A graph-cut approach for pulmonary artery-vein segmentation in noncontrast CT images

Daniel Jimenez-Carretero[a,1,1], David Bermejo-Peláez[a,1], Pietro Nardelli[b], Patricia Fraga[c], Eduardo Fraile[c], Raúl San José Estépar[b], Maria J Ledesma-Carbayo[a,*]

[a]Biomedical Image Technologies, Universidad Politécnica de Madrid & CIBER-BBN, Madrid, Spain

[b]Applied Chest Imaging Laboratory, Brigham and Womens' Hospital, Boston, Massachusetts, United States

[c]Unidad Central de Radiodiagnóstico, Madrid, Spain

Abstract

Lung vessel segmentation has been widely explored by the biomedical image processing community; however, the differentiation of arterial from venous irrigation is still a challenge. Pulmonary artery-vein (AV) segmentation using computed tomography (CT) is growing in importance owing to its undeniable utility in multiple cardiopulmonary pathological states, especially those implying vascular remodelling, allowing the study of both flow systems separately. We present a new framework to approach the separation of tree-like structures using local information and a specifically designed graph-cut methodology that ensures connectivity as well as the spatial and directional consistency of the derived subtrees. This framework has been applied to the pulmonary AV classification using a random forest (RF) pre-classifier to exploit the local anatomical differences of arteries and veins. The evaluation of the system was performed using 192 bronchopulmonary segment phantoms, 48 anthropomorphic pulmonary CT phantoms, and 26 lungs from noncontrast CT images with precise voxel-based reference standards obtained by manually labelling the vessel trees. The experiments reveal a relevant improvement in the accuracy (~ 20%) of the vessel particle classification with the proposed framework with respect to using only the pre-classification based on local information applied to the whole area of the lung under study. The results demonstrated the accurate differentiation between arteries and veins in both clinical and synthetic cases, specifically when the image quality can guarantee a good airway segmentation, which opens a huge range of possibilities in the clinical study of cardiopulmonary diseases.

Keywords

Artery-vein segmentation; Lung; Graph-cuts; Random forest; Arteries; Veins; Noncontrast CT; Phantoms

[1] Corresponding authors. daniel.jimenezc@die.upm.es (D. Jimenez-Carretero), mledesma@die.upm.es (M.J. Ledesma-Carbayo). [1]These authors contributed equally to this work and are considered co-first authors.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.media.2018.11.011.

1. Introduction

Computed tomography (CT) technology has been improving during the last decades and currently it is possible to obtain near-isotropic, sub-millimetre resolution acquisition of the complete chest in a single breath-hold, avoiding partial volume effects and breathing artefacts. For that reason CT has become the reference modality in pulmonary imaging and the radiological study of complex biological structures such as pulmonary vessel trees (Hsieh, 2009; Sluimer et al., 2006).

Although general vessel segmentation in lungs has been widely explored (Van Rikxoort and Van Ginneken, 2013), the differentiation of arterial from venous irrigations is still an open problem. The main drawbacks found in AV segmentation in CT include: extreme complexity and density of vessel tree structure, arteries and veins being located in close proximity (often intertwined), similar intensity values that make both irrigations indistinguishable (especially in noncontrast CT), and difficulties in differentiation between branching points (bifurcations) and vessel crossings.

AV segmentation would be useful in different medical scenarios and multiple diseases or pathological states. It would allow the study of arterial and venous irrigations separately, displaying the complex anatomical relationship between pulmonary lesions and vasculature, which is critical in oncology studies and management of cancer treatments, for example. Moreover, it would provide additional valuable information for surgeons to evaluate the arterial and venous conduits for vascular surgery. Features such as density, structure, and size of vessels could be studied in diseases that imply vasculature remodelling. This is the case for chronic obstructive pulmonary disease (COPD) and especially pulmonary arterial hypertension (PAH), where processes such as growth, cellular apoptosis, inflammation, and fibrosis contribute to changes in arterial structure (Crosswhite and Sun, 2014). The study of these features would allow physicians to diagnose and follow the pathological state of patients, and aid in the decision making for treatment. Additionally, AV information would ease and speed up the design and test of drugs in clinical trials (monitoring the effects and efficiency) for this kind of disease. A segmentation of the two vessel trees could also be of interest in computer-aided detection (CAD) systems. In the particular case of CAD for detection of pulmonary embolism, the identification and discarding of the whole venous tree would dramatically decrease the number of erroneous detections outside the arterial tree.

Nevertheless, AV segmentation has been clinically infeasible despite its objective usefulness. The huge complexity of pulmonary vascular trees makes a manual segmentation of both structures within a reasonable time frame unrealistic. Thus, the interest in the design of automatic or semiautomatic tools is undeniable. However, only a few image processing methods have been published for AV segmentation of CT images and most of them require user interaction for initialization or guiding, and in general they do not demonstrate their usefulness with a proper validation scheme.

For these reasons, the number of studies using pulmonary AV information is very limited, and therefore much knowledge about the relationship between cardiopulmonary disease and vessel geometry, topology, branching patterns, and other aspects remains hidden. The potential opportunity to fill this gap by creating new biomarkers or extracting measurements from analysis or statistics is promising, and it would be accessible through an automatic AV segmentation tool.

1.1. Background and state of the art on AV segmentation

Despite the difficulty in classifying pulmonary vessels in CT images, several contributions to this problem have been proposed during recent decades, which have relied on two main sources of information: prior knowledge and vessel connectivity.

Biological knowledge and CT image-based information can be exploited to ease this classification task (Fig. 1). Especially relevant is the existence of a vascular pattern inside the lung: the bronchopulmonary segment, revealing a preferential relationship between arteries, veins, and airways inside this functional region. The main-stem arteries, veins, and airways lie together in loose connective tissue sheaths at the hilum. Once inside the lung, arteries run along with the ventilating bronchi, developing close to each other and in parallel, whereas veins follow Miller's dictum, generally being as far away from the airways as possible (Miller, 1947). Arteries and veins are closely situated, often intertwined down to the capillary networks, where both vascular trees are connected. CT images display pulmonary vessels and airways with specific characteristics typical of the modality, where the intensity values in arteries and veins (and also airway walls) are very similar when no contrast media is used, yielding unclear boundaries between them. Moreover, high but still limited resolution in CT imaging produces partial volume effects that avoid the visualization of high-generation vessels, capillaries and, therefore, the connection between arteries and veins. This fact, together with the intensity values of airway lumen and the small thickness of airway walls cause airways to become indistinguishable from lung parenchyma earlier than vessels.

The spatial relationship between the bronchus and lung vasculature has been exploited by several researchers by using prior knowledge asserting that airways run in parallel with arteries, making necessary a previous airway tree segmentation. Tozaki et al. (1998, 2001) used intensity filters, information about distances between vessel segments and bronchus, as well as pixel-based connections. Only a gross qualitative analysis of results was mentioned, based on the visual assessment of relevant large vessels from the AV outputs in eight cancerrelated CT cases, performed by a medical specialist. Buelow et al. (2005) proposed a method of front-propagation from seed points at the voxel, segment, and tree levels, based on distance to bronchus (detecting homologous bronchi candidates) and orientation (using the eigenvalues of a Frangi-like tensor structure (Frangi et al., 1998)). Only large vessels (excluding main-stem) were analysed and no quantitative validation was performed, but it showed systematic misclassifications at distal branches of the vascular tree. Yonekura et al. (2007) did not present an AV segmentation algorithm, but studied anatomical features of vessels in crossings and bifurcation points, as well as the relation between bronchial and vessel pulmonary trees from manual AV segmentations, with the objective to discover features for the future development of a classification algorithm.

Other approaches do not use prior knowledge of the bronchial tree and basically trust the connectivity information to lead the separation process. The independent use of connectivity information in a dedicated approach has demonstrated quite limited performance owing to the high proximity of arteries and veins in multiple locations of the development of pulmonary tree irrigations, commonly displaying touching areas and intertwined events. Some authors tried to address these problems by applying some modifications or improvements over the main

connectivity strategy. However, all connectivity-based methods require a starting point: some labelled seeds to begin and drive the connectivity classification. That is the reason why approaches

mainly based on connectivity tend to be semiautomatic methods, needing user intervention to set

the seed points for arteries and veins.

Wala et al. (2011) presented a tracking-based automated method for low-dose CT scans to separate arteries from other surrounding isointensity structures. The algorithm follows arteries starting from automatically detected seed points located in the basal pulmonary areas and tracking the vessel, detecting bifurcations. The validation of results computes an evaluation metric proposed by the authors: sparse surface. A quantitative evaluation of results reported 64% sensitivity and 90% specificity values when using sparse markings on 10 low-dose CT cases (210 arteries and 205 veins in total) as a reference standard. Park et al. (2006) applied a similar approach based on sphere-inflation tracking from terminal points to the heart in six CT angiographies, trying to differentiate between bifurcations and crossings and automatically detecting the pulmonary trunk. However, they only provide gross quantitative evaluation measures: average number of errors in branches (five incorrectly classified and six missing per case), not detailing the validation process and the branches used to that end. More recent semiautomatic works combine fuzzy connectedness using a fuzzy distance transform together with topomorphological features such as multiscale opening and morphological reconstruction to perform the AV segmentation from seed points placed interactively by the user (Saha et al., 2010; Gao et al., 2010; 2012), embedded later into a 2D–3D interconnected graphical user interface. The validation of these approaches included three real human noncontrast CT images, one phantom of CT contrast-separated pulmonary AV casting of a pigs lung, two in vivo contrast-enhanced CT images of a single pig, and mathematically generated tubular phantoms. The results reported that 91–95% accuracy can be achieved with this method iteratively placing around 35 seed points; this can be improved to 99% with more than 60 seed points. They also provide 92–98% of interuser agreement using the semiautomatic tool proposed by the authors, and only 74–87% without it, when a more manual segmentation process is followed.

