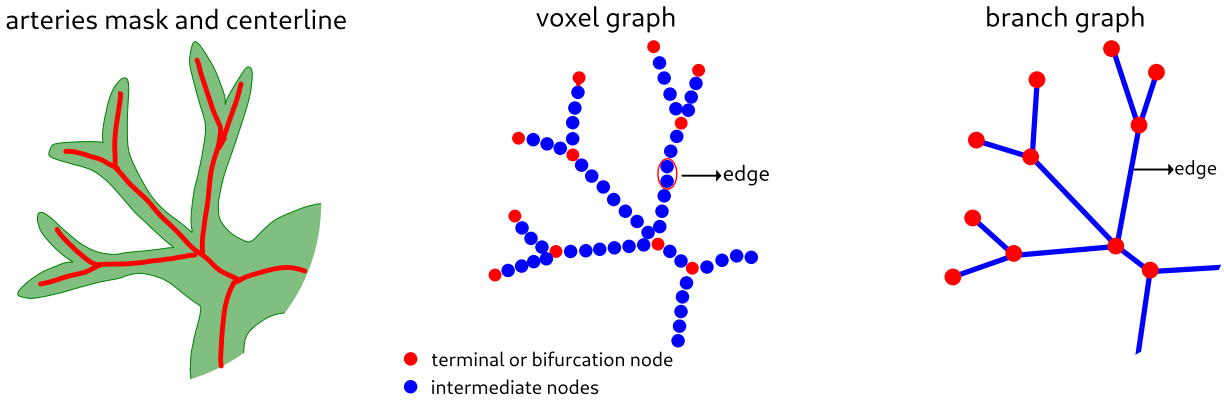


Fig. A.3: Illustration de deux représentations par graphes basées sur la ligne centrale, extraites d'une segmentation vasculaire.

Les observations motivent l'approche développée dans cette thèse, qui vise à construire un modèle de graphe de l'AAP anatomiquement cohérent, intégrant la géométrie vasculaire, la localisation spatiale et le positionnement des thrombus, afin de permettre le calcul automatique de biomarqueurs cliniquement pertinents tels que les scores de Qanadli et de Mastora et le volume thrombotique total.

Chapitre 4: Développement d'un plugin semi-automatique d'annotation des structures vasculaires

L'annotation volumétrique précise des réseaux vasculaires constitue un prérequis indispensable pour l'entraînement et l'évaluation des méthodes de segmentation vasculaire basées sur l'apprentissage profond. Dans le cadre du projet PERSEVERE, qui regroupe des acquisitions

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

de CTPA de patients atteints d'EP, l'absence d'annotations manuelles et la complexité des images pathologiques rendent cette tâche particulièrement difficile.

Nous évaluons dans un premier temps plusieurs outils de segmentation disponibles dans 3D Slicer sur un patient de la base de données PARSE afin d'en identifier les limitations. Nous testons les modèles automatiques `Lungmask` et `TotalSegmentator`, qui, bien que produisant des masques pulmonaires visuellement satisfaisants, ne permettent pas de distinguer les artères des veines et génèrent des segmentations vasculaires incomplètes ou mal connectées, notamment au niveau du hile pulmonaire. Nous évaluons ensuite l'outil `Grow from Seeds`, combiné à `Draw Tube`, dans le cadre d'une approche semi-automatique par croissance de région. L'idée est de placer un équivalent de ligne centrale dans les vaisseaux avec un Draw Tube comme graine de segmentation, et de contraindre la croissance de région de ce label par un second label correspondant à une opération morphologique sur des masques Lungmask : on soustrait au masque des poumons le masque des vaisseaux dilatés de 2 mm. Cette approche s'avère extrêmement coûteuse en temps : le placement manuel de chaque tube le long des branches artérielles nécessite plusieurs heures pour un seul patient, et les résultats restent bruités et incomplets par rapport à la vérité terrain. Nous testons enfin l'extension `RVesselX` (Lamy et al. 2022), conçue pour la segmentation du réseau vasculaire hépatique, qui démontre l'intérêt des outils interactifs exploitant des a priori topologiques. Cependant, les segmentations obtenues sur l'arbre artériel pulmonaire demeurent incomplètes et limitées aux vaisseaux proximaux, confirmant que cet outil ne peut pas être directement transposé à d'autres territoires vasculaires.

Face à ces limitations, nous proposons VSMOD (des Ligneris et al. 2025), une extension open-source de 3D Slicer reposant sur deux modules complémentaires. Le premier module, dédié à l'extraction des lignes centrales, utilise un algorithme de suivi itératif fondé sur RANSAC pour ajuster des cylindres successifs le long des vaisseaux. L'utilisateur initialise chaque vaisseau en plaçant deux points dans la lumière artérielle, un point de départ et un point de direction, et l'algorithme propage automatiquement la ligne centrale en estimant le rayon local à chaque étape. Les vaisseaux sont organisés sous la forme d'un graphe hiérarchique, exportable au format JSON via la bibliothèque NetworkX. Plusieurs paramètres avancés permettent d'adapter le comportement de l'algorithme à différentes anatomies, notamment l'angle maximal entre cylindres consécutifs, le pourcentage de points inliers, ou encore le nombre maximum de tentatives de placement de cylindre. Le second module exploite ce graphe pour placer automatiquement des graines à l'intérieur des vaisseaux, dont la taille est adaptée au rayon estimé localement. Une bordure est également générée autour des vaisseaux afin de contraindre la croissance de région et d'éviter les fuites dans les tissus environnants, produisant ainsi une annotation volumétrique complète et topologiquement cohérente de l'arbre vasculaire.

Nous évaluons VSMOD sur le premier échantillon de 10 patients du jeu de données PERSEVERE, contenant en moyenne 269 vaisseaux identifiables par patient. Le temps moyen d'annotation par image avec VSMOD est de 3 heures et 35 minutes, contre une estimation de 21 heures et 22 minutes pour une annotation entièrement manuelle avec les outils natifs de 3D Slicer, soit une réduction du temps d'annotation d'environ 80% (voir tableau A.1). VSMOD produit en parallèle un graphe de lignes centrales avec des estimations de rayons, informations non disponibles avec les outils conventionnels.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Méthode d'annotation	Temps par vaisseaux	Temps total par patients
Outils manuels (3D Slicer)	4m46s	21h22m52s
VSMOD	48s	3h35m12s

Tab. A.1: Temps de segmentation moyens par patient pour un réseau vasculaire pulmonaire contenant en moyenne 269 vaisseaux.

Nous évaluons également la généralité de l'approche sur une image d'angiographie par résonance magnétique (ARM) cérébrale issue de la base de données BraVa. Sans modifier le pipeline algorithmique, il suffit d'augmenter le paramètre d'angle maximal inter-cylindres de 60° à 90° pour adapter le suivi aux vaisseaux cérébraux, qui présentent une tortuosité plus élevée et des calibres plus faibles que les artères pulmonaires. Les résultats qualitatifs montrent que VSMOD produit des arbres de lignes centrales cohérents ainsi que des segmentations volumétriques fidèles sur ce territoire vasculaire distinct.

Les annotations produites par VSMOD sur le jeu de données PERSEVERE constituent les vérités terrain sur lesquelles reposent les contributions suivantes de cette thèse.

Chapitre 5 : Entraînement des modèles de segmentation vasculaire et des thrombus

Dans ce chapitre, nous présentons la stratégie d'entraînement mise en œuvre pour obtenir des segmentations fiables de l'AAP et des thrombus sur la cohorte PERSEVERE.

Dans un premier temps, nous évaluons quatre modèles afin de tester différentes stratégies de transfert à partir du jeu de données public PARSE 2022 , composé de 100 scanners CTPA annotés pour la segmentation de l'artère pulmonaire. Le Modèle 1 est un TinyUNet supervisé, entraîné uniquement sur les données PARSE annotées. Le Modèle 2 est un Mean Teacher semi-supervisé, combinant les données PARSE annotées et les images PERSEVERE non annotées. Le Modèle 3 reprend cette même architecture, mais est cette fois pré-entraîné en initialisant les poids à partir du Modèle 1. Enfin, le Modèle 4 est un TinyUNet supervisé, entraîné exclusivement sur un petit échantillon de 10 patients PERSEVERE annotés manuellement. Les résultats mettent en évidence un écart de domaine important entre les sujets sains du jeu PARSE et les patients atteints d'EP de la cohorte PERSEVERE. Ces expériences préliminaires confirment qu'un petit ensemble d'annotations spécifiques au domaine est insuffisant pour produire des segmentations topologiquement consistantes, motivant ainsi le développement d'une stratégie d'annotation itérative dédiée à PERSEVERE.

Face aux limitations identifiées, nous mettons en place une stratégie d'annotation itérative inspirée de l'apprentissage actif, illustrée en Figure A.4. Ce protocole alterne entre l'entraînement de modèles et la correction manuelle guidée par un expert pneumologue, afin de construire progressivement un ensemble d'annotations fiables. Les structures ciblées sont l'arbre artériel pulmonaire, annoté via le plugin VSMOD, et les thrombus, annotés entièrement manuellement

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

et validés par un expert. Les masques pulmonaires sont générés automatiquement avec TotalSegmentator.

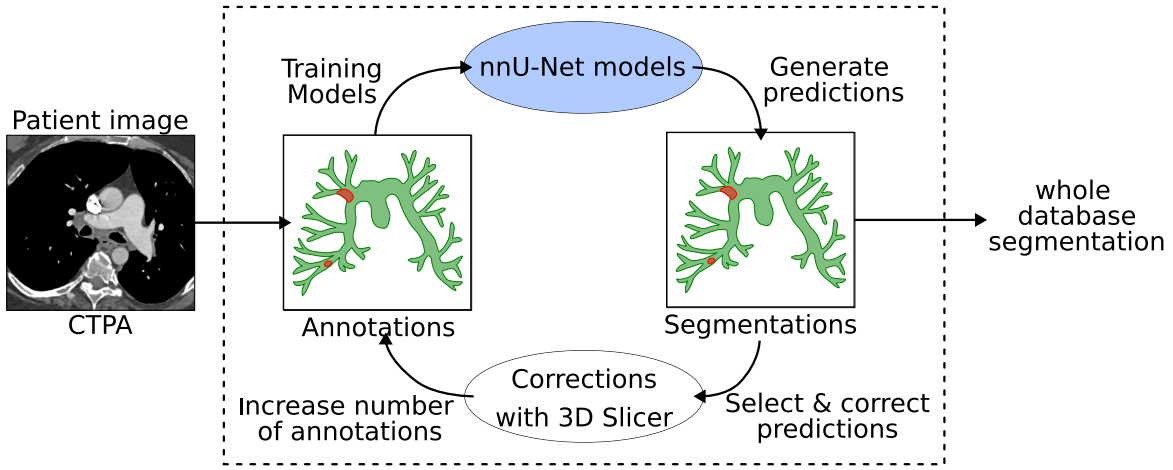


Fig. A.4: Protocole d'annotation itératif inspiré de l'apprentissage actif, utilisé pour obtenir des segmentations fiables de l'arbre artériel pulmonaire et des thrombus sur l'ensemble de la base de données PERSEVERE.