A different method, trusting only in connectivity, used high-order potentials to label vessel voxels from the automatic detection of the pulmonary artery trunk and left atrium (Kitamura et al., 2013; 2016), comparing the outputs with manual annotations only in most of the important vessels (with intensities higher than −200 Hounsfield Units) measuring, for each vessel, the percentage of centerline voxels correctly classified, achieving an accuracy rating of 90.7%. Another interesting approach was recently published, proposing the construction of a minimum-spanning tree using local information from vessels that can be split in a set of subtrees sharing the same AV classification (Park et al., 2013). Although this method does not address the AV segmentation itself, authors evaluated the potential utility of the tool in a semiautomatic scheme where the user could label each subtree manually, by checking how accurately each subtree is composed by only arteries or only veins in 10 different pulmonary segments. A visual pseudo-score was reported in 10 noncontrast CT images (9.18 out of 10), and volumetric-based accuracy scores were computed in two of the cases by manually extracting the volume of vessel segments mis-classified, obtaining accuracies between 75 and 100% in each pulmonary segment.

The most recent contributions regarding AV segmentation (Payer et al., 2015; 2016; Charbonnier et al., 2016) explored the same idea of dividing the whole vessel tree into multiple subtrees, and classifying them using the interesting fact that arteries and veins tend to be uniformly distributed

within the lung, approaching a common alveolar sac. Payer et al. (2015, 2016) used integer programming to extract subtrees, and also to label them from an arterialness score based on closeness to airways and distribution within the lung. The evaluation using 25 contrast-enhanced and only one noncontrast CT image, where voxelbased labels were manually created (calibre higher than 2 mm and −400 HU), reported 91.1% of voxel-based, and a fairer not calibre-dependent 90.9% of centerline-based labelling agreement (using only manual and automatic common voxels). In Charbonnier et al. (2016), an iterative pruning and detection of AV attachments based on branching angles perform the subtrees extraction, which are later classified based on their distribution and the difference between final arterial and venous volumes. Evaluation was performed in 2365 annotated vessels from 55 noncontrast CT images, and full vascular manual separations (annotated from the extracted subtrees) on a subset of 10 scans, achieving a mean accuracy of 89%.

1.2. Our contribution

The main objective of this work is the design, development, and validation of a new framework to approach the separation of tree-like structures using local information and a specially designed graph-cut methodology that ensures connectivity as well as spatial and directional consistency of the derived subtrees. This framework has been applied to the AV classification of pulmonary vessels images from CT, addressing the clinical need of retrieving detailed vascular information in the study of cardiopulmonary diseases. This automated approach should allow to pull out critical information from the massive amount of data produced by new biomedical imaging technologies, that is unfeasible to extract during the daily clinical routine. The output could become a potential new source of knowledge for prognostic, diagnostic and treatment decisions in cardiopulmonary pathologies.

We present a fully automatic framework for AV classification in pulmonary CT images. The approach starts from an initial pulmonary vessel extraction using scale-space particles, and performs their posterior AV classification based on Graph-Cuts (GC) incorporating: i) AV similarity scores obtained exploiting local information with a random forest (RF) preclassification step and ii) spatial and directional consistency as well as particle-connectivity information. We evaluate the system by assessing the accuracy of vessel particles classification in a set of clinical and synthetically generated noncontrast CT cases.

In Section 2, we describe the methodology proposed to address the AV segmentation/ separation problem in CT images, including a first segmentation of structures and the posterior classification of pulmonary vessels into arteries or veins. Section 3 presents the experimental setup, comprising the description of the datasets, validation strategy, parameter optimization of the method, and the three experiments carried out. Finally, the results and conclusions are detailed in Sections 4 and 5, respectively.

2. Methods

In this work, we present a method combining both local prior knowledge and connectivity information along with spatial, and directional consistency to perform the classification of detected particles located in tree-like structures. Concretely we approach the classification of the pulmonary

vascular arterial and venous subtrees. This framework can be formulated as an energy minimization problem with the form of a binary first-order Markov random field (MRF):

$$E(p) = \sum_i U_i(p_i) + \alpha \sum_{i,j} \mathbb{P}_{i,j}(p_i, p_j) \quad (1)$$

where p_i are vessel particles (a centreline discretization of arterial and venous trees) that need to be classified as arteries or veins ($p_i \in \{0, 1\}$); U_i in the unary term representing the compatibility of the data with a specific label; and $\mathbb{P}_{i,j}$ is the pairwise term representing the compatibility of neighbouring labels.

This formulation can also be expressed in terms of GC (Boykov et al., 2001): MRF variables can be described as a set of vertices V in an undirected graph $G = \langle V, \mathcal{E} \rangle$, where the weights of the set of edges G are derived from a pairwise term, also named as boundary term. Two special nodes (sink t, and source s) are connected to V creating edges

($\mathcal{E}_t = \{(s, v_i) \cup (v_i, t) | v_i \in V\}$) with weights derived from a unary term, named as regional term. There is a one-to-one correspondence between binary classifications, configurations of the MRF, and cuts (partitions of the graph into two connected components, S and T, containing s and t nodes, respectively). Solving the MRF problem is equivalent to finding the min-cut, C_{min}, with the minimum cost $E(S, T)$ (sum of all weights from edges that are removed to obtain a partition). Thus, the solution of Eq. (1) can be reformulated to minimize the combination of two energies correspondent to boundary and the regional terms of the graph G, that can be solved exactly using linear programming (Boykov and Veksler, 2006):

$$C_{min} = \underset{S,T}{\operatorname{argmin}}(E(S,T)) = \underset{S,T}{\operatorname{argmin}}(E_{bound}(S,T) + \alpha \cdot E_{reg}(S,T)) \quad (2)$$

where the α parameter balances the weight of the regional and boundary terms:

$$E_{bound}(S,T) = \sum_{p1 \in S,\, p2 \in T} W_{bound}(p_1, p_2)$$

$$(p1, p2) \in \mathcal{E} \quad (3)$$

$$E_{reg}(S,T) = \sum \left(W_{reg}(p_1, t) + \sum W_{reg}(p_1 \in S \quad p_2 \quad (p_2, s)) \right) \in T$$

The use of GC in the presented AV problem is a quite direct and reasonable strategy: a) arterial and venous irrigations follow a natural tree structure, which is a specific type of graph; b) the particle-based representation of vessel segmentations provide a direct correspondence between particles and nodes; c) the spatial relationship between particles can be represented in the boundary term –n-links, neighborhood links– ; d) prior information about particle classes can be incorporated in the regional term –t-links, terminal links– (s and t representing the ideal artery and the ideal vein); e) the minimal-cut provides a direct binary classification of particles into arteries and veins. In our

system, the boundary term involves connectivity and structural information of particles, and regional term comes from an RF-based pre-classification step based on local features, providing an AV similarity score.

In the following subsections, we define the discretization of vessels into a set of particles to be classified, and the construction of the graph G, edges and their corresponding weights representing the boundary and the regional terms, to form a suitable energy function for pulmonary AV classification (Fig. 2).

2.1. Segmentation of Pulmonary structures

Prior to the differentiation between arteries and veins, a segmentation stage is necessary to separate structures of interest (vessels) from the background, simplifying and constraining the task. Moreover, lung and airway segmentations are also important in order to include constraints and extract relevant information that will be used into the system.

Lung segmentations obtained with Ross et al. (2009) comprise lung parenchyma including small vessels and airways, but excluding mediastinum and principal vessels and airways passing through it.

The method used for segmenting airways is based on a constrained energy minimization problem using scale-space particles to densely sample intraparenchymal airway locations (Estépar et al., 2012a). A posterior clustering step using connected components is applied, filtering results in a set of candidate airway points situated in both physical space and scale. The output is a discretization of the airways in centerline points with a radius associated with the size of the airway structure in these locations.

A similar strategy of scale-space particle filtering based on a Hessian matrix response was used to extract the pulmonary vasculature (Estépar et al., 2012b).

The discretization of vessels (and airways) into a set of particles allows not only the simplification of structures, but a better representation for classification purposes, conferring the same importance to every point in the vessel tree independently of the calibre size, but dense enough to properly evaluate the algorithm in the whole structure.

2.2. Pulmonary AV classification with GC

Once the pulmonary vessels are segmented and represented with a set of scale-space particle points, we can proceed to their classification into arteries and veins using the proposed GC approach. Briefly the steps in the proposed framework are the following:

- Given the airway and vessel tree particle segmentations, the area enclosing the bronchial tree is computed as the convex hull of the airways subtree particles.
- Within this convex hull the arteries run in parallel and close to the airways. A pre-classification step based on these local properties assigns to the vessel particles within this mask more probability to belong to either arteries or veins. The particles outside the convex hull will have equal prior probability to belong to arteries or veins subtrees

(0.5 for to each of them). This process allows the computation of the regional term of the GC (Section 2.2.1).

- Once the particles within the convex-hull are pre-classified, the graph-cut is applied to all the vessel particles of the lung ensuring spatial and directional consistency as well as connectivity forming tree-like structures (Boundary Term Section 2.2.2). Therefore, the GC expands the pre-classification from the area where airways are detected to all the vessel particles in the lung or lobe (Section 2.2.5).
- As the last step, particles that remain isolated are connected in a refinement step to ensure that the vessel trees are complete and no particles are disconnected (Section 2.2.4).

In the following subsections, we detail the construction of the boundary and regional terms (Eq.

(3)), establishing criteria to create edges between nodes of the graph G and their weights.

In Section 2.2.1, we define the regional term from AV similarity scores obtained with an RF pre-classification. The individual classification of vessel particles may result in spatial labelling inconsistencies, connectivity/structural information is then incorporated in the boundary term (Section 2.2.2) to define the final GC architecture for AV segmentation.

2.2.1. Regional term: RF pre-classification based on local features—As

explained before, the regional term of the GC, represented by the edges $\mathcal{E}_t$, must include prior likelihood information about the classification model. This role can be played by AV similarity scores derived from imaging features that take advantage of the prior knowledge (Section 1.1). An RF pre-classification using these local features of particles is used to obtain the scores and, therefore, the weights of t-links in the regional term of the graph (w_{reg}).

Feature extraction.: In our approach, the main source of information for AV separation is the relation between vessel irrigations and the bronchial system (Fig. 3). Moreover, some structural and image-based vessel features provide additional important knowledge and the combination of all of them would allow the study of vessel-airway relation differently depending on the variation of the other properties. Because the classification is performed over the individual vessel particles obtained in the segmentation stage, the selected measurements are computed for each one of them.

- Scale. The local radius of the vessel provides information about the branch generation in the tree structure.
- Hessian matrix eigenvalues. The local 3D geometry of the structure eases the distinction between bifurcations, segmental points, and vessel crossings.
- Local intensity histogram. The local distribution of intensities improves the differentiation between arteries and veins, particularly when contrast is used (in Computed Tomography Pulmonary Angiography-CTPA databases). The normalized intensity histogram of N_{bins} bins is computed in the 3D spherical area defined by the local radius and centred on the particle location. The normalized histogram is represented by 10 bins leading to 10 features to describe the local distribution of the intensity around the particle.
- Distance to closest airway points. Mean and standard deviation of Euclidean distance between the vessel particle and its N_{neigh} closest airway particles provide the most critical information for the classification task.
- The Distance between closest airway points. Mean and standard deviation of cross-distances between the selected N_{neigh} airway points associated with each vessel particle can also be valuable, providing information about the existence of single/multiple proximal bronchial branches.

- Orientation similarity between vessels and closest airway points. The study of the Hessian matrix also allows the computation of a local preferential direction in each particle location (at a specific scale) that is useful to assess the parallelism between arteries/veins and their closest airways. By using the preferential local direction of vessel, v_{ves}, and each neighbour airway particle, v_{air}, we can compute an orientation similarity score (OSS) based on the cosine-like distance, with values between 0 (perpendicular vectors) and 1 (parallel vectors), that can be joined in an orientation similarity histogram of N_{bins_O} bins.

$$OSS(v_{ves}, v_{air}) = \frac{|v_{ves} \cdot v_{air}|}{\|v_{ves}\| \cdot \|v_{air}\|} \quad (4)$$

Predictive model.: An RF approach (Breiman, 2001) was selected to perform the preclassification of particles owing to their advantages compared with other machine learning methods: robustness, excellent generalization capability, high speed, low memory use, facility to understand and interpret the results and the management of probabilistic estimates and, therefore, probabilistic similarity scores for the different classes. Because the principal disadvantage is the high dependence of the results depending on the parameter selection, we decided to constrain the number of features by using only those with high probability to be informative for the AV classification (the ones presented above, coming from prior biological and image-based knowledge).

T independent estimators based on decision trees using $n_{feat} << N_{feat}$ random features (being N_{feat} the total number of features) were used to constitute the RF predictive model.

AV similarity scores.: Once the model is created from the features extracted in vessel particles from the training set, using the label artery as positive class, it can be used to classify particles belonging to new cases.

The final classification scores obtained for each new particle are probabilistic estimations, and therefore, they can be interpreted as AV similarity measures, where values represent the probability of a particle belonging to the arterial (P_{art}) or venous (P_{vein}) irrigations, with

$P_{art} + P_{vein} = 1$.

The probabilistic nature of the AV similarity measures allows a direct conversion into a binary classification by simple thresholding. A visual examination clearly reveals the existence of spatial inconsistencies (Supplementary Figure S1), which support the addition of connectivity information and the use of the proposed GC strategy, instead of a simple machine learning classification using local features.

t-links.: The weights of the special edges (t-links), connecting vessel particles and the virtual nodes, are directly fixed to the probabilistic estimations:

$$W_{reg}(p_i,s) = P_{art}(p_i) \quad (5) \quad W_{reg}(p_i,t) = P_{vein}(p_i) = 1 - P_{art}(p_i)$$

where p_i are vessel particles, and s and t are the virtual nodes representing the ideal artery and vein.

However, because local features used in the machine learning pre-classification step essentially exploit the topological and structural relation of vessels with respect to airways, the probabilistic similarity scores obtained in peripheral areas of the lung are less reliable. Segmentations of the bronchial tree cannot reach these areas, and thus the absence of airway information implies huge limitations on the classification power of small vessel particles and those located far from airways. Similarly, particles located in mainstem vessels offer low confidence levels of AV similarity scores owing to their abnormal spatial location at the hilum, where arteries, veins, and airways lie together.

For these reasons, as previously introduced, we decided to limit the use of pre-classification scores to areas where reliable bronchial information can be found: the convex hull (plus a safety margin, m_{CHull}) of the available bronchial tree segmentation (Fig. 4) and vessel particles not belonging to very small vessels (scale lower than r_{ves}^{min}). The exclusion of hilum region can be performed by imposing a limitation on the vessel calibre r_{ves}^{min}. Only vessel particles with sizes between these values and located inside the convex hull are considered to construct their t – links from the probability scores. Note that the values of r_{ves}^{min} and r_{ves}^{min} need to be carefully selected (or omitted) owing to the high inter-subject variability and the probable dependence of vessel sizes on pathological states.

Therefore, the final weights of t-links connecting vessel particles with the ideal artery and vein, are fixed to the probabilistic estimations only if the particles comply with the previous restrictions; if not, we assume that there is not a preferential classification as artery or vein for the particle, and the connectivity information will drive its final classification:

$$W_{reg}(p_i,s) = \begin{cases} P_{art}(p_i), & \text{if } p_i \in CHull(V_{air}) \;\&\; r_{ves}^{min} \le \sigma_i < r_{ves}^{max} \\ 0.5, & \text{otherwise} \end{cases} \quad (6)$$
$$W_{reg}(p_i,t) = 1 - W_{reg}(p_i,s)$$

where σ_i denotes the scale of vessel particle p_i, and $CHull(V_{air})$ represents the dilated convex hull of all airway particle points $V_{air} = \{p_i^{air} \forall i\}$.

2.2.2. Boundary term: connectivity and structural information—As described before, the boundary term of the GC, represented by the edges $\mathcal{E}$, must include information about the coherence between neighbouring nodes. In our system, this corresponds to the connectivity information between vessel particles. The rules to establish connections between nodes and the energy function to set up their weights need to be defined in order to conform the final n-links of the graph. Both steps are important to minimize possible missconnections and to enhance the real

topology of the vessel trees, taking into account that events such as bifurcations or crossings are not explicitly defined in our system based on particles.

Connectivity rules.: Unawareness about real structural information of particles, together with the previously reported problems related with the topology of arterial and venous trees (proximity of irrigations, intertwined structures difficult to differentiate from bifurcation points, etc.), hamper a proper definition of the ideal connectivity in our system (i.e., twoneighbour links except for terminal and bifurcation points, with one and three connections, respectively). For that reason, we follow a conservative structural connectivity strategy by initially allowing the generation of links between each particle and all the particles that are within a certain distance r_{neigh} which is big enough to assure abundance of links. However, excessive n-links could lead to intra-irrigation (artery-artery / vein-vein) and inter-irrigation (artery-vein) particle connections when irrigations are close. Two strategies are used to overcome this difficulty:

- Anisotropic connections: In order to ensure that the neighbouring particles are located along the local preferential direction $\vec{v}$ of the specific vessel particle p, we constrain the search inside a cylinder along this preferential direction (Fig. 5a).
- Stabilizing the number of links: fix a maximum number of connections between particles. We create N_{con} maximum n-links between a vessel particle and its N_{con} nearest neighbours inside the anisotropic area. In that way, each node in the graph will have approximately N_{con} edges, avoiding over-connections of particles in vessel-intertwined areas. The cut of an edge will depend more on the strength (weight) of the edges than on the number of them.

In summary, the final set of n-links is defined as:

$$\mathcal{E} = \{(p_1, p_2 | p_1, p_2 \in V_{n-t} \& p_2 \in kNN(p_1, \vec{v}_1))\} \quad (7)$$

where $k = N_{con}$ and:

$$kNN(p, \vec{v}) = \{p_i' | p_i', p_j' \in B_{rneigh}(p) \cap \mathbb{C}_{rneig_2}(p, \vec{v}) \} \quad (8)$$
$$\& dist(p, p_i') < dist(p, p_j'), \forall i = 1.k, \forall j > i$$

being B_r and $\mathbb{C}_r(p, \vec{v})$ the spherical area of radius r around location p, and the cylindrical area of radius r centred at p and with orientation $\vec{v}$, respectively.

n-links.: The strength of the connections between particles, depicted in the graph by the weight of the n-links, will represent a probability-like estimation of being neighbours in the vascular tree for couples of vessel particles. These weights are defined using three critical features:

- Scale consistency: difference between local scales of particles. The more similar the scale of close particles, the higher the probability of being neighbours.

$$w_\sigma(p_1, p_2) = |\sigma_1 - \sigma_2| \quad (9)$$

- Particle proximity: Euclidean distance between them. The closer two particles are located, the higher the probability of being neighbours in the arterial or venous trees.

$$w_{dist}(p_1, p_2) = \| p_1 - p_2 \| \quad (10)$$

- Direction consistency: the angle between preferential directions. The parallelism between vectors can be represented numerically by a cosine-based distance varying from 0 (perpendicular vectors) to 1 (parallel vectors):

$$w_{\vec{u}}(\vec{u}_1,\vec{u}_2) = \cos\left(\min\left(\varphi(\vec{u}_1,\vec{u}_2), \pi - \varphi(\vec{u}_1,\vec{u}_2)\right)\right) = \left|\cos(\varphi(\vec{u}_1,\vec{u}_2))\right| = \frac{|\vec{u}_1 \cdot \vec{u}_2|}{\|\vec{u}_1\| \|\vec{u}_2\|} \quad (11)$$

where $\varphi(\vec{u}_1, \vec{u}_2)$ represents the angle between both vectors in the range $[0, \pi]$.

The direction of tubular structures changes smoothly and continuously along the vessels, so neighbouring particles belonging to the same tree branch should share a similar preferential direction. However, we also need to take care of some problems regarding bifurcation points, where the definition of preferential direction is unstable. In these areas, the criterion based only in parallelism between preferential directions could lead to small weights in bifurcation edges. The use of a connectivity vector between particles (representing the direction of the potential vessel passing through both particle locations) overcomes the problem (Fig. 5b):

$$\vec{v}_{con}(p_1, p_2) = p_1 - p_2 \quad (12)$$

Therefore, a proper and robust way to define the direction consistency is using the parallelism between the connectivity vector and the individual preferential local directions at the corresponding particles:

$$w_{\parallel}(p_1, p_2) = g\left(w_{\parallel}(\vec{v}_1,\vec{v}_2), w_{\parallel}(\vec{v}_1,\vec{v}_{con}(p_1, p_2)), w_{\parallel}(\vec{v}_2,\vec{v}_{con}(p_1, p_2))\right) \quad (13)$$

where $\vec{v}_i$ represents the local preferential direction of particle P_i, and g any function combining all or some of these variables.