Pour cette phase, nous adoptons le pipeline nnU-Net en remplacement des architectures personnalisées utilisées lors des expériences préliminaires. Sa capacité à adapter automatiquement son architecture, son prétraitement et sa configuration d'entraînement aux caractéristiques du jeu de données en fait un choix particulièrement adapté, même avec un nombre limité d'annotations. Nous entraînons quatre modèles à chaque itération, recevant des combinaisons différentes de masques anatomiques en entrée : artères seules, thrombus seuls, artères et thrombus combinés, et artères, thrombus et poumons combinés.

Lors de la première itération, nous nous appuyons sur les 10 patients PERSEVERE annotés manuellement avec le plugin VSMOD. Les résultats montrent que la combinaison d'artères, de thrombus et de poumons (Modèle 4) permet d'obtenir les meilleures performances pour la segmentation des thrombus, avec un score Dice de 0,67. Pour la segmentation artérielle, nous n'observons aucune amélioration significative en ajoutant du contexte anatomique supplémentaire, les meilleures performances étant obtenues avec le Modèle 1, avec un score Dice de 0,91. À la suite de cette itération, nous sélectionnons 20 patients supplémentaires, dont les segmentations artérielles sont corrigées via VSMOD et validées par l'expert. Les annotations de thrombus sont également révisées, bien que 6 patients aient dû être exclus en raison d'une qualité d'image insuffisante.

Lors de la deuxième itération, nous intégrons l'ensemble des 24 annotations complètes disponibles. Les 6 patients supplémentaires disposant de masques artériels sont utilisés pour entraîner un modèle additionnel, le Modèle 1bis. Les résultats confirment une amélioration par rapport à la première itération. Pour la segmentation artérielle, le Modèle 1bis atteint un score Dice de 0,94, identique à celui du Modèle 1, mais est retenu pour le déploiement final en raison de son entraînement sur le plus grand ensemble d'annotations disponible. Pour la segmentation des thrombus, le Modèle 4 obtient à nouveau les meilleures performances avec

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

un score Dice de 0,82, et est choisi pour être appliqué à l'ensemble de la base de données PERSEVERE.

Nous appliquons ensuite le Modèle 1bis et le Modèle 4 à l'ensemble des 353 patients PERSEVERE sélectionnés. Des post-traitements sont appliqués afin de garantir la cohérence anatomique et topologique des masques : pour les artères pulmonaires, nous ne conservons que la plus grande composante connexe, assurant un arbre artériel unique par patient. Pour les thrombus, les voxels situés en dehors du masque artériel sont supprimés, de sorte que l'ensemble des thrombus prédits reste strictement à l'intérieur de la lumière artérielle.

Ces segmentations constituent la base de la chaîne d'extraction de biomarqueurs décrite dans le chapitre suivant.

Chapitre 6 : Modélisation sous forme de graphe et extraction des biomarqueurs

Dans ce sixième chapitre, nous motivons d'abord le recours à une représentation par graphe de l'AAP pour l'évaluation de la sévérité de l'EP. Nous soulignons les limites des scores cliniques existants, tels que les scores de Qanadli et de Mastora, qui reposent sur des modèles anatomiques simplifiés, sont chronophages et ne tiennent pas compte de l'architecture vasculaire propre à chaque patient. Nous proposons ensuite un pipeline automatique qui modélise l'AAP sous la forme d'un graphe orienté, à partir des segmentations issues du CTPA, et extrait des biomarqueurs locaux et globaux pertinents pour l'évaluation de la charge thrombotique de l'EP.

Nous décrivons le pipeline en quatre étapes principales. Les trois premières génèrent le graphe et produisent des biomarqueurs locaux, qui sont ensuite agrégés lors de la quatrième étape pour produire des biomarqueurs à l'échelle du patient (voir Figure A.5). La première étape consiste à convertir la segmentation binaire des artères pulmonaires en un graphe dirigé, anatomiquement cohérent, enrichi d'informations géométriques. Nous extrayons le squelette, c'est-à-dire la ligne centrale, et déduisons le profil de rayon à partir d'une carte de distance pour chaque voxel de cette ligne. Nous appliquons ensuite une étape de correction afin de supprimer les cycles non physiologiques et les artefacts de squelettisation, en traitant deux types d'erreurs : les branches terminales spurieuses, identifiées grâce à un critère combinant la longueur de la branche et le diamètre de l'artère parente, et les cycles, causés par des contacts entre artères adjacentes dans le masque de segmentation. Nous comparons les performances de cette étape de correction à celles du pipeline . Le graphe corrigé est ensuite orienté à partir de la racine, correspondant à l'APP, en utilisant un algorithme de parcours en largeur.

Nous enrichissons ensuite le graphe orienté avec un contexte anatomique par labelisation hiérarchique, en assignant à chaque arête une latéralité, une appartenance lobaire et un niveau hiérarchique. Nous proposons un critère morphologique basé sur le rapport de diamètre et le rapport de longueur pour distinguer les troncs artériels des bifurcations, avec des seuils optimisés sur un sous-ensemble de 24 patients.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

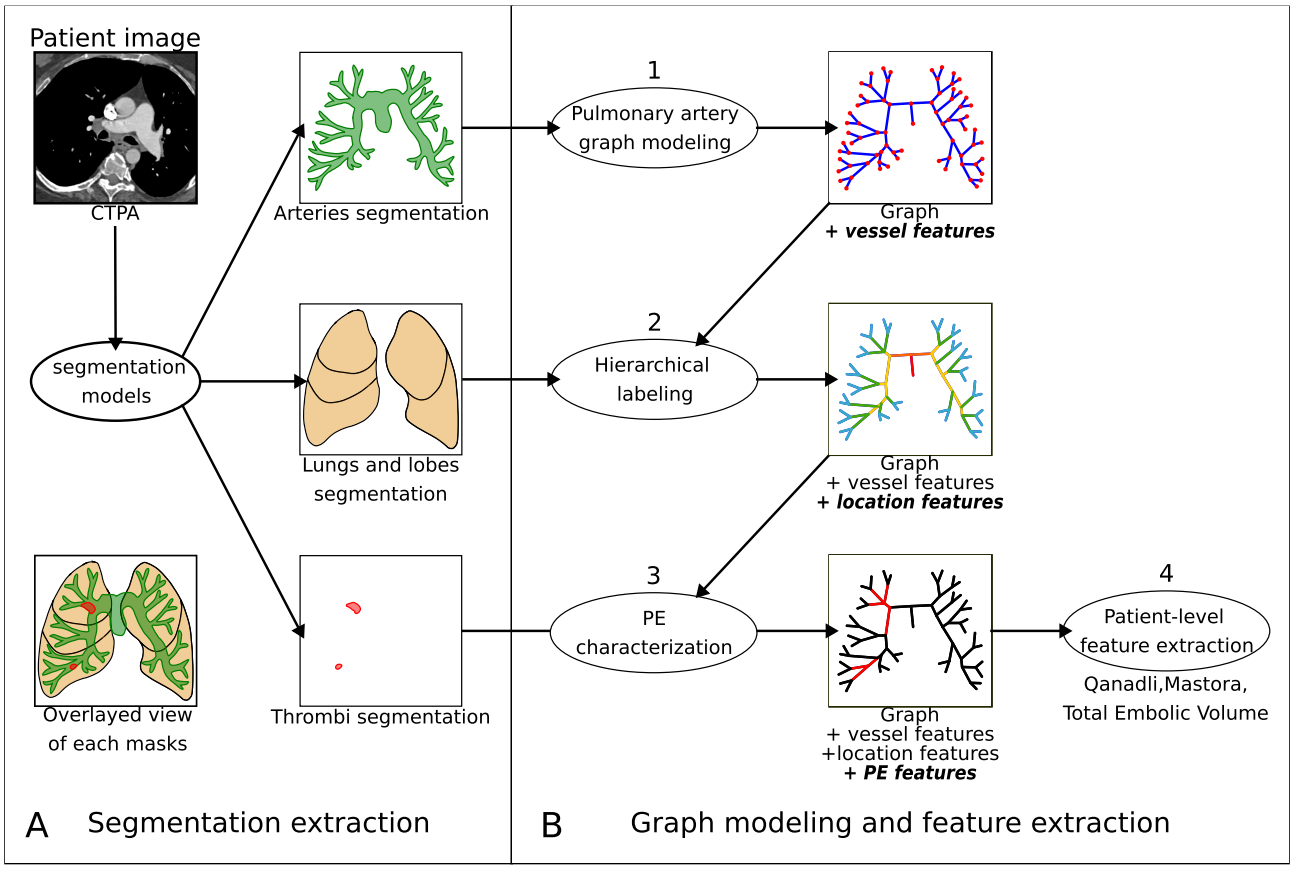


Fig. A.5: Schéma général du pipeline proposé.
A) Segmentation des artères, des poumons, des lobes et des thrombi décrite dans le Chapitre 5. B) Pipeline de modélisation par graphe et d'extraction de caractéristiques. Les étapes (1-3) génèrent le graphe et produisent des biomarqueurs locaux, qui sont ensuite agrégés lors de l'étape (4) pour produire des biomarqueurs à l'échelle du patient.

Une fois le graphe hiérarchiquement labellisé, nous extrayons, pour chaque arête, des caractéristiques spécifiques aux thrombi à partir de l'intersection du masque de thrombus avec le graphe artériel. Nous calculons le volume de thrombus par intersection directe du masque de segmentation avec le volume artériel correspondant, et nous appliquons une reformation planaire courbe (CPR) afin de mesurer l'obstruction transversale maximale le long de chaque arête.

Les caractéristiques locales sont ensuite agrégées pour produire des biomarqueurs à l'échelle du patient, notamment le volume embolique total (TEV) et sa distribution par lobe et par niveau hiérarchique, ainsi que des adaptations automatisées des scores de Qanadli et de Mastora s'appuyant sur la structure du graphe patient-spécifique.