By using the three sources of information, we can construct different weighting functions w_{bound} to define the strength of the n-links:

$$w_{bound}(p_1, p_2) = f(w_{\sigma}(p_1, p_2), w_{dist}(p_1, p_2), w_{\vec{p}}(p_1, p_2)) \quad (14)$$

where f is any function combining all or some of these variables. Section 3.3 describes how 9 different functions were tested and which function was finally used.

2.2.3. Refining the GC—The restrictions introduced in the construction of n-links may impede the full connection of all vessel particles into a graph with a single connected component, resulting in a forest $\mathcal{F}$, with some isolated subtrees. These isolated subtrees of particles occur mainly in peripheral areas where vessel segmentation may present discontinuities specifically due to the small scale of the vessels. If a subtree is located outside the convex hull of airways (where no suggestion of AV labels is included using t-

links) several cuts of the graph with minimum energy exist, being equally probable the classification of every particle in these isolated subgraphs as either arterial or venous ones (Fig. 6a).

The assurance of only one connected graph G would avoid these uncertainty problems related with isolated subtrees, providing a unique classification coming from a global minimum for all particles in the vascular tree. Converting the forest $\mathcal{F}$ into a one-connectedcomponent graph $\mathcal{F}$ also allows for the application of the proposed AV method with vessel segmentations presenting large discontinuities between segments in the tree, a common problem with many methods in bifurcations and high-generation vessels.

The employed strategy connects subtrees iteratively depending on the minimum distances between them, repeating the following steps until the graph is composed of only one connected component (Fig. 6b): a) select the biggest connected component as the principal one; b) compute the Euclidean distances from particles of secondary subtrees to particles belonging to the principal one; c) select the minimum distance and create the edge connecting both particles with the cost determined by the specific weighting function w_{bound} (p1 , p2) (Eq. (14)).

2.2.4. GC classification: min-cut/max-flow algorithm—Once the final graph is completed, the minimum-cut, C_{min}, is computed using the min-cut/max-flow algorithm

(Boykov et al., 2001) versioned by Delong et al. (2012) to solve Eq. (2), providing a partition of the graph G into two connected components G_1 and G_2 and, therefore, the final binary AV classification.

3. Experimental methods

In the following subsections, we detail the datasets used in our experiments (Section 3.1), present the strategy used for validation (Section 3.2), and describe the different experiments performed to properly adjust the parameters (Section 3.3) and to evaluate the AV segmentation results (Section 3.4).

3.1. Data description

The extraordinary complexity and density of pulmonary vessels and airways render the performance of manual segmentations difficult, limiting the access to accurate AV reference standards, which are critical for training and validation. The use of synthetic CT images of the lung could overcome these difficulties by providing a database of pseudorealistic cases in a constrained and controlled scenario where each part of the lung is differentiated unequivocally. Jimenez-Carretero et al. (2016) presented a method to generate computational pulmonary CT phantoms automatically, by synthesizing lung structures (arteries, veins, airways) and parenchyma in a realistic way. In that work, two versions of each phantom were synthesized, with different

curvature conditions for vessels and airways obtained by the exclusion/inclusion of a deformation stage. The database of the final pulmonary CT phantoms is available online via Zenodo (Jimenez-Carretero et al., 2015a; 2015b).

In our experiments, a set of clinical and synthetic cases were used to evaluate the performance of the proposed AV segmentation method in a relatively large number of cases (Fig. 7). Real cases correspond to noncontrast CT images, and the synthetic ones include: a) simpler bronchopulmonary segment phantoms and b) realistic and complete anthropomorphic pulmonary CT phantoms. For each individual lung, a voxel-based AV reference standard was already available (phantoms) or manually created (clinical cases).

3.1.1. Bronchopulmonary phantoms—A bronchopulmonary segment represents a repetitive anatomical and functional structure that exhibits a vascular pattern inside the lung, where arteries run along with the ventilating bronchi developing close and in parallel whereas veins lie in the interval between segments, as far away from the airways as possible.

For that reason, simulated CT of bronchopulmonary segment phantoms presented in Jimenez-Carretero et al. (2016) were suitable for initial tests in the development of the algorithm, including the possibility of studying the dependency of results regarding different curvature conditions for vessels and different densities of bronchial information.

Six groups of phantoms were used (64 cases/group), combining three different densities of synthetic airways ($N_{art} = N_{vein} = N_{air}/FA$ with $FA \in \{1, 2, 3\}$, where N_i represents the number of terminal points for arterial, venous, and bronchial trees, respectively) and two different curvature conditions for vessels and airways (nonde-formed/deformed). $N_{art} = N_{vein}$ varies from 30 to 100.

3.1.2. Anthropomorphic pulmonary CT phantoms—The dataset of synthetic CT images representing whole lungs presented in Jimenez-Carretero et al. (2016) were used to mitigate the limitation in the availability of clinical CT cases with proper AV reference standards, allowing to test the effect of different vessel curvatures.

A total of 48 synthetic lungs were used: 24 pulmonary phantoms and their corresponding versions with a realistic curvature of vessels and airways (nondeformed/deformed).

3.1.3. Noncontrast CT cases: dataset 1 fully AV-labelled—Three CT cases used in this work were acquired as part of the COPDGene study (Regan et al., 2011). COPDGene centres obtained approval from their Institutional Review Boards and all subjects provided written informed consent. The data were anonymized by the COPDGene study data coordinating centre before sharing them.

Images correspond to noncontrast CT from patients suffering from COPD. A total of six different lungs were available, with pixel spacings between 0.5859 and 0.7031 mm, and slice thicknesses between 0.5 and 0.625 mm.

A slow but necessary voxel-based manual labelling process was used to obtain the AV reference standards for these six lungs. This extremely time-consuming task was performed using ITK-Snap[2] in two stages. In the first step, a pre-classification was carried out by two trained engineers (≈15 effective hours per lung) using a slice-by-slice manual labelling process over the CT images with the original sub-millimetre resolution and following a deep-first strategy. In the second step, a radiologist with broad expertise in lung imaging checked and corrected classification errors using the same strategy. During the entire process, the labelling was constrained to the voxel-based vessel segmentation obtained after sphere-inflation from the particle-based segmentation described in Section 2.1. As a consequence, the accuracy of the general vessel segmentation could not be evaluated, but we assure the abundance of vessels from multiple generations (up to 4th or 5th branching levels), implying the enrichment in the evaluation of the AV classification method.

3.1.4. Noncontrast CT cases: dataset 2 partially AV-labelled—An additional independent dataset was created for repeatability and generalization purposes, comprised of twenty new cases corresponding to right lungs of patients coming from COPDGene study (Regan et al., 2011), where vessels and airways were segmented in both the upper and lower-right lobes. A denser segmentation of vessels was performed up to 5th or 6th branching levels. The reference standard was created using the strategy explained in Section

3.1.2, but performing segmentations and labelling particles using Chest Imaging Platform (CIP[3], Brigham and Women's Hospital, Boston, USA).

3.2. Validation strategy

In this work, the AV segmentation has been addressed like a binary particle-based classification problem. Thus, the Fl-scores allow the study of the exactness and completeness of the specific algorithm in a very robust way.

Although the desired readout is the final AV classification, the intermediate classifications can be also examined to evaluate different stages of the algorithm. In our case, three evaluations have been performed corresponding to: i) pre-classification using RF_{CHull} based on local features (only evaluated for the particles inside the convex hull of the airways); ii) classification with GC before applying the refinement stage; and iii) final AV classification with refined GC. Additionally, to understand the benefit of the graph-cut approach with respect to an RF classification in the full lung, (RF) is reported as reference for the experiments dealing with clinical cases. The RF_{CHull} provides the accuracy of the preclassification step that will be used as regional term in the GC stage. Differences between the refined and nonrefined versions of GC illustrate the effect of graph disconnections in the results at the full lung levels.

3.3. Parameter optimization

The whole AV workflow depends on several parameters that influence the results as well as the performance or the efficiency of the algorithm. Therefore, these parameters need to be optimized

[2] www.itksnap.org/.

[3] www.chestimagingplatform.org.

to obtain acceptable classification results. We used one half of the set of the bronchopulmonary phantoms, 32 cases/group (Section 3.1.1), to perform this task.

Local features parameters.— $N_{neig} = 20$, $N_{binsI} = 10$ and $N_{binsO} = 5$ were fixed experimentally by running several tests. The selected final values showed a good balance between efficiency and classification power.

RF parameters.— A grid search strategy was followed to test and select the best combination of parameters: $nfeat \in \{N_{feat}, \log_2(N_{feat}) + 1, \sqrt{N_{feat}}\}$ and $T \in \{2^n \mid n = 0, \ldots, 10\}$. A leave-one-out cross-validation strategy in the dataset of bronchopulmonary phantoms reported that the best model was created with $nfeat = \sqrt{N_{feat}}$ and an appropriate balance between efficiency and performance can be reached with $T = 500$ (Supplementary Fig. S2).

GC parameters.— Some parameters used in the construction of the GC were selected empirically after a relatively small number of tests, taking into account mean distances between particles and prior knowledge from human lungs: $r_{neigh} = 3$ mm, $N_{con} = 5$, $m_{CHull} = 20$ mm, $r_{vesmax} = 4$ mm, $r_{vesmin} = 1.5$ mm. A grid search for the combination of 9 different cost functions (Supplementary Subsection: Weighting functions w_{bound}) $w = \{w_i(p_1, p_2) \mid i = 1, \ldots, 9\}$ and $\alpha = 6, \ldots, 12$ revealed that the tendency of results remains very similar for every group of phantoms (different combinations of airways density and curvature condition) reaching the best results when $\alpha = 8$ and $w_{bound} = w_8(p_1, p_2)$ (Supplementary Fig. S3):

$$w_8(p_1, p_2) = w_2(\mathbf{v}_1, \mathbf{v}_2) \max\left(\left\|\mathbf{v}_{conn}(p_2 + 1, p_2)\right\|, \left\|\mathbf{v}_{conn}(p_1, p_2)\right\|\right)^2 \quad (15)$$

3.4. Experiments

3.4.1. Experiment 1: bronchopulmonary phantoms— Bronchopulmonary phantoms used during parameter optimization served as a training set for the construction of the RF classifier, treating each of the six groups independently. The rest of the bronchopulmonary phantom cases were tested to evaluate and analyse the effect of the curvature of pulmonary structures as well as the density of airways in the final AV segmentation results.

Taking advantage of the knowledge about the actual label of each part of the synthetic image, we constrained the particle-based vessel segmentation method into the regions defined by each specific structure (arteries and veins), improving the accuracy of segmentations and, consequently, reducing the misclassification errors derived from segmentation problems, resulting in a more specific and accurate evaluation of the AV classification problem.

3.4.2. Experiment 2: anthropomorphic pulmonary CT phantoms— All the 48 cases were used to properly validate the AV segmentation approach under a realistic scenario. Deformed and non-deformed phantoms were treated separately using two strategies: leave-one-out cross-validation (CV) and two-fold (keeping balance in the number of right and left lungs, randomly selected for each group) to confirm previous results using a lower amount of data for

training. Again, the arterial and venous segmentations were conducted separately using the knowledge about actual labels for each voxel inside the synthetic lungs.

3.4.3. Experiment 3: noncontrast CT cases—The six lungs corresponding to clinical noncontrast CT images were evaluated using a leave-one-out CV and three-fold CV ('leave-one-subject-out to exclude from training both the left and right lungs coming from the same patients). In this case, all vessels were segmented simultaneously to create a more realistic scenario for the processing and evaluation of new clinical cases, where segmentation and classification errors can occur.

3.4.4. Experiment 4: importance of airway segmentation—A way to check the consequences of missing bronchial information in clinical cases is to evaluate the AV separation method using pruned versions of the original airway tree segmentation. To that end, we iteratively removed airway particles with scale lower than a specific value from the segmented bronchial trees of these six lungs and reran the leave-one-out CV validation.

3.4.5. Experiment 5: replication set—A final test to check the repeatability of results and generalization of the presented approach was performed on a completely independent set of clinical cases. Furthermore, this experiment allows to explore and validate the possibility of performing AV classification in incomplete lung cases. A leave-one-out cross validation experiment was performed.

For these 20 cases, airway particles have been initialized using a binary mask created from the method recently proposed in Nardelli et al. (2017). In particular, the airway feature strength is first extracted using Nardelli et al. (2017). Then, a simple thresholding of 0.7 is used to obtain a binary mask that is used as initialization of particles. This greatly improves the number of airway particles normally extracted when using the filter proposed by Frangi et al. (1998) to extract airway strength.

4. Results

The automatic algorithm was developed combining Python and Matlab and executed on an AMD Athlon II X4 630 @ 2.8GHz with 12GB RAM under Ubuntu 12.04. The complete workflow excluding the initial segmentation of pulmonary structures was approximately 4 min for each complete lung; where half of the time was used for the GC refinement stage.

Table 1 presents the classification results of the six groups of bronchopulmonary phantoms (Experiment 1). The performance behaves as expected when varying conditions: the accuracy of the method significantly decreases together with the density of the generated airways (higher FA) owing to the reduction in the available information of the vessel–airway relation that mainly drives the RF classification step. Similarly, a higher deformation and curvature in vessels and airways increases the uncertainty in the computation of preferential directions and, therefore, a reduction in the orientation similarity and direction consistency of neighbouring particles, explaining the quite small but appreciable deterioration of the algorithm performance. Similar results (Table 2) were obtained when using the anthropomorphic pulmonary CT phantoms (Experiment 2), where the structure and density of vessels and airways are more realistic. In both experiments, we can

clearly appreciate that subsequent stages of the algorithm incrementally improve the AV classification results.

Although the pre-classification based on RF reports fairly good results (F1 > 74% and > 72% in experiments 1 and 2, respectively), the incorporation of connectivity information into the GC strategy highly increases the accuracy of the classification (F1 > 91% and > 90%, respectively). GC not only overcomes the spatial inconsistencies coming from the RF stage, but also extends the classification to particles located outside the convex hull of the airways in a suitable way. The final refined version of the GC reports slightly better results (F1 > 92%), demonstrating the improvement obtained when ensuring the proper connectedness of the graph, but also showing the relatively small prevalence of isolated subtrees in the phantoms. Some examples are shown in Fig. 8.

Experiment 3 (Table 3) reports better results for clinical cases in the pre-classification step using RF_{CHull} (around 78–80%). This is because of the not fully realistic nature of air and vessel flow systems in synthetic cases which render difficult the proper management of local features and vessel–airway relationship, which is more stable in real lungs. Contrarily, although GC highly improves the results (around 87–89% with GC and 89–91% with the refined version), the accuracy of the classification is slightly lower than in the phantoms, suggesting a higher complexity of air and blood flow systems especially in peripheral areas in real lungs compared with synthetic cases. The results using a three-fold cross validation

(leaving one subject out) resulted in slightly worse performance than using CV. Figs. 9 and 10 show the performance of the AV segmentation algorithm in two interesting lungs corresponding to the best and worst resulting classifications.

Fig. 11 presents the results of experiment 4 checking the importance of using an accurate segmentation of airways to reach a good AV segmentation by progressively pruning the airway tree. Reasonable results are obtained when airways down to 1 mm of radius are properly segmented.

Table 4 presents results from experiment 5. The evaluation on a completely independent dataset reports lower performance, especially noticeable in the RF pre-classification stage, providing the GC stages with a more unsuitable starting point to get an optimized final AV classification. This fact was further analyzed by supervising the quality of airways segmentation by two experts blinded to the validation results. Three different groups of cases were identified: cases with complete airway tree, cases with partial airway tree (part of a main branch was missing) and incomplete (one or more main branches missing). Table 4 shows the results within these categories showing that the effect of the quality of the airway segmentation is important. This experiment confirms the possibility of using the presented tool to process incomplete lungs with a denser representation of the vessel trees (up to 5th or 6th branching levels) with competitive results especially when a good segmentation of the airway tree can be guaranteed.

5. Discussion and conclusions

This work presents an innovative framework to separated tree structures from image data proposing a specifically designed graph-cut methodology that exploits local features imposing the

connectivity and the spatial and directional consistency required to ensure complete trees. The proposed framework enables a completely automatic AV segmentation/ separation of the pulmonary vasculature represented as discrete particles that are preclassified based on the vessel-airway relationship characteristic of the central areas of the lung using an RF approach. However, the spatial inconsistency of results when only a preclassifier is used demands the inclusion of connectivity information and spatial consistency into the system. As a consequence, a GC approach was followed to merge both sources of information. Previous works used a similar graph-based strategy for retinal AV segmentation in 2D fundus photography (Joshi et al., 2014; Eppenhof et al., 2015) achieving very nice results. However, the classification is essentially based on color-intensity properties, and connectivity information is introduced mostly to distinguish between vessel crossings and bifurcations that are typical for this 2D image modality. Our framework exploits these approaches in 3D overcoming different challenges such as the complexity of 3D tree structures, and scarce information coming from image intensities to distinguish in our case arteries and veins. For that reason, energy functions need to include more complex local features and deal with more general and smooth structural consistency. The same strategy and framework could be applied to other problems, for example, the separation of liver vasculature defining the specific local relations among the different vessel trees.

The discretization of vessels with scale-space particles allows the simplification of structures and a better representation for classification purposes as well as a representation of the vessel calibre. This representation is flexible, and any external segmentation could be transformed into a particle-like segmentation. We suggest an adjustment of parameters if the structure, distances, and disposition of the discrete vessel/airway particles change, in order

to achieve an optimal performance of the method. In this work, we did not carry out a comprehensive comparative study against other recent methods proposed in the literature owing to the lack of open access to the corresponding algorithms. A metaevaluation of the results presented in this work and other suggests that our approach is competitive in terms of a combination of usability, validation, and performance, by providing a: i) completely automatic tool; ii) strong and rigorous evaluation using precise AV ground truths covering the whole lung and reaching high generations of the vasculature; and iii) validation of the method in clinical and synthetic scenarios, using a large number of data including 26 lungs from clinical noncontrast CT images and 48 synthetic lungs from anthropomorphic pulmonary CT phantoms. Furthermore, the proposed AV segmentation method has been fully incorporated in the CIP (Brigham and Women's Hospital, Boston, USA), a Python/C+ + platform, allowing and easing future studies and comparative evaluations.

Visual examinations (Figs. 8–10) revealed that classification mistakes appear in the peripheral and distal areas of the lung (isolated subtrees may exist), and in areas close to the hilum (where air and blood flow systems are close). Most bifurcations are properly connected thanks to the connectivity rules included in the boundary weighting function that takes into account both the parallelism with the vessel preferential direction and the new potential vessel connection (connectivity vector). Given the imposed connectivity and spatial consistency, if errors exist they appear in the same

subtrees keeping a classification consistency through subsequent generations, which could be easily corrected with an interactive tool requiring relatively low user burden.

F1-scores have been reported in three different stages of the algorithm: a first preclassification of particles located in the convex hull of airways, using RF based on local features (RF_{CHull}); a complete vessel particle classification using the RF model used as reference (RF), a complete vessel particle classification using GC combining the RF preclassification step with connectivity information(GC); and a final AV classification using a refined version of GC that addressed the disconnection of small-scale distal vessels (Ref. GC). These results show a relevant improvement in the accuracy (~20%) of the vessel particle classification with the proposed framework (GC and Ref. GC) with respect to using only the pre-classification based on local information applied to the whole area of the lung under study (RF).

Bronchopulmonary segment phantoms were used to set up the system and select the best parameters for the algorithm. Results of these data reveal that the proposed approach can obtain almost perfect AV separations in pseudo-realistic and simplified lung scenarios (Table 1).

An actual validation of the system was performed with anthropomorphic pulmonary CT phantoms and clinical noncontrast CT images. Results of both datasets suggest very good performance of the algorithm in real and complex synthetic scenarios, reporting F1-scores up to 90% (Tables 2 and 3). The disconnection of small vessels subtrees, addressed by the refined version of the GC, achieved important improvements at peripheral areas.