Nous présentons ensuite les résultats sur les 24 patients annotés du jeu de données PERSEVERE ainsi que sur les segmentations des 353 patients obtenues au chapitre précédent. Sur les 24 patients annotés, nous montrons que la correction du graphe réduit le nombre de nœuds et d'arêtes de 22% en moyenne, élimine l'intégralité des artefacts sur l'APP et réduit le nombre moyen de cycles par patient de 1,2 à 0,1, ce qui dépasse le pipeline `VesselGraph`. Nous

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

montrons également que les seuils morphologiques retenus permettent d'obtenir sur les segmentations des 353 patients un nombre d'artères segmentaires proche des valeurs anatomiques de référence, avec un nombre moyen de 20,2 artères segmentaires détectées pour une valeur attendue de 19 selon la littérature anatomique. Nous évaluons l'accord entre les scores calculés automatiquement et les scores manuels sur 10 patients par analyse de Bland-Altman et corrélation de Spearman, et nous obtenons des coefficients de Spearman de 0,83 pour le score de Qanadli et 0,86 pour le score de Mastora. Enfin, nous illustrons la capacité du graphe à caractériser finement la distribution spatiale de la charge thrombotique, en distinguant des profils d'obstruction proximale et distale à volume embolique total comparable, et nous reportons un temps de traitement moyen d'environ neuf minutes par patient.

La modélisation par graphe constitue ainsi une représentation patient-spécifique de la vascularisation pulmonaire, capable de prendre en compte la variabilité anatomique tout en supportant l'extraction d'un ensemble de biomarqueurs. Par ailleurs, les composants de modélisation vasculaire en graphe du pipeline, notamment l'extraction de la ligne centrale et le calcul des caractéristiques vasculaires, s'appliquent à d'autres structures vasculaires segmentables, ouvrant la voie à des applications dans d'autres contextes cliniques. Dans le chapitre suivant, nous essayons de les appliquer à l'imagerie neurovasculaire.

Chapitre 7 : Quantification de l'apport du SPCCT pour la visualisation des vaisseaux cérébraux

L'imagerie neurovasculaire nécessite une résolution submillimétrique pour délimiter avec précision les vaisseaux intracrâniens de petit calibre. Le scanner à double énergie (DECT) est actuellement la modalité de référence pour ce type d'imagerie, mais souffre d'une résolution spatiale limitée et d'artefacts métalliques importants en présence d'implants tels que des coils ou des stents. Le scanner à comptage photonique (SPCCT) est une modalité récente qui combine une résolution spatiale ultra-élevée offrant ainsi la visualisation de détails vasculaires jusqu'alors indétectables (Greffier, Villani, et al. 2023).

Dans ce chapitre, nous présentons une étude pilote visant à quantifier l'apport du SPCCT par rapport au DECT pour la visualisation des artères cérébrales, réalisée en collaboration avec le Dr Salim Si-Mohamed. Les résultats ont été acceptés pour publication dans le *Journal of Neuroradiology*. Nous exploitons des acquisitions SPCCT et DECT d'un même patient, un homme de 68 ans porteur de plusieurs coils cérébraux, ayant bénéficié d'une angiographie par scanner cérébral dans les deux modalités à trois jours d'intervalle. Quatre types de reconstructions sont comparés : des images polychromatiques et des images monoénergétiques virtuelles à 40 keV, pour chacune des deux modalités (DECT-p, DECT-40keV, SPCCT-p, SPCCT-40keV). Le SPCCT offre une résolution nettement supérieure au DECT avec une taille de voxel de 0,25 x 0,23 x 0,23 mm^3 contre 0,48 x 0,48 x 0,80 mm^3.

Nous définissons trois régions d'intérêt (ROI), sélectionnées par des experts en neuroradiologie pour leur importance clinique dans le suivi post-opératoire : le polygone de Willis, qui contient de petites artères perforantes difficiles à visualiser avec les modalités conventionnelles;

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

la région à proximité d'un coil, où les artefacts métalliques limitent sévèrement l'évaluation post-opératoire en DECT; et les artères lenticulostriées gauches et droites (LAs), de diamètre inférieur à 1 mm, qui constituent l'une des cibles les plus difficiles en imagerie vasculaire cérébrale.

Nous produisons les segmentations vasculaires initiales automatiquement à l'aide d'un modèle de segmentation des artères cérébrales (Rougé, Passat, and Merveille 2024), puis nous les corrigeons manuellement sous 3D Slicer et les faisons valider par un neuroradiologue senior. Nous appliquons ensuite le pipeline de modélisation par graphe, développé initialement pour l'imagerie pulmonaire, aux masques de segmentation cérébraux. Chaque segment vasculaire est représenté par une arête du graphe, ce qui nous permet d'extraire des mesures par vaisseau, notamment le volume total segmenté, le nombre de segments artériels distincts, ainsi que le diamètre moyen, minimal et maximal par vaisseau.

Nous montrons que le SPCCT permet de détecter plus du double du nombre de segments artériels par rapport au DECT dans le polygone de Willis (53 segments pour SPCCT-p contre 26 pour DECT-p). Le diamètre minimal détectable atteint 0,62 mm avec le SPCCT, contre 1,04 mm avec le DECT, et le volume total segmenté est supérieur d'environ 20% en faveur du SPCCT. Ces résultats indiquent que le SPCCT permet de capturer des branches distales et des artères perforantes qui restent en dessous du seuil de détection du DECT. Concernant les LAs, nous démontrons qu'elles ne peuvent être segmentées de manière fiable que sur les acquisitions SPCCT, étant totalement indétectables sur les deux reconstructions DECT. Le SPCCT-p permet de détecter respectivement 10 et 30 segments artériels dans les LAs gauches et droites, avec des diamètres minimaux atteignant 0,57 mm. Nous observons que les reconstructions à 40 keV réduisent légèrement le nombre de segments détectés, ce qui s'explique par une meilleure discrimination entre artères et veines, permettant d'éliminer certains faux positifs. Dans la région du coil, alors que les artefacts métalliques rendent la lumière artérielle adjacente indiscernable en DECT, nous montrons que le SPCCT permet une délimitation précise du réseau artériel avec le coil inclus, avec 234 mm^3 de volume artériel segmenté et 19 segments identifiés, présentant des diamètres compris entre 0,74 mm et 1,46 mm (voir Figure A.6).

Nous mettons également en évidence un effet différentiel notable des reconstructions à 40 keV selon la modalité. En DECT, les images à 40 keV tendent à augmenter le volume vasculaire apparent, probablement en raison d'un élargissement artificiel des contours vasculaires. En SPCCT, en revanche, les images à 40 keV améliorent la spécificité et la visibilité des lumières distales, sans entraîner de surestimation des diamètres. Cela suggère que la combinaison d'une haute résolution spatiale et de l'imagerie spectrale permet de bénéficier de l'amélioration du contraste iodé sans les inconvénients observés sur les systèmes à intégration d'énergie.

En résumé, dans ce chapitre, nous confirmons également que le pipeline de modélisation par graphe, initialement conçu pour les artères pulmonaires, est adaptable aux structures vasculaires cérébrales, plus tortueuses et géométriquement complexes. De plus, les résultats obtenus à partir des graphes suggèrent que le SPCCT pourrait devenir la modalité non invasive de référence pour le suivi des patients traités par coils ou stents, en réduisant potentiellement le recours à l'artériographie invasive.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

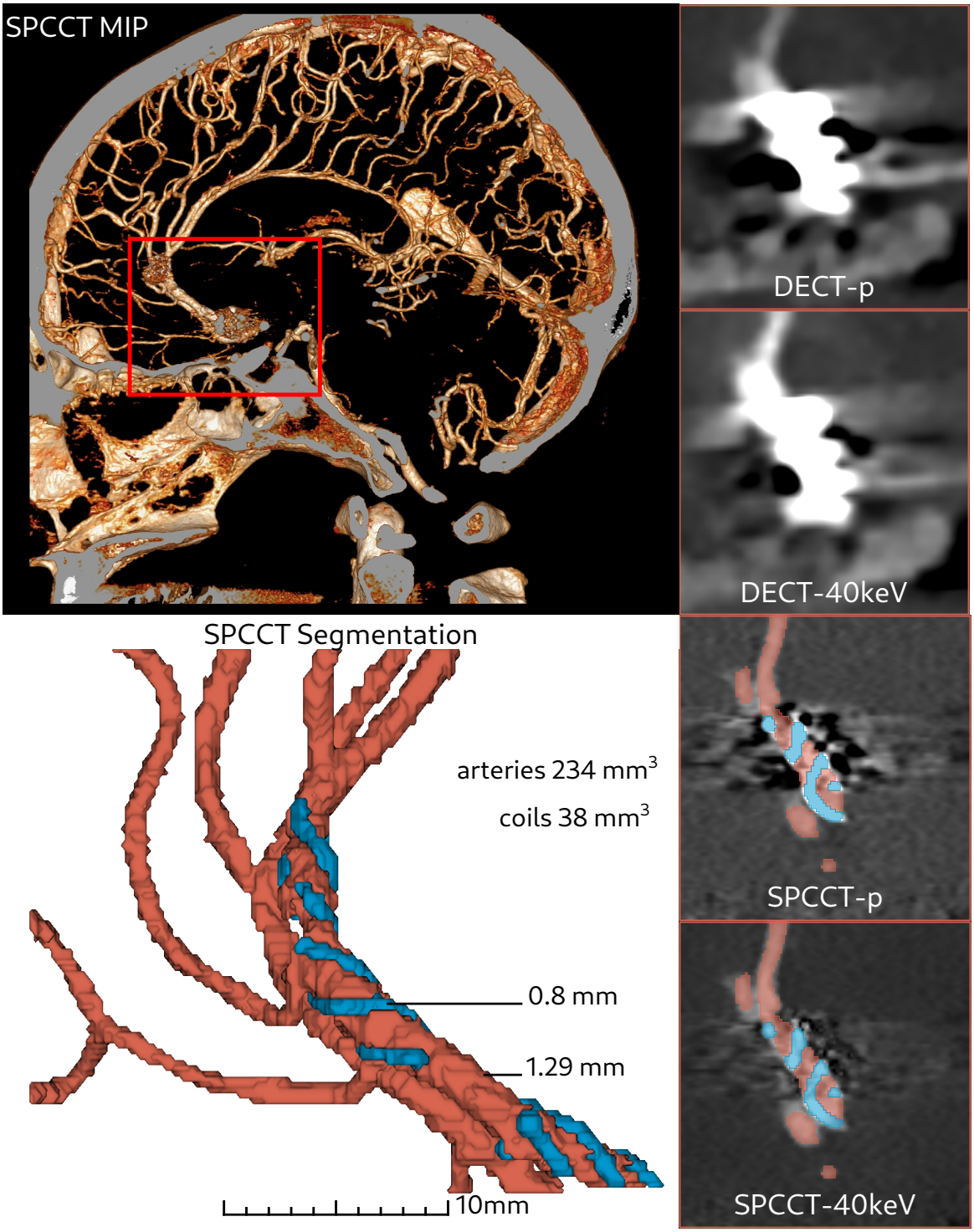


Fig. A.6: Reconstructions en projection d'intensité maximale (MIP) selon un plan sagittal oblique (épaisseur de dalle de 15 mm) et coupes agrandies pour la visualisation de la lumière artérielle en présence d'un coil, sur les images DECT-p, DECT-40keV, SPCCT-p et SPCCT-40keV avec les segmentations correspondantes. Les artères sont représentées en rouge et les coils en bleu.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Conlcusion

Cette thèse est motivée par le besoin clinique d'outils automatisés et quantitatifs pour la stratification du risque dans l'EP. L'objectif principal était de concevoir un pipeline complet, depuis les acquisitions CTPA jusqu'à l'extraction de biomarqueurs à partir de représentations en graphe, capable de caractériser l'AAP et la charge thrombotique. Ce travail a nécessité de relever trois défis : produire des annotations de référence fiables, entraîner des modèles de segmentation par apprentissage profond robustes et modéliser une représentation en graphe anatomiquement cohérente de l'AAP.