As expected, experiments with phantoms (Tables 1 and 2) probe that curvature of structures introduce changes in the definition of connectivity between particles that influences the accuracy of the algorithm. At the same time, it shows that the robustness of vessel segmentation and the detection of principal directions, are vital for the proper definition of connectivity and therefore, to maximize the classification power of the AV segmentation method. Similarly, tests with clinical cases have demonstrated the importance of an accurate and sensitive method for airway segmentation to maximize the computation of vessel-airway relationships that drive the final AV classification (Fig. 11). When comparing results obtained with synthetic and real lungs (Tables 2 and 3), the AV segmentation method performed comparatively, reinforcing the similarity between anthropomorphic pulmonary CT phantoms and clinical non-contrast CT images. However, results also suggest the existence of some differences between the performance on clinical and synthetic cases that are related to higher complexity in the real lungs, the lack of information from airway in some principal bronchial locations, and the noise introduced by the vessel segmentation step.

The replication experiment using a completely independent set of lungs partially segmented (two lobes) reinforces the evaluation of the method and demonstrates the possibility of performing reasonable AV segmentations in cases with incomplete information, as well as a denser representation of the vessel trees (up to 5th or 6th branching levels). The cross validation experiment confirms the advantage of a larger training database to properly capture the huge variability present in clinical cases, in order to improve the performance of the method, especially of the RF pre-classification stage.

This independent dataset also allowed testing the methodology in a realistic setting considering images of different quality and derived airway segmentations. The dependence on the airway segmentation is clearly a limitation of the concretely proposed approach. As shown, subsegmentation of airways influence the results decreasing the accuracy (Fig. 11), and analogously oversegmentation of airways could lead to classification errors in the AV separation introducing a bias towards arteries in the corresponding areas. The prerequisite of detecting bronchi down to 1 mm to get good results is a reasonable requirement for any airway segmentation algorithm.

Our experiments have shown that the framework performs very well if the image quality and the initial segmentation of airways guarantees that all the principal airway branches are segmented. However, as this dependency could hamper the application of the current framework to other datasets with less quality to ensure proper airway segmentation, in parallel to this work we have been exploring other pre-classification steps (Nardelli et al., 2018) that do not rely on an airway segmentation but exploit directly the local image content through convolutional neural networks (LeCun et al., 2015). Those results also report the improvement in performance thanks to the spatial consistency imposed by the graph-cut as detailed and described in this paper.

Errors in vessel segmentation could also affect the results of our method. Oversegmentation could influence the results only when false positives lay out in a given arrangement that could lead to a branch. Vessel subsegmentations are tackled by our algorithm thanks to the GC and the refined GC that ensure connectivity even in the presence of gaps and isolated trees. Recent developments of vessel segmentation techniques achieve high accuracy even for small vessels (> 97%, Rudyanto et al. (2014)) guaranteeing proper input to our algorithm. In this sense another possible future improvements includes the incorporation of global information in order to extend the current pre-classification of particles belonging to first vessel generations to more distal particles close to the alveoli. Some examples are proposed in Payer et al. (2016) and Charbonnier et al. (2016), including global features related to the organized peripheral distribution of arteries and veins. However, we decided not to incorporate this specific information to keep the method relatively robust against segmentation errors (miss-segmentations may destabilize the topology), abnormal AV distributions (pathological cases such as some isolated arterial/venous pulmonary malinosculations (Irodi et al., 2015), for example), and/or incomplete vessel information (lobe-based studies with partial vessel segmentations, such as those evaluated in the experiment of Section 3.4.4, as well as abnormal or surgically resected lungs). Further studies should be performed on data with these conditions as in this work we have only tackled pathological cases with COPD.

The performance of the algorithm has been tested on a set of 20 independent clinical cases showing good overall results specifically when the quality of the airway segmentation is good. We also considered the evaluation of the proposed method in the data from CARVE challenge (Charbonnier et al., 2016). However, the quality of the data would not guarantee a proper airway segmentation, as confirmed by imaging experts. Further analysis on data of higher quality would allow a more objective comparison of different methods on a bigger set. Additionally, the proposed methods could be used in CTPA studies including contrast media to enhance pulmonary vessels.

We would expect similar or even better results in this scenario, where the local intensity distributions of vessels would play an important role in the performance of the algorithm as it may distinguish the two irrigations. The extension of the evaluation database would also allow the exploration and validation of different applications of the AV segmentation in the clinical field. For instance, AV separation could be used to avoid false positive detections of PE using CAD systems, by removing venous locations that should not contain clots. Moreover, the proposed AV segmentation system, and its availability through the CIP, opens the possibility of studying different features such as density, distribution, structure, and size of arteries and veins quantitatively and objectively in an automated fashion. Diseases that imply vasculature remodeling, such as PAH, could use these measurements to check or even discover the phenotypical consequences on arterial and venous systems, that could lead to the definition of new biomarkers. The potential benefits for the analysis of pulmonary diseases in the future are undeniable.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This work has been supported by the National Institutes of Health, USA (R01HolL116473 , R01HL116931) and by the Spanish Ministry of Economy and Competitiveness (TEC2013-48251-C2-2-R). DJC and DBP were also supported by an FPU grant by the Spanish Ministry of Education.

References

Boykov Y, Veksler O, 2006 Graph cuts in vision and graphics: theories and applications In: Handbook of Mathematical Models in Computer Vision. Springer, pp. 79–96.

Boykov Y, Veksler O , Zabih R, 2001 Fast approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Mach. Intell. 23 (11), 1222–1239.

Breiman L, 2001 Random forests. Mach. Learn. 45 (1), 5–32.

Buelow T , Wiemker R , Blaffert T , Lorenz C , Renisch S , 2005 Automatic extraction of the pulmonary artery tree from multi-slice CT data In: SPIE Medical Imaging, pp. 730–740.

Charbonnier J-P, Brink M, Ciompi F, Scholten ET , Schaefer-Prokop CM, van Rikxoort EM, 2016 Automatic pulmonary artery-vein separation and classification in computed tomography using tree partitioning and peripheral vessel matching. IEEE Trans. Med. Imaging 35 (3), 882–892. [PubMed: 26584489]

Crosswhite P, Sun Z, 2014 Molecular mechanisms of pulmonary arterial remodeling. Mol. Med. 20

(20), 191–201. [PubMed: 24676136]

Delong A, Osokin A , Isack HN, Boykov Y , 2012 Fast approximate energy minimization with label costs. Int. J. Comput. Vis. 96 (1), 1–27,

Eppenhof K, Bekkers E, Berendschot TT, Pluim JP, ter Haar Romeny BM, 2015 Retinal artery/vein classification via graph cut optimization In: Proceedings of the Ophthalmic Medical Image Analysis Second International Workshop - OMIA 2015. University of Iowa, pp. 121–128.

Estépar RSJ, Ross JC, Kindlmann G, Diaz A, Okajima Y, Kikinis R, Westin C, Silverman E, Washko

G, 2012a Automatic airway analysis for genome-wide association studies in COPD. In: 9th IEEE International Symposium on Biomedical Imaging-ISBI 2012, pp. 1467–1470. doi:10.1109/ ISBI.2012.6235848.

Estépar RSJ , Ross JC, Krissian K, Schultz T , Washko GR, Kindlmann GL, 2012b Computational vascular morphometry for the assessment of pulmonary vascular disease based on scale-space particles. In: IEEE 9th International Symposium on Biomedical Imaging-ISBI 2012, pp. 1479– 1482.

Frangi AF, Niessen WJ, Vincken KL, Viergever MA, 1998 Multiscale vessel enhancement filtering In: Medical Image Computing and Computer-Assisted Interventation-MICCAI 1998, pp. 130–137.

Gao Z, Grout RW, Holtze C, Hoffman EA, Saha PK, 2012 A new paradigm of interactive artery/vein separation in noncontrast pulmonary CT imaging using multiscale topomorphologic opening. IEEE Trans. Biomed. Eng. 59 (11), 3016–3027. [PubMed: 22899571]

Gao Z, Holtze C, Grout R, Sonka M, Hoffman E, Saha PK, 2010 Multi-scale topo-morphometric opening of arteries and veins: an evaluative study via pulmonary CT imaging In: Advances in Visual Computing. Springer, pp. 129–138 .

Hsieh J , 2009 Computed Tomography: Principles, Design, Artifacts, and Recent Advances, PM188. SPIE Press,

Irodi A, Prabhu SM, John RA , Leena R, 2015 Congenital bronchopulmonary vascular malformations,sequestration and beyond. Ind. J. Radiol. Imaging 25 (1), 35,

Jimenez-Carretero D, Estepar RSJ, Cacio MD, Ledesma-Carbayo MJ, 2015a Automatic synthesis of anthropomorphic pulmonary CT phantoms - DICOM database. doi:10.5281/zenodo.32740.

Jimenez-Carretero D, Estepar RSJ, Cacio MD, Ledesma-Carbayo MJ, 2015b Automatic synthesis of anthropomorphic pulmonary CT phantoms - NRRD database. doi:10.5281/zenodo.20766.

Jimenez-Carretero D , Estepar RSJ, Cacio MD, Ledesma-Carbayo MJ, 2016 Automatic synthesis of anthropomorphic pulmonary CT phantoms. PLoS ONE 11 (1), e0146060. [PubMed: 26731653]

Joshi VS, Reinhardt JM, Garvin MK, Abramoff MD, 2014 Automated method for identification and artery-venous classification of vessel trees in retinal vessel networks. PLoS ONE 9 (2), e88061, [PubMed: 24533066]

Kitamura Y , Li Y , Ito W, Ishikawa H , 2013 Adaptive high-order submodular potentials for pulmonary artery-vein segmentation In: Proc. Fifth International Workshop on Pulmonary Image Analysis, pp. 53–61 .

Kitamura Y, Li Y, Ito W, Ishikawa H, 2016 Data-dependent higher-order clique selection for arteryvein segmentation by energy minimization. Int. J. Comput. Vis. 117 (2), 142–158,

LeCun Y, Bengio Y, Hinton GE, 2015 Deep learning. Nature 521 (7553), 436–444. doi:10.1038/ nature14539. [PubMed: 26017442]

Miller WS, 1947 The lung, Second Charles C. Thomas, Springfield, III.