Première contribution. Face à l'absence d'annotations de référence dans le jeu de données PERSEVERE, nous avons développé VSMOD, un plugin open-source pour la segmentation vasculaire semi-automatique. VSMOD réduit le temps d'annotation de 80% par rapport aux approches entièrement manuelles, produit des masques de segmentation topologiquement connexes, et génère directement une représentation en graphe du réseau vasculaire. À partir de ces propriétés, un ensemble d'annotations de haute qualité de l'AAP a été produit depuis la cohorte PERSEVERE, constituant le socle de toutes les contributions suivantes. Ce travail a fait l'objet d'une publication dans le *Journal of Open Source Software (JOSS)*.

Deuxième contribution. Nous avons développé et évalué des modèles d'apprentissage profond pour la segmentation automatisée de l'artère pulmonaire et des thrombus, entraînés sur les annotations générées avec VSMOD. Ces modèles ont démontré de solides performances sur la cohorte PERSEVERE et ont permis l'extraction automatisée à grande échelle des informations vasculaires et thrombotiques.

Troisième contribution. Nous avons proposé un pipeline complet de modélisation en graphe de l'AAP, incluant l'extraction de lignes centrales, la construction du graphe, l'étiquetage hiérarchique et le calcul de biomarqueurs. Ce pipeline fournit une représentation structurée et quantitative de l'arbre vasculaire et de la charge thrombotique. Ce travail a été soumis au *Journal of Biomedical and Health Informatics (JBHI)*.

Quatrième contribution. Nous avons démontré la transférabilité de ce cadre de modélisation en graphe à un territoire anatomique différent, à savoir l'imagerie neurovasculaire. Le pipeline a été appliqué à des angiographies cérébrales par scanner afin de comparer la qualité de segmentation entre des acquisitions DECT et SPCCT en présence d'implants métalliques, illustrant son adaptabilité au-delà du contexte pulmonaire. Ce travail a fait l'objet d'une publication dans le *Journal of Neurobiology*.

Malgré ces contributions, plusieurs limitations doivent être reconnues.

Concernant les annotations, VSMOD a principalement été conçu pour la segmentation de l'AAP en CTPA. Bien qu'un transfert préliminaire à l'imagerie cérébrovasculaire ait démontré que la stratégie utilisant RANSAC et la croissance de région n'est pas intrinsèquement limitée au territoire pulmonaire, la paramétrisation actuelle reste optimisée pour la géométrie des artères pulmonaires, caractérisée par des calibres relativement grands et une faible tortuosité. Un ajustement manuel demeure nécessaire pour des structures vasculaires aux propriétés différentes.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Par ailleurs, la génération automatique de graines dans le module de segmentation repose sur des valeurs de rayon local estimées par l'algorithme RANSAC. Dans les régions à faible contraste ou présentant des déformations pathologiques, ces estimations peuvent être inexactes, entraînant une initialisation erronée des graines qui nécessite une correction manuelle.

Concernant la segmentation, les modèles entraînés sur la cohorte PERSEVERE ont montré une transférabilité limitée à des jeux de données externes tels que PARSE. Cette limitation découle à la fois de différences liées à l'acquisition, notamment le fabricant du scanner et les protocoles de reconstruction, et de disparités liées à la pathologie, la cohorte PERSEVERE étant composée exclusivement de patients atteints d'EP avec une morphologie vasculaire altérée. Les modèles risquent ainsi d'apprendre des représentations liées à des patterns pathologiques et à des conditions d'acquisition spécifiques, ce qui limite leur généralisation. Par ailleurs, les jeux de données annotés publiquement disponibles manquent soit de cas pathologiques, soit restent limités en taille, et les tentatives d'exploiter des stratégies semi-supervisées telles que Mean Teacher n'ont pas conduit à des améliorations satisfaisantes. Pour la segmentation des thrombus, la variabilité d'apparence des défauts de remplissage hypodenses reste un défi.

Concernant la modélisation en graphe, le pipeline rencontre encore des difficultés en présence de déformations pathologiques sévères. Ces difficultés peuvent provenir de deux sources distinctes : des erreurs de segmentation en amont, telles que des parois vasculaires fusionnées ou des connexions parasites introduites par le modèle d'apprentissage profond, et des limitations des algorithmes de construction et de correction du graphe eux-mêmes, qui peuvent échouer à résoudre correctement des configurations topologiques ambiguës dans des arbres vasculaires fortement déformés.

L'application neurovasculaire demeure une preuve de concept basée sur un seul patient, ce qui démontre la faisabilité du transfert du pipeline à un territoire anatomique différent, mais ne constitue pas une validation de sa généralisabilité à des cohortes neurovasculaires plus larges. De plus, le pipeline actuel est conçu pour des réseaux vasculaires à topologie arborescente, traitant tous les cycles comme des artefacts. Cette hypothèse est valide pour l'AAP, mais limite l'applicabilité du pipeline à des réseaux vasculaires présentant des topologies différentes, comme le réseau cérébrovasculaire, qui contient le cycle anatomiquement valide du polygone de Willis.

Ces limitations ouvrent plusieurs directions pour les travaux futurs.

Du point de vue des annotations, deux directions complémentaires peuvent être envisagées. D'une part, VSMOD devrait être évalué sur des structures vasculaires au-delà du territoire pulmonaire afin d'évaluer sa généralisabilité. D'autre part, le module de segmentation par croissance de région pourrait être remplacé ou complété par un modèle d'apprentissage profond, ce qui améliorerait la robustesse dans les cas pathologiques et les images de faible qualité. Du point de vue de la segmentation, améliorer la généralisation entre les protocoles d'acquisition et les populations de patients nécessitera à la fois l'accès à des jeux de données multicentres et le développement de stratégies d'entraînement qui traitent explicitement les décalages de domaine et de concept. Pour la segmentation de l'artère pulmonaire, une approche prometteuse combine un pré-entraînement sur de larges jeux de données sains, comme AirRC (J. Liu et al. 2025), avec

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

un affinage sur les cas pathologiques de la cohorte PERSEVERE. Les outils de segmentation polyvalents récents, tels que la version mise à jour de TotalSegmentator, offrent également une base prometteuse pour le transfert et l'affinage. À plus long terme, des modèles fondationnels vasculaires (Wang, Meng, and Z. Xu 2025) entraînés sur plusieurs territoires anatomiques et modalités d'imagerie pourraient réduire le besoin de larges jeux de données annotés.

Du point de vue de la modélisation, évoluer vers des approches basées sur l'apprentissage et sensibles à la topologie représente une direction prometteuse. Des travaux récents, comme l'Implicit Point-Graph Network (Xie, Yang, et al. 2025; Xie, Y. Zhu, et al. 2026), proposent de modéliser conjointement la segmentation et la topologie, permettant une gestion plus robuste des branches déconnectées et des structures anatomiquement incohérentes. De telles approches pourraient améliorer la construction du graphe et permettre un étiquetage anatomique plus fin au niveau des artères segmentaires.

Du point de vue clinique, la représentation automatisée en graphe de l'AAP ouvre la voie à des études pronostiques à grande échelle. Les biomarqueurs extraits, incluant les scores de charge thrombotique et les distributions spatiales du volume embolique, fournissent une description riche de l'AAP des patients atteints d'EP qui pourrait améliorer la stratification du risque au-delà des scores cliniques existants. En particulier, l'organisation spatiale des caillots pourrait aider à différencier les patients présentant des volumes emboliques totaux similaires mais des évolutions cliniques différentes.

La disponibilité conjointe des représentations en graphe et des données cliniques ouvre des perspectives plus générales. La disponibilité de biomarqueurs cliniques tels que le BNP et la troponine dans le jeu de données PERSEVERE permet d'investiguer les corrélations entre les caractéristiques dérivées du graphe et les évolutions des patients. Enfin, la représentation en graphe est naturellement adaptée à l'utilisation de réseaux de neurones sur graphes (GNN) pour la prédiction des résultats cliniques. La cohorte de 353 graphes de patients constitue un jeu de données structuré dans lequel les arêtes encodent des caractéristiques des vaisseaux, de leur localisation et de l'EP. Les GNN peuvent exploiter cette structure pour capturer les interactions au sein de l'arbre vasculaire et apprendre des patterns prédictifs allant au-delà des biomarqueurs scalaires agrégés. Les travaux préliminaires (Paetzold et al. 2021; Vaida and Z. Huang 2026) menés dans des domaines similaires laissent entrevoir le potentiel de cette approche.

Cette thèse établit un cadre complet et automatisé pour l'analyse en graphe de l'AAP dans l'EP, de l'annotation à l'extraction de biomarqueurs cliniques. En fournissant une représentation structurée et quantitative de la charge thrombotique, elle pose les bases de futurs développements méthodologiques et d'aide à la décision clinique, et ouvre la voie vers des approches plus personnalisées et sensibles à la topologie pour l'analyse des maladies vasculaires.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Appendix

I Excluded PERSEVERE patients list by criteria

In this annex, we list each patient that were excluded for one of the three motives : (i) Manufacturer homogeneity, (ii) incomplete clinical blood biomarkers, and (iii) image quality.

I.1 Different manufacturer

The patient removed are 0016 (GE MEDICAL SYSTEMS BrightSpeed QX/i); 0045, 0077 (SIEMENS Sensation 16); 0092 (GE MEDICAL SYSTEMS Optima CT540); 0141 (Philips Brilliance 40); 0149, 0160 (SIEMENS Sensation 16); 0212 (GE MEDICAL SYSTEMS Revolution EVO); 0236, 0239 (GE MEDICAL SYSTEMS Optima CT660); 0262, 0287 (GE MEDICAL SYSTEMS Revolution EVO); 0306, 0373, 0400 (SIEMENS Sensation 16).