Nardelli P, Jimenez-Carretero D, Bermejo-Pelaez D, Washko GR, Rahaghi FN, Ledesma-Carbayo MJ, Estpar RSJ, 2018 Pulmonary arteryvein classification in ct images using deep learning. IEEE

Trans. Med. Imaging 37 (11), 2428–2440. doi:10.1109/TMI.2018.2833385. [PubMed: 29993996]

Nardelli P, Ross JC, Estepar RSJ, 2017 CT image enhancement for feature detection and localization In: Medical Image Computing and Computer Assisted Intervention - MICCAI, pp. 224–232.

Park S, Bajaj C, Gladish G, 2006 Artery-vein separation of human vasculature from 3D thoracic CT angio scans. Comput. Model. Objects Present. Images 23–30.

Park S, Lee SM, Kim N, Seo JB, Shin H, 2013 Automatic reconstruction of the arterial and venous trees on volumetric chest CT. Med. Phys. 40 (7), 071906, [PubMed: 23822443]

Payer C, Pienn M, Balint Z, Shekhovtsov A, Talakic E, Nagy E, Olschewski A, Olschewski H, Urschler M, 2016 Automated integer programming based separation of arteries and veins from thoracic CT images. Med. Image Anal. 34, 109–122 . [PubMed: 27189777]

Payer C , Pienn M , Bálint Z , Olschewski A, Olschewski H, Urschler M, 2015 Automatic artery-vein separation from thoracic CT images using integer programming In: Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015, pp. 36–43.

Regan EA , Hokanson JE, Murphy JR, Make B, Lynch DA, Beaty TH, Curran-Everett D, Silverman

EK, Crapo JD, 2011 Genetic epidemiology of COPD (COPDGene) study design. COPD 7 (1), 32– 43,

Ross JC, Estepar RSJ, Diaz A, Westin C-F, Kikinis R, Silverman EK, Washko GR, 2009 Lung extraction, lobe segmentation and hierarchical region assessment for quantitative analysis on high resolution computed tomography images In: Medical Image Computing and Computer-Assisted Intervention-MICCAI 2009, pp. 690–698.

Rudyanto RD, Kerkstra S , van Rilowort EM , Fetita C , Brillet P-Y, Lefevre C, Xue W,Zhu X, Liang J Oksuz I , Unay D, Kadipasoglu K, Estepar RSJ ,Ross JC, Washko GR, Prieto J-C, Hoyos MH, Orkisz M, Meine H , Huellebrand M, Stoecker C , Mir FL , Naranjo V , Villanueva E , Staring M ,

Xiao C , Stoel B.C., m. Fabijanska A , Smistad E, Elster AC, Lindseth F, Foruzan AH, Kiros R,

Popuri K, Cobzas D, Jimenez-Carretero D, Santos A, Ledesma-Carbayo MJ, Helmberger M,

Urschler M , Pienn M, Bosboom DGH, Campo A, Prokop M , de Jong PA, de Solorzano CO , Munoz-Barrutia A, van Ginneken B, 2014 Comparing algorithms for automated vessel segmentation in computed tomography scans of the lung: the VESSEL12 study. Med. Image Anal. 18 (7), 1217–1232. [PubMed: 25113321]

Saha PK, Gao Z, Alford SK, Sonka M, Hoffman EA , 2010 Topomorphologic separation of fused isointensity objects via multiscale opening: separating arteries and veins in 3-D pulmonary CT. IEEE Trans. Med. Imaging 29 (3), 840–851, [PubMed: 20199919]

Sluimer I , Schilham A, Prokop M , van Ginneken B, 2006 Computer analysis of computed tomography scans of the lung: a survey. IEEE Trans. Med. Imaging 25 (4), 385–405. [PubMed: 16608056]

Tozaki T, Kawata Y, Niki N, Ohmatsu H, 2001 Extraction and classification of pulmonary organs based on thoracic 3D CT images. Syst. Comput. Jpn. 32 (9), 42–53.

Tozaki T, Kawata Y , Niki N , Ohmatsu H , Kakinuma R , Eguchi K , Kaneko M , Moriyama N, 1998 Pulmonary organs analysis for differential diagnosis based on thoracic thin-section CT images. IEEE Trans. Nucl. Sci. 45 (6), 3075–3082.

Van Rikxoort EM, Van Ginneken B, 2013 Automated segmentation of pulmonary structures in thoracic computed tomography scans: a review. Phys. Med. Biol. 58 (17), R187–R220 . [PubMed: 23956328]

Wala J, Fotin S, Lee J, Jirapatnakul A, Biancardi A, Reeves A, 2011 Automated segmentation of the pulmonary arteries in low-dose CT by vessel tracking. arXiv:1106.5460.

Yonekura T , Matsuhiro M , Saita S , Kubo M , Kawata Y , Niki N , Nishitani H , Ohmatsu H, Kakinuma R, Moriyama N, 2007 Classification algorithm of pulmonary vein and artery based on multi-slice CT image. In: Proc. SPIE Vol. 6514, Medical Imaging, p. 65142E,

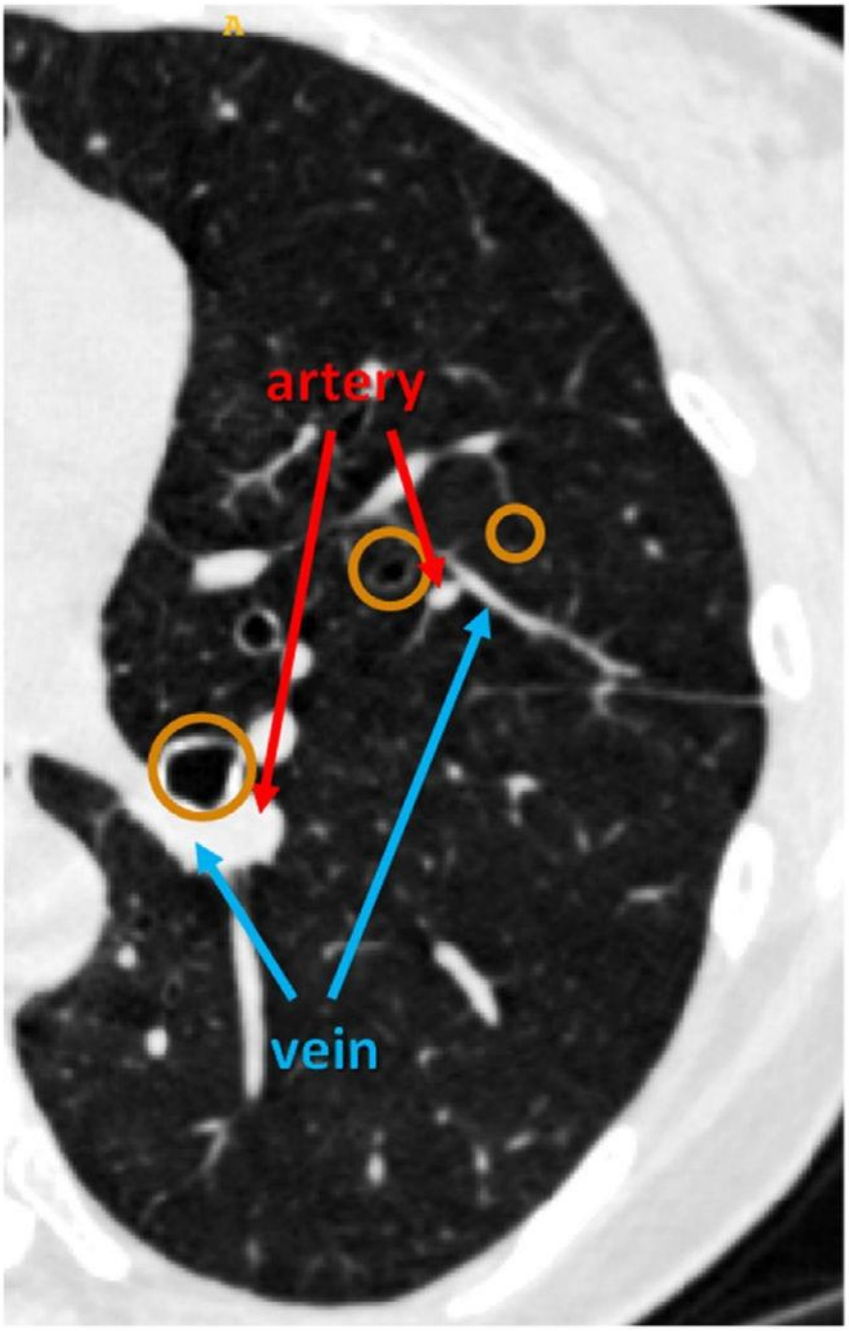

Fig. 1.

Example of noncontrast pulmonary CT image, where arteries and veins are closely situated at different generation levels and share the same intensity values. Airways (orange circles) become indistinguishable at high generations because of partial volume effects that avoid the visualization of airway walls. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

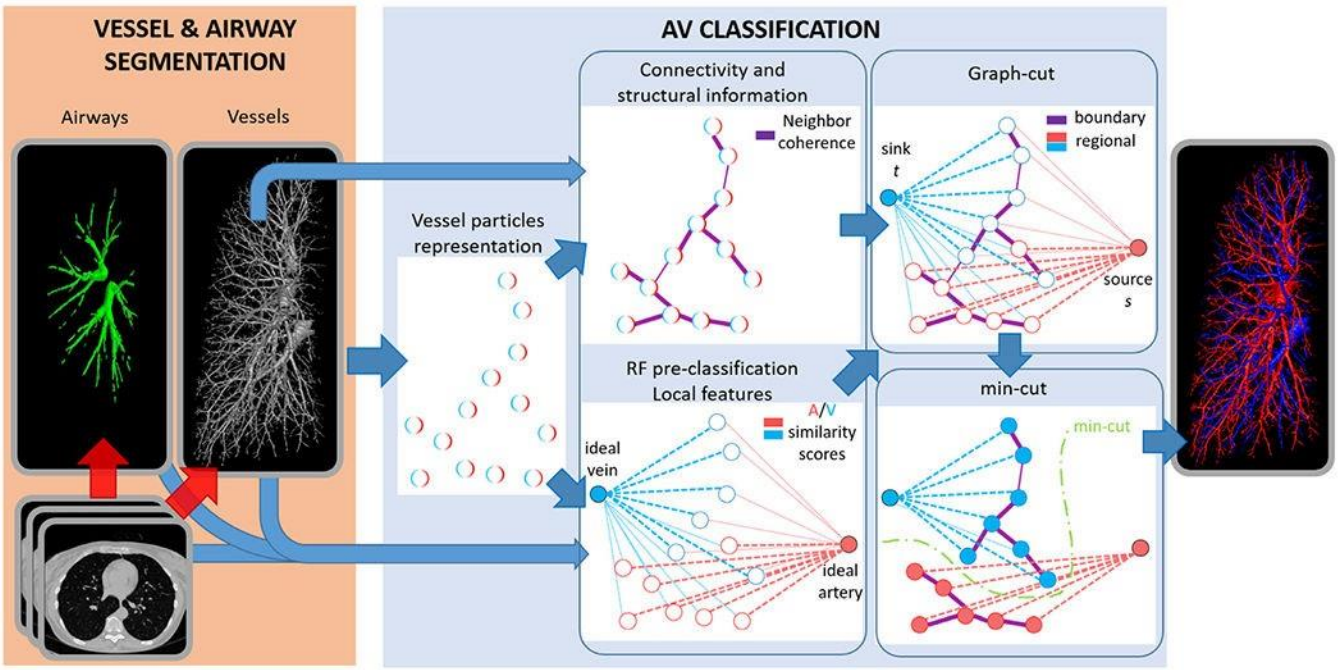

Fig. 2.