I.2 Incomplete clinical blood biomarkers

The patient removed here are 0004, 0019, 0020, 0021, 0028, 0048, 0052, 0054, 0069, 0070, 0075, 0091, 0113, 0117, 0121, 0151, 0166, 0186, 0218, 0226, 0228, 0231, 0245, 0285, 0286, 0292, 0295, 0298, 0301, 0302, 0316, 0318, 0329, 0342, 0369, 0377, 0379, 0401, 0428.

I.3 Image quality

The exclusion criteria on the motif of image quality concerns unusable scans, truncated heart or lungs, absence of visible emboli, missing arterial phase, severe imaging artefact, damaged lungs or abnormalities inside lungs making arteries hard to be detected. The detail of the number of patient in each cases is described after :

- Two patients were removed because no mediastinal image was available (0097, 0431), and one because the image is truncated (0041).
- Ten patients had a difficult diagnosis for the radiologist; no PE was observed in the image (0002, 0189, 0229, 0276, 0314, 0321, 0338, 0376, 0393, 0396).
- One patient had no visible arterial time on the image (0146).
- Four patients had too many artefacts, caused by motion, beam hardening, or linked to the injection 0027, 0035, 0114, 0195.
- Four patients were removed for a very damaged lung, or abnormalities present next to or inside the lung (0015, 0143, 0252, 0381).
- Two patients (0412, 0424) had a large truncated part on the segmentation mask we obtained later on due to the difficulty in distinguishing arteries because of low contrast in the surrounding tissues.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

I.4 Lung and lobe segmentation

We require accurate segmentation of the lobe for the PERSEVERE dataset. We observe that although the left and right lung mask are correct with the `lungmask` model, it fails on this non-healthy subject dataset to segment the lobes, a task where `TotalSegmentator` is efficient (see Figure B.1).

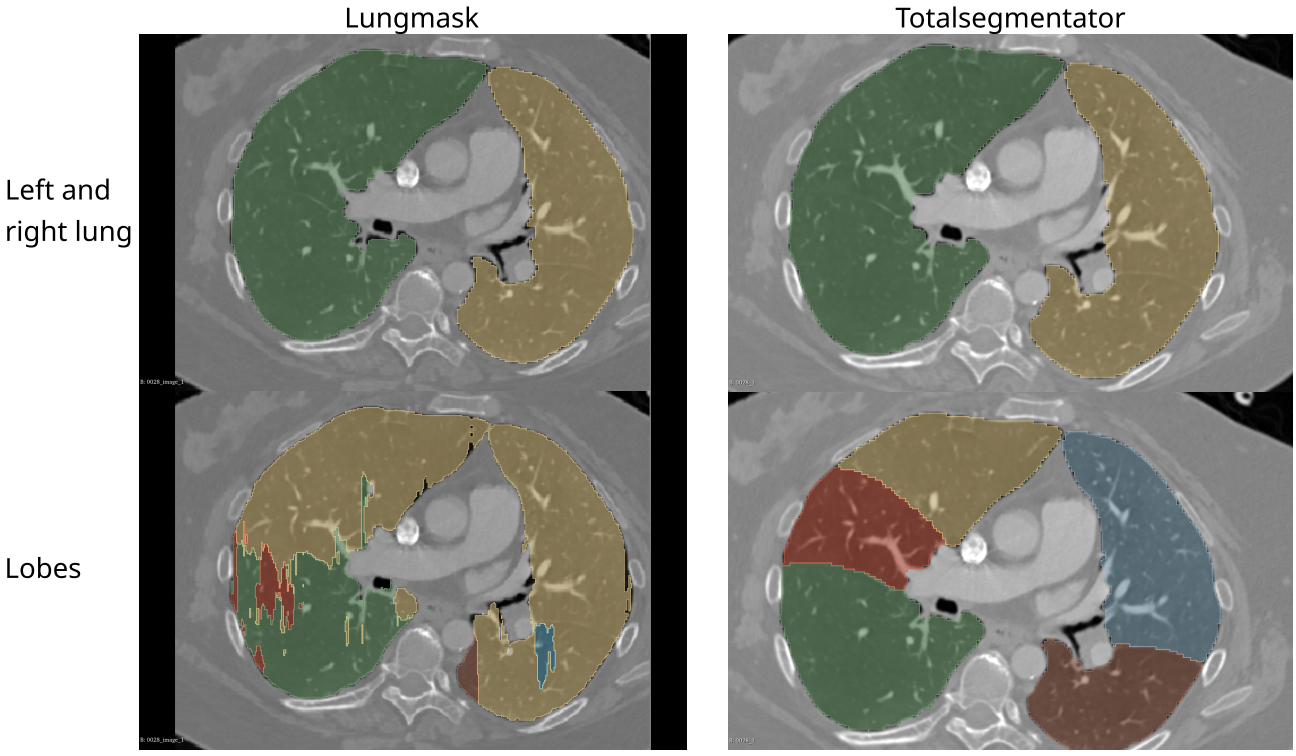


Fig. B.1: Illustration of the lungs and lobes segmentation from Lungmask and Totalsegmentator on patient 0028 from the PERSEVERE dataset

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

II PERSEVERE dataset reconstruction kernels

The two included scanner models differ in their reconstruction kernels. The SOMATOM Definition AS predominantly uses iterative reconstruction kernels (I40f in 96.9% of cases, I30f in 2.9%), whereas the SOMATOM Definition exclusively uses filtered back-projection kernels (B40f: 62.1%, B30f: 20.7%, B70f: 17.2%) (see Table B.1).

Scanner	**Kernel**	**Number of patient (%)**
SOMATOM Definition AS	I40f / thickness 3	372 (96.6)
	I30f / thickness 3	10 (2.6)
	I30f / thickness 2	1 (0.3)
	I70f / thickness 3	1 (0.3)
	I40f / thickness 2	1 (0.3)
SOMATOM Definition	B40f	18 (62.1)
	B30f	6 (20.7)
	B70f	5 (17.2)

Tab. B.1: Distribution of reconstruction kernels by SOMATOM Definition AS and SOMATOM Definition scanner model.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

III PERSEVERE statistics

The descriptive statistics of the resulting arterial segmentations are summarized in Table B.2. Pulmonary artery volumes range from 55.26 cm^3 to 295.37 cm^3, with a mean of 133.28 cm^3 and a median of 127.71 cm^3. All 353 patients present a β_0 value of exactly 1, confirming that the post-processing successfully enforced a single connected arterial tree in every case. A β_1 value greater than zero, indicating the presence of topological cycles in the segmentation, is observed in 111 patients (31.4%), with values ranging from 1 to 5. No enclosed cavities were detected in any patient ($\beta_2 = 0$ uniformly). The handling of residual cycles is described in Chapter 6.

Statistic	**Volume (cm^3)**	**Voxel Count**	β_0	β_1	β_2
Minimum	55.26	211,636	1	0	0
Maximum	295.37	1,131,198	1	5	0
Mean	133.28	510,416	1.00	0.47	0.00
Median	127.71	489,112	1.00	0.00	0.00
Std. Dev.	39.89	152,783	0.00	0.85	0.00
Patients > 0	–	–	353 (100%)	111 (31.4%)	–

Tab. B.2: Descriptive statistics of arterial segmentation volumes and topological Betti numbers across the full dataset ($n = 353$ patients).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

IV PARSE statistics

Table B.3 reports descriptive statistics on pulmonary artery segmentation volumes and topological Betti numbers across the 90 CTPA training cases of the PARSE dataset. Pulmonary artery volumes range from 80.94 cm^3 (min) to 357.7 cm^3 (max), the mean volume is 140.69 cm^3. 100% of patient have a $\beta_0 = 1$ which confirms that all annotations of the arterial trees form a single connected component. 88.9% have a $\beta_1 > 1$ which means that a majority of patients present segmentations with non-topological cycles. And 45.6% of patients have $\beta_2 > 0$ meaning that the segmentations present cavities.

Statistic	**Volume (cm^3)**	**Voxel Count**	β_0	β_1	β_2
Minimum	80.94	126,821	1	0	0
Maximum	357.70	568,109	1	91	205
Mean	140.69	308,616	1.00	12.76	3.69
Median	130.34	297,517	1.00	3.00	0.00
Std. Dev.	41.96	78,239	0.00	20.91	22.16
Patients > 0	–	–	90 (100%)	80 (88.9%)	41 (45.6%)

Tab. B.3: Descriptive statistics of arterial segmentation volumes and topological Betti numbers across the PARSE dataset($n = 90$ patients).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

V Cerebral Artery Segmentation 2023 challenge

V.1 CAS dataset

Although the CAS2023 challenge focuses on cerebral artery segmentation rather than pulmonary arteries, we participated in it at the beginning of this thesis as an opportunity to gain hands-on experience with deep learning segmentation models before tackling the more complex pulmonary arterial tree. This work is described in Chapter 5, section V.3.

The Cerebral Artery Segmentation (CAS) challenge was part of MICCAI 2023. This challenge focuses on automatic segmentation of cerebral arteries from Time-of-Flight Magnetic Resonance Angiography (TOF-MRA) images of patients with symptomatic artery stenosis. The aim is to help the diagnosis of cerebral vascular disease, such as arteriosclerosis, dissection, arteritis, moyamoya disease, and reversible cerebral vasoconstriction syndrome, with the identification and characterization of stenosis.

The dataset includes 100 3D TOF-MRA volumes (training set) with cerebral artery annotations (14349 slices total), publicly available in NIfTI format. The dimensions of this dataset are [208×320×96] to [784×784×255] voxels (median: 640×640×150), with isotropic 1.0 mm voxel spacing. (see Figure B.2).

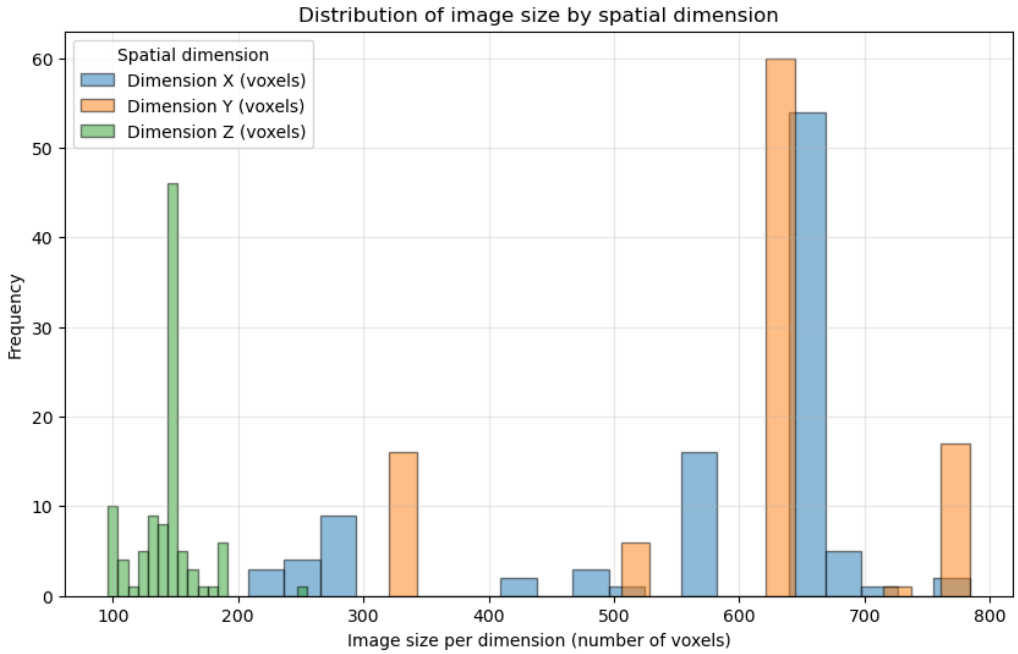


Fig. B.2: Shape distribution on the 100 CAS patients

Figure B.3 illustrates an example of an image and an annotation on the first patient of the CAS training dataset.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

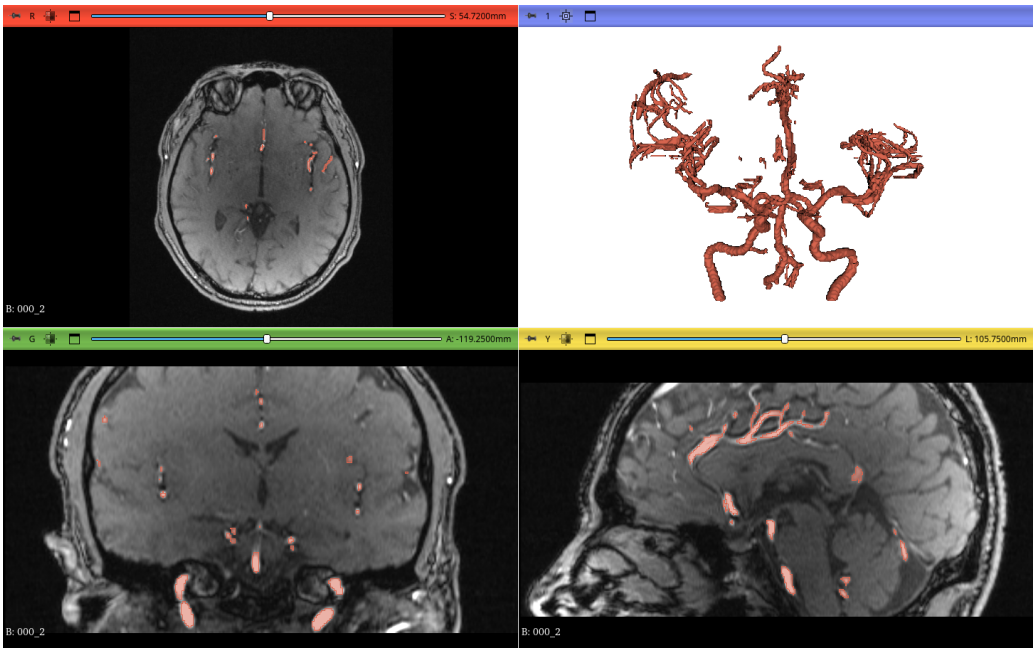


Fig. B.3: Representative example of a TOF MRA volume with its corresponding cerebral artery annotations for patient 000 from the CAS training dataset, displayed in 3D Slicer.

V.2 CAS statistics

Table B.4 reports descriptive statistics on cerebral artery segmentation volumes and topological Betti numbers across the 100 3D TOF-MRA volumes training cases of the CAS dataset. Cerebral artery volumes range from 4.91 cm^3 (min) to 16.88 cm^3 (max), with mean 9.49 cm^3 (see Table B.4). Unlike pulmonary arteries, 100% of patients have $\beta_0 > 1$ (mean 57 and median 30 connected components), indicating severe fragmentation. 100% have $\beta_1 > 0$ (mean 75 circles) reflecting more circles than the only one expected, *ie.* the circle of Willis, and 100% have $\beta_2 > 0$ (mean 49 cavities).

Statistic	**Volume (cm^3)**	**Voxel Count**	β_0	β_1	β_2
Minimum	4.91	17,034	5	21	1
Maximum	16.88	331,294	307	217	284
Mean	9.49	158,279	57.46	75.57	49.19
Median	9.00	170,061	30.00	70.50	34.00
Std. Dev.	2.61	74,500	63.31	37.37	49.45
Patients > 0	–	–	100 (100%)	100 (100%)	100 (100%)

Tab. B.4: Descriptive statistics of cerebral artery segmentation volumes and topological Betti numbers across the CAS2023 dataset ($n = 100$ patients).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

V.3 Preliminary training experiments on vascular segmentation (CAS2023)

We participated in the MICCAI 2023 CAS challenge, which focuses on automatically segmenting cerebral arteries from TOF-MRA images to aid in diagnosing vascular diseases like arteriosclerosis and stenosis (see details of the Challenge dataset in chapter 2 I.2). A training set of 100 annotated patients is provided. The challenge evaluation was done using several metrics: the Dice (II.1.4), the Average Hausdorff Distance (AHD), the Dice in the stenosis region, and the AHD in the stenosis region. The ranking is done with a score calculated as :

$$\text{Final score} = 35\%\ \text{DSC} + 35\%\ \text{AHD} + 15\%\ \text{DSCstenosis} + 15\%\text{AHDstenosis}. \quad \text{(B.1)}$$

V.3.1 Preprocessing

The images were all resampled to the median spacing of the cohort 0,28125 mm x 0,28125 mm x 0,6 mm. Then, brain masks were created with the FSL brain extraction tool HD-BET Smith et al. 2004 and a fractional intensity threshold of 0.1. Finally, with the help of the brain mask, a bounding box was created around the brain with an X-axis border of 30 pixels, a Y-axis border of 10 pixels, and a Z-axis border of 10 pixels (except for the end of the border). The preprocessing is illustrated for patient 000 of the training step (Figure. B.4).

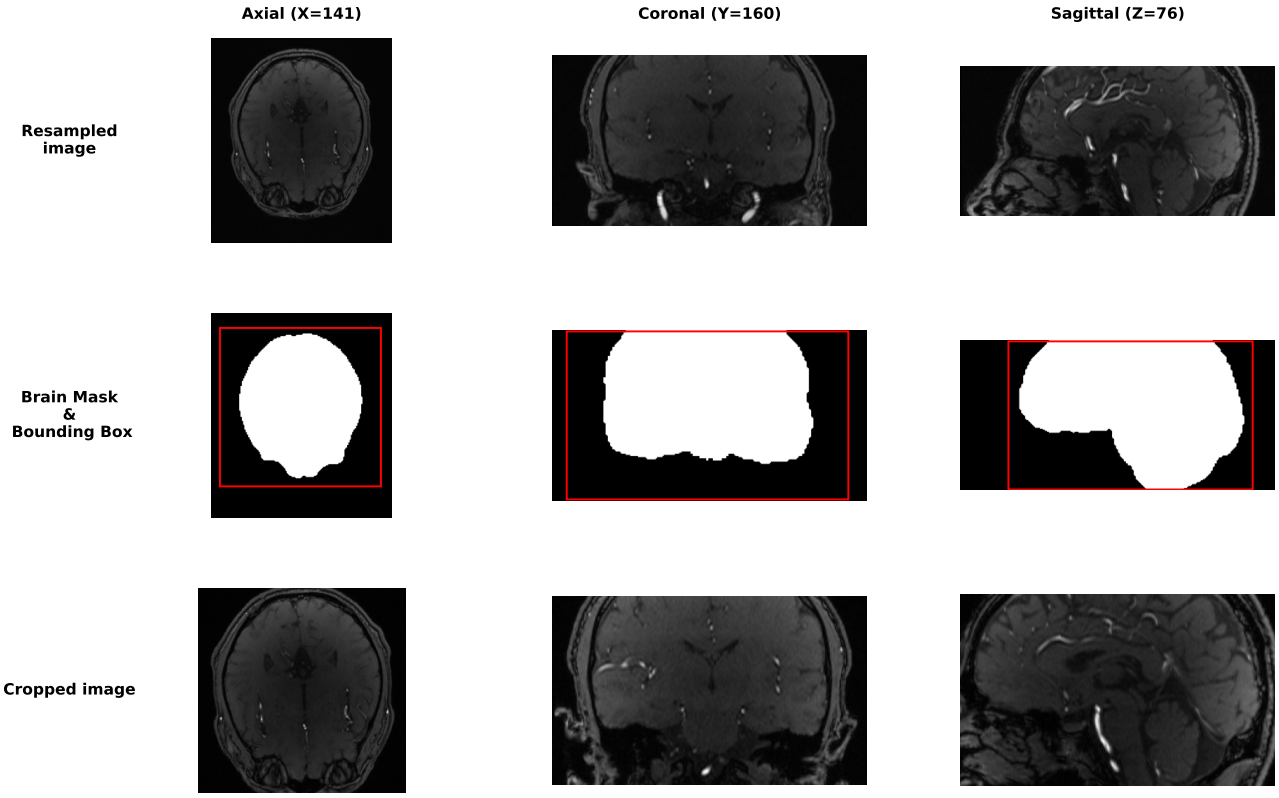


Fig. B.4: Illustration of the preprocessing on the patient 0000 of the training set.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

V.3.2 Model architectures

We choose to train several UNet-based models and then try ensembling techniques. The models trained were a TinyUNet, a BasicUNet (see I.1.3), and a cascadedUNet (see I.1.5).

The TinyUNet is a lighter version of the U-Net with 4 encoder-decoder levels. The BasicUNet is the standard UNet implementation, with 6 encoder-decoder levels, serving as a baseline for medical image segmentation. The cascaded UNet consists of two UNets: one predicts the segmentation, the other predicts the vessel skeleton. The skeleton captures topological information to help regularize the segmentation output.

V.3.3 Training configuration

All models were trained with patch size 256 x 256 x 64 pixels. The batch size was fixed to 2, and the number of batches per epoch was 250. Models were trained wih 500 epochs. It was a combined dice and cross-entropy loss. The learning rate was fixed to 10^{-3}. The optimizer was the SGD with Nesterov momentum. The validation was done using a 5-fold cross-validation strategy on 80 of the 100 patients. Each model was trained using data augmentation inspired by nnUnet, with spatial, intensity, and geometric modifications. The spatial modifications include rotation of more or less 30°, scaling between 0.7 and 1.4, and elastic deformations. The intensity changes include Gaussian noise, blurring, brightness, contrast, gamma, and low-resolution simulation. The geometric operations are mirror flips on all axes.

V.3.4 Results of ensembling

The results of the different ensembling strategies are presented in Table B.5. We tested (i) Single model ensembling, which is the average of each model's K-fold, (ii) cross-model per-fold ensembling, where each K-fold of the three models is averaged, and (iii) cross-model per-model ensembling, where each Single model ensembling is averaged. It is the single model ensembling of the cross-model per-fold ensembling that gave the best DICE performance. Using this models, we then applied the post-processing.

V.3.5 Post-processing

The post-processing we chose was a deletion of small components under a minimum pixels size, with the function *remove_small_object* from the library `skimage.morphology`. We tested several sizes, 50, 100 and 150 and the best mean dice was with the removed component of under 100 pixels (Table B.6).

V.3.6 Results of the challenge

Once the training model and post-processing strategy were decided, we relaunched the training of all the models with this time the whole training set of the 100 patients, so that we hope we can achieve the best segmentation possible on the blind test set of 50 patients, given for the evaluation procedure. The challenge final ranking is presented in table B.7, and we ended up in 16th place.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Model	Category	Mean Dice ± Std	Mean AHD± Std
TinyUNet			
	Fold 0	0.8449 ± 0.0479	85.74 ± 25.07
	Fold 1	0.8384 ± 0.0753	89.47 ± 30.80
	Fold 2	0.8503 ± 0.0494	90.83 ± 30.39
	Fold 3	0.8474 ± 0.0467	94.95 ± 29.38
	Fold 4	0.8497 ± 0.0495	98.05 ± 33.05
	TinyUNet Ensembling (5 folds)	0.8522 ± 0.0463	87.52 ± 27.83
BasicUNet			
	Fold 0	0.8470 ± 0.0480	90.78 ± 28.04
	Fold 1	0.8370 ± 0.0885	84.34 ± 25.74
	Fold 2	0.8508 ± 0.0478	85.80 ± 26.54
	Fold 3	0.8455 ± 0.0493	90.80 ± 26.05
	Fold 4	0.8499 ± 0.0490	88.23 ± 26.02
	BasicUNet Ensembling (5 folds)	0.8522 ± 0.0474	82.58 ± 23.74
CascadedUNet			
	Fold 0	0.8309 ± 0.0601	101.01 ± 28.65
	Fold 1	0.8321 ± 0.0759	104.78 ± 26.15
	Fold 2	0.8457 ± 0.0510	89.77 ± 25.77
	Fold 3	0.8264 ± 0.0558	87.83 ± 23.37
	Fold 4	0.8373 ± 0.0526	111.53 ± 28.32
	CascadedUNet Ensembling (5 folds)	0.8436 ± 0.0503	92.73 ± 24.36
Tiny+Basic+CascadedUNet per fold ensembling			
	Fold 0	0.8462 ± 0.0493	86.11 ± 25.90
	Fold 1	0.8399 ± 0.0770	88.72 ± 28.44
	Fold 2	0.8523 ± 0.0482	83.61 ± 28.93
	Fold 3	0.8490 ± 0.0460	85.85 ± 24.87
	Fold 4	0.8510 ± 0.0482	88.84 ± 26.10
	Tiny+Base+CascadedUNet per fold ensembling (5 folds)	**0.8530 ± 0.0467**	**85.55 ± 26.02**
Tiny+Basic+CascadedUNet per model			
	Tiny+Basic+CascadedUNet per model ensembling (3 ensembling)	0.8524 ± 0.0469	85.33 ± 26.21

Tab. B.5: Results Summary for Single Models folds, Single-Model Ensembling, Multi-model per fold ensembling, and Multi-model per-model Ensembling on 20-patient test set

Note: All results evaluated on 20-patient hold-out test set. Single-model ensembling average predictions across 5 folds of the same model. Cross-model per-fold ensembling: for each fold, average predictions from TinyUNet, BasicUNet, and CascadedUNet trained on that K-fold. Cross-model per-model ensembling: averaging the 3 single-model ensembles.

Configuration	Mean Dice ± Std	Mean AHD ± Std
No post-processing	0.8530 ± 0.0467	85.55 ± 26.02
Min size = 50 pixels	0.8533 ± 0.0466	81.60 ± 25.85
Min size = 100 pixels	**0.8538 ± 0.0465**	**82.47 ± 25.53**
Min size = 150 pixels	0.8536 ± 0.0465	81.48 ± 26.14

Tab. B.6: Post-Processing Optimization: Impact of Small Object Removal on Tiny+Base+WNet Ensemble

Note: Post-processing applied to Tiny+Basic+CascadedUNet per fold ensembling. Small connected components below threshold are removed to improve connectivity.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Rank	Team Name	SCORE	DSC	DSCstenosis	AHD	AHDstenosis
1	PhiWCH	0.88284	0.8760	0.8893	0.8976	0.8483
2	SVCC	0.88276	0.8769	0.9012	0.8902	0.8584
3	BIT-BUAA-BNU	0.88250	0.8754	0.8938	0.8957	0.8504
4	BITer	0.88180	0.8772	0.9032	0.8888	0.8515
5	Infervision	0.88170	0.8770	0.9004	0.8906	0.8501
6	SZTU-IME	0.88168	0.8777	0.8999	0.8897	0.8493
7	GZHU	0.88143	0.8780	0.8987	0.8895	0.8488
8	blackbean	0.88138	0.8757	0.8956	0.8923	0.8497
9	UW	0.88010	0.8765	0.8999	0.8874	0.8482
10	XB	0.87980	0.8737	0.8961	0.8931	0.8400
11	Davide Resta	0.87960	0.8637	0.8840	0.9047	0.8441
12	NexToU	0.87800	0.8745	0.8983	0.8840	0.8472
13	BUAA_MI	0.87620	0.8714	0.8918	0.8874	0.8415
14	sicong	0.87490	0.8693	0.8770	0.8914	0.8365
15	UWUT	0.86880	0.8642	0.8915	0.8757	0.8455
16	**CREATIS-MYRIAD**	**0.8660**	**0.8648**	**0.8835**	**0.8726**	**0.8222**
17	Zehan Zhang	0.84750	0.8496	0.8756	0.8412	0.8236
18	Theresa Nolte	0.82750	0.8215	0.8522	0.8366	0.7865
19	Los Moritos	0.78760	0.8332	0.8085	0.7418	0.7430
20	SKMR	0.75250	0.8093	0.8762	0.6267	0.8101
21	Chinese Academy of Sciences Brain Imaging Platform	0.6492	0.6802	0.7161	0.6121	0.5776

Tab. B.7: Team rankings and performance metrics of CAS 2023 challenge.

During the training of our models, we focused on the Dice metric to evaluate performance. Comparing our Dice score to that of the challenge winner, the difference is 0.0112. Looking at the overall score, the difference is 0.01684. For our first experience with this type of challenge, we are relatively satisfied with achieving a result reasonably close to the best performance.

We intended to examine the winning challenge participant's architectural choice and evaluate their solution; however, as the challenge organizers stated that there were no publication requirements and that participating teams were free to publish their own results independently. The challenge organizer stated that no challenge paper was planned for publication. As a result, the winning team's methodology was never formally documented in a publicly available publication, and we were therefore unable to evaluate or compare against their solution in detail.

Nevertheless, this experiment provided valuable experience in training vascular segmentation models. While CAS2023 focuses on cerebral vasculature, the developed training pipeline can be transferred to other vascular territories. We therefore applied similar training strategies to pulmonary artery segmentation using the PARSE dataset.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

V.4 Training efficiency

The training comparison between the four models is reported in Table B.8.

Metric	Model 1	Model 2	Model 3	Model 4
Best Val Dice	0.7451	0.7772	0.7584	0.6322
Num Epoch	401	303	300	200
Epochs to Val Dice > 0.70	233	72	**5**	Not reached
Time per Epoch	5.7 min	19.0 min	19.2 min	1.48 min
Total Training Time	38.1 h	96 h	96 h	4.93 h
Training Data	PARSE only	PARSE + PERSEVERE	PARSE + PERSEVERE	PERSEVERE only
Initialization	Random	Random	Pre-trained	Random
Train/Val Split	72/18	72/18	72/18	6/2

Tab. B.8: Training efficiency comparison across four models

Models 1, 2, and 3 were all trained for approximately 300 to 400 epochs on PARSE data, while Model 4 was trained for only 200 epochs on the much smaller PERSEVERE subset, resulting in a total training time of less than 5 hours. Models 2 and 3 required approximately 19 minutes per epoch due to the dual-network Mean Teacher architecture processing both labeled and unlabeled images simultaneously, compared to 5.7 minutes per epoch for Model 1. In terms of validation, Dice on their respective training domains, Model 2 achieved the highest value (0.7772), closely followed by Model 3 (0.7584) and Model 1 (0.7451), while Model 4 reached only 0.6322 on its two-patient validation set. The pretraining strategy used in Model 3 had an effect on convergence speed: by initializing the student network from the supervised baseline of Model 1, Model 3 reached a validation Dice above 0.70 in only 15 epochs, compared to 76 epochs for Model 2 and 329 epochs for Model 1. Even though the supervised initialization accelerates semi-supervised training, the final validation performance of Models 2 and 3 remains comparable.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

VI Advanced parameters for RANSAC: the threshold

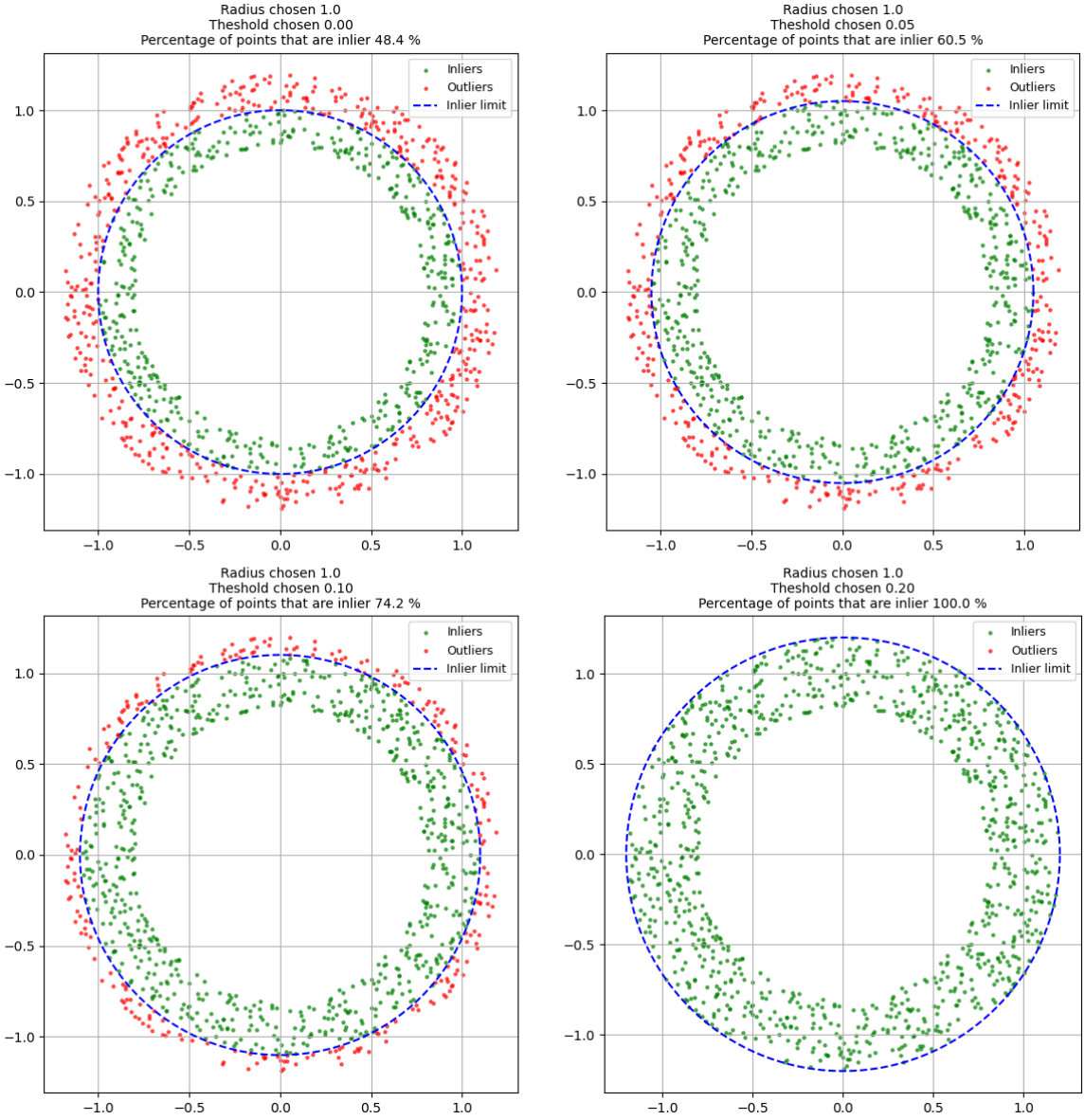


Fig. B.5: Illustration of a point cloud between 0.8 and 1.2mm from a cylinder axis, and the impact of the *Threshold (%)* parameter on the *Inlier Points (%)* parameter. Figure done by Gabriel Husak

We illustrated the impact of the threshold in Figure B.5. This is an example of how many points are considered inliers in a simulated point cloud spread between 0.8 and 1.2mm around a chosen radius of 1mm. In this example, with a 10% threshold, we obtain 60.5% of inliers, which passes our default Inlier Points (%) parameter of 60%, and the cylinder fitting is validated.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

VII Details of the quality check during the hierarchical labeling of the PAT

Unknown Edges and Connectivity Issues Failures in the automated labeling of the proximal tree (levels 1-2) were primarily attributed to:

- **Morphological Irregularities:** Irregularities in the Main Pulmonary Artery (MPA) mask (e.g., Patients 0039, 0342) created artifacts that prevented the root from being correctly identified.
- **Segmentation Discontinuities:** "Holes" or gaps in the arterial mask (e.g., Patients 0170, 0179, 0319, 0406) led to incorrect skeletonization, which often broke the MPA centerline, preventing the propagation of labels from the MPA to the distal branches.

Where possible, a "quick fix" was applied using simple techniques to refine the masks (e.g., smoothing the MPA or closing small gaps in the arterial mask). However, cases with severe mask degradation due to damaged lung parenchyma rendering the image unusable (e.g., Patient 0252) were excluded from the final analysis.

Missing Lobes and Pathological Challenges The absence of labels within specific lobes was cross-referenced with the original CT scans to distinguish between algorithmic failures and clinical reality:

- **Anatomical Justification:** In several cases, a lobe was correctly identified as having no arterial edges because it was severely damaged or collapsed (e.g., Patient 0011 LLL), or because the vessels were too small to be discerned even manually (e.g., Patient 0123 RML). These cases were retained to maintain anatomical fidelity.
- **Segmentation Failures:** Missing lobar branches were often caused by poor contrast injection timing (e.g., Patient 0302) or massive pulmonary emboli (EP) completely obstructing the arterial lumen (e.g., Patient 0412).
- **Exclusion Criteria:** Patients were removed from the cohort if the lobe/artery segmentations were deemed too poor for reliable analysis (e.g., Patients 0015, 0143, 0381, 0424), or if the lobar segmentation itself was incorrect due to large masses or atypical anatomy (e.g., Patient 0318).

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés

FOLIO ADMINISTRATIF

THESE DE L'INSA LYON, MEMBRE DE L'UNIVERSITE DE LYON

NOM : DES LIGNERIS
Prénoms : Morgane, Maud, Catherine

DATE de SOUTENANCE : 21/05/2026

TITRE (EN) : *Pulmonary embolism risk stratification based on vascular network modeling*
TITRE (FR) : *Stratification du risque de l'embolie pulmonaire basée sur la modélisation du réseau vasculaire*

NATURE : Doctorat

Numéro d'ordre : 2026ISAL0031

École Doctorale : Électronique, Électrotechnique et Automatique de Lyon (N° 160)

Spécialité : Traitement du signal et de l'image

RÉSUMÉ (EN) :

Pulmonary Embolism (PE), the third most frequent acute cardiovascular syndrome, results from the obstruction of pulmonary arteries by blood clots. Assessing its severity requires characterizing the location, extent, and distribution of thrombi within the Pulmonary Arterial Tree (PAT). Current clinical scores, such as Qanadli and Mastora, rely on simplified anatomical templates and time-consuming manual measurements, limiting their routine use. Additionally, key cardiac biomarkers required for risk stratification are sometimes unavailable in emergency settings. This thesis addresses these limitations by developing an automated pipeline for graph-based modeling of the PAT from Computed Tomography Pulmonary Angiography (CTPA), enabling the extraction of patient-specific imaging-derived biomarkers for PE severity assessment. This work was conducted within the PERSEVERE project, using a cohort of 353 CTPA acquisitions. Four main contributions are presented. First, VSMOD, an open-source 3D Slicer plugin, enables semi-automated vascular annotation, reducing annotation time by approximately 80% while preserving topological consistency. Second, a deep learning-assisted iterative annotation strategy compensates for the lack of annotated data, producing robust segmentation models for pulmonary arteries and emboli across the full cohort. Third, a fully automated graph-based pipeline converts vascular and thrombus masks into an anatomically oriented directed graph, from which local and global biomarkers are extracted, including automated adaptations of existing thrombotic burden scores showing strong agreement with manual measurements. Fourth, the framework is applied to neurovascular imaging, demonstrating its broader applicability by quantitatively comparing vessel segmentation between Dual-Energy CT and Spectral Photon Counting CT. Together, these contributions establish a complete, reproducible workflow from raw CTPA data to structured clinical biomarkers, laying the foundation for automated, patient-specific PE severity assessment and future prognostic studies.

RÉSUMÉ (FR) :

L'embolie pulmonaire (EP), troisième syndrome cardiovasculaire aigu le plus fréquent, nécessite une évaluation précise de la charge thrombotique et de sa distribution au sein de l'arbre artériel pulmonaire (AAP). Les scores cliniques existants, tels que ceux de Qanadli et de Mastora, restent limités par leur caractère manuel et peu reproductible, tandis que les biomarqueurs cardiaques nécessaires à la stratification du risque peuvent être indisponibles dans les situations d'urgence. Cette thèse propose un pipeline automatisé de modélisation de l'AAP par graphes à partir de l'angiographie pulmonaire par scanner, permettant l'extraction de biomarqueurs d'imagerie spécifiques au patient. Ces travaux s'inscrivent dans le cadre du projet PERSEVERE, à partir d'une cohorte de 353 patients.Quatre contributions principales sont présentées. Premièrement, VSMOD, un plugin open source pour 3D Slicer, permet une annotation semi-automatisée des structures vasculaires, réduisant le temps d'annotation d'environ 80% tout en préservant la cohérence topologique. Deuxièmement, face au manque d'annotations et à un décalage de domaine identifié entre données saines et pathologiques, une stratégie d'annotation itérative assistée par apprentissage profond a permis d'entraîner des modèles nnU-Net robustes pour la segmentation des artères pulmonaires et des emboles sur l'ensemble de la cohorte. Troisièmement, un pipeline de modélisation par graphes transforme les masques vasculaires et emboliques en un graphe orienté anatomiquement cohérent, encodant pour chaque arête la morphologie vasculaire, la localisation hiérarchique et le degré d'obstruction. Ce pipeline permet le calcul automatique d'adaptations des scores de Qanadli et de Mastora, fortement corrélées aux évaluations manuelles, ainsi que la distribution spatiale du volume embolique. Quatrièmement, le cadre de modélisation est appliqué à l'imagerie neurovasculaire pour comparer quantitativement les capacités de segmentation vasculaire entre scanner à double énergie et scanner à comptage photonique spectral, démontrant la généricité de l'approche. L'ensemble de ces contributions établit un workflow complet et reproductible, des données brutes d'imagerie jusqu'à des biomarqueurs cliniquement interprétables, posant les bases d'une évaluation automatisée et personnalisée de la sévérité de l'EP et d'études pronostiques futures.

MOTS-CLÉS (EN) : pulmonary embolism ; machine learning ; segmentation
MOTS-CLÉS (FR) : embolie pulomnaire ; apprentissage automatique ; segmentation

Laboratoire(s) de recherche : Centre de Recherche en Acquisition et Traitement de l'Image pour la Santé (CREATIS)

Directrice de thèse : Carole Frindel
Co-Directrice de thèse : Odyssée Merveille

Présidente du Jury : Maria A. Zuluaga
Composition du Jury : Pierre-Henri Conze, Hervé Delingette, Laurent Bertoletti, Joseph Jacob.

Thèse accessible à l'adresse : https://theses.insa-lyon.fr/publication/2026ISAL0031/these.pdf © [M. Morgane Des Ligneris], [2026], INSA Lyon, tous droits réservés