Graphical scheme of the proposed framework for the automatic AV segmentation algorithm. AV classification is performed over a set of scale-space vessel particles coming from the initial segmentation. Connectivity and structural information is used to compute neighbour coherence, forming the boundary term (n-links, purple lines); and RF pre-classification based on local features extracted from vascular and bronchial information is used to compute arterial/venous similarity scores, forming the regional term (t-links , red/blue links) of the graph-cut. The final AV classification is obtained by computing the minimum-cut. Circles depict vessel particles, thickness of connections represent the weight of the links, and circumference colors show the preferential class for each particle. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

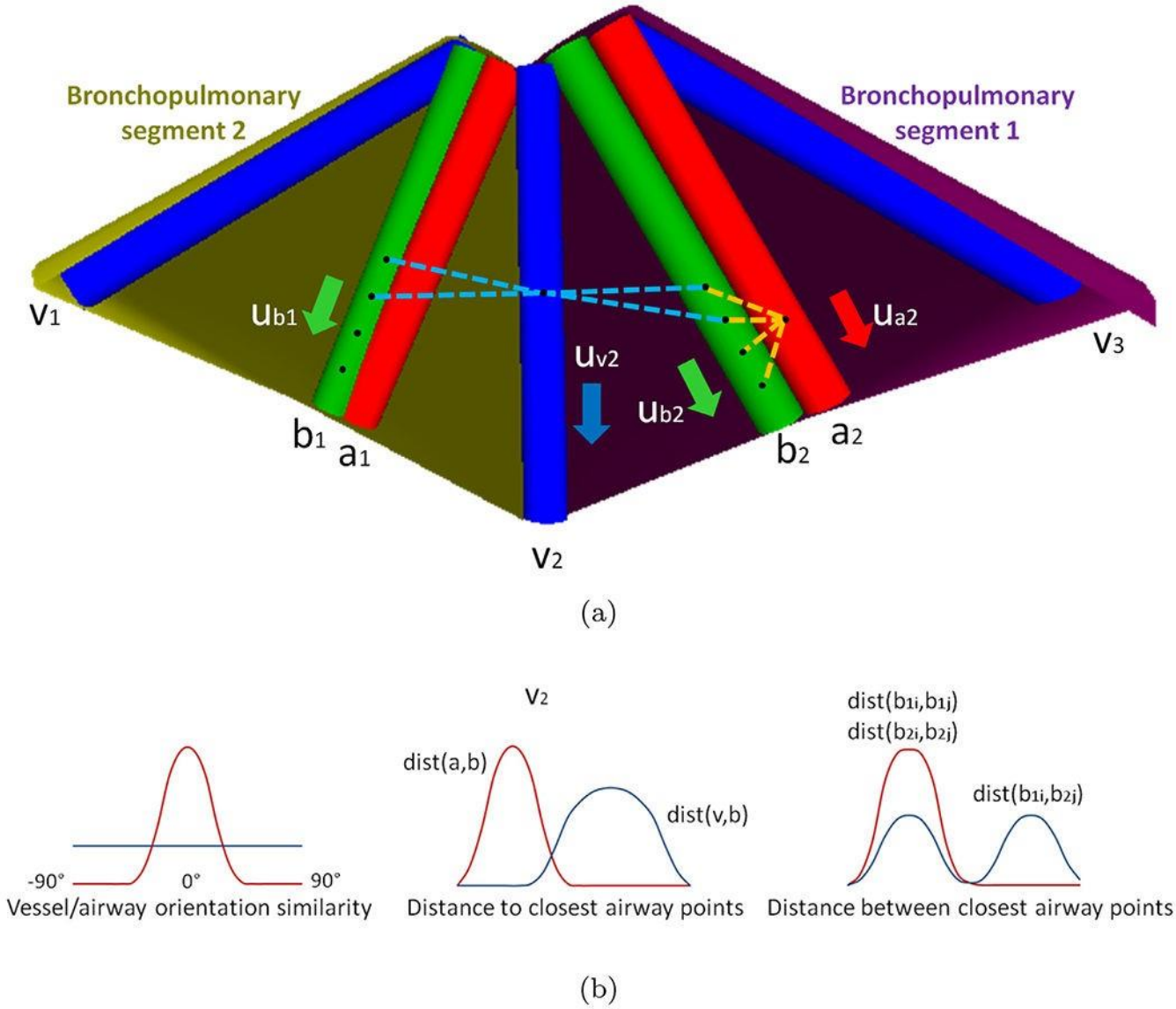

Fig. 3.

a) Representation of two bronchopulmonary segments displaying the spatial relationship between arteries a_i (red), veins v_i (blue) and airways b_i (green), and their preferential directions u_{ai}, u_{vi} , and u_{bi} , respectively. b) Distributions for arteries and veins are different for different vessel features, such as orientation with respect to airways, distance to closest airway locations, and distance between these closest bronchial locations. Given a particle in a vessel the N_{neigh} closest airway points are computed. In this scheme, blue dotted lines identify those points (N_{neigh}=4) for a vein particle, and the orange dotted lines the ones for an artery particle. As it can be observed, the mean distance to N_{neigh} points and the mean distance among them are smaller for artery particles than for vein particles. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

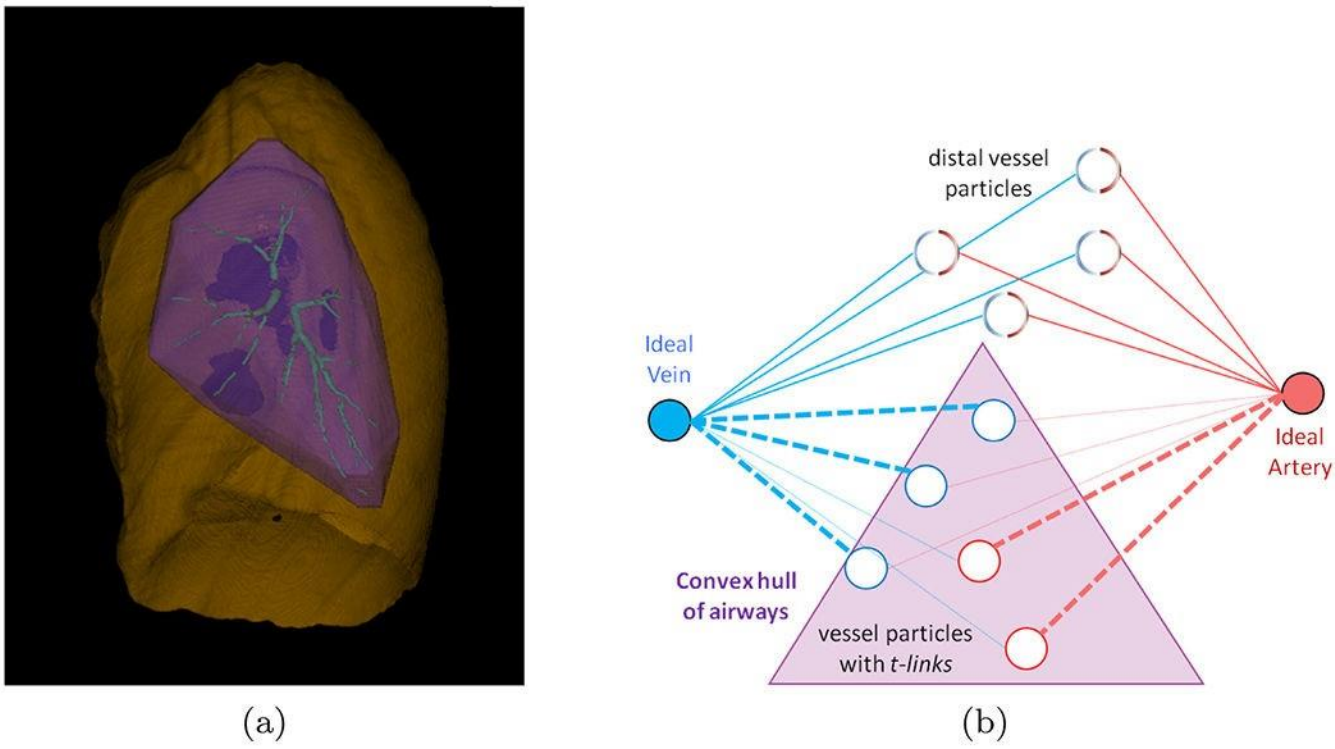

Fig. 4.

a) The construction of the regional term is constrained to the convex hull, plus a security margin (purple volume), of the airway segmentation (green structure). b) Example with a schematic GC representation of the regional term where vessel particles (circles) inside the convex hull of the bronchial tree (purple area) are connected with the two special nodes representing ideal artery and vein with weights determined by AV similarity scores; and particles outside the convex hull do not incorporate any preferential AV classification in their

t-links ($w_{reg}(p_i, t) = (w_{reg}(p_i, s) = 0.5$ - solid thick lines). Thickness of connections

represent the weight of the links and circumference colors show the preferential class for each particle. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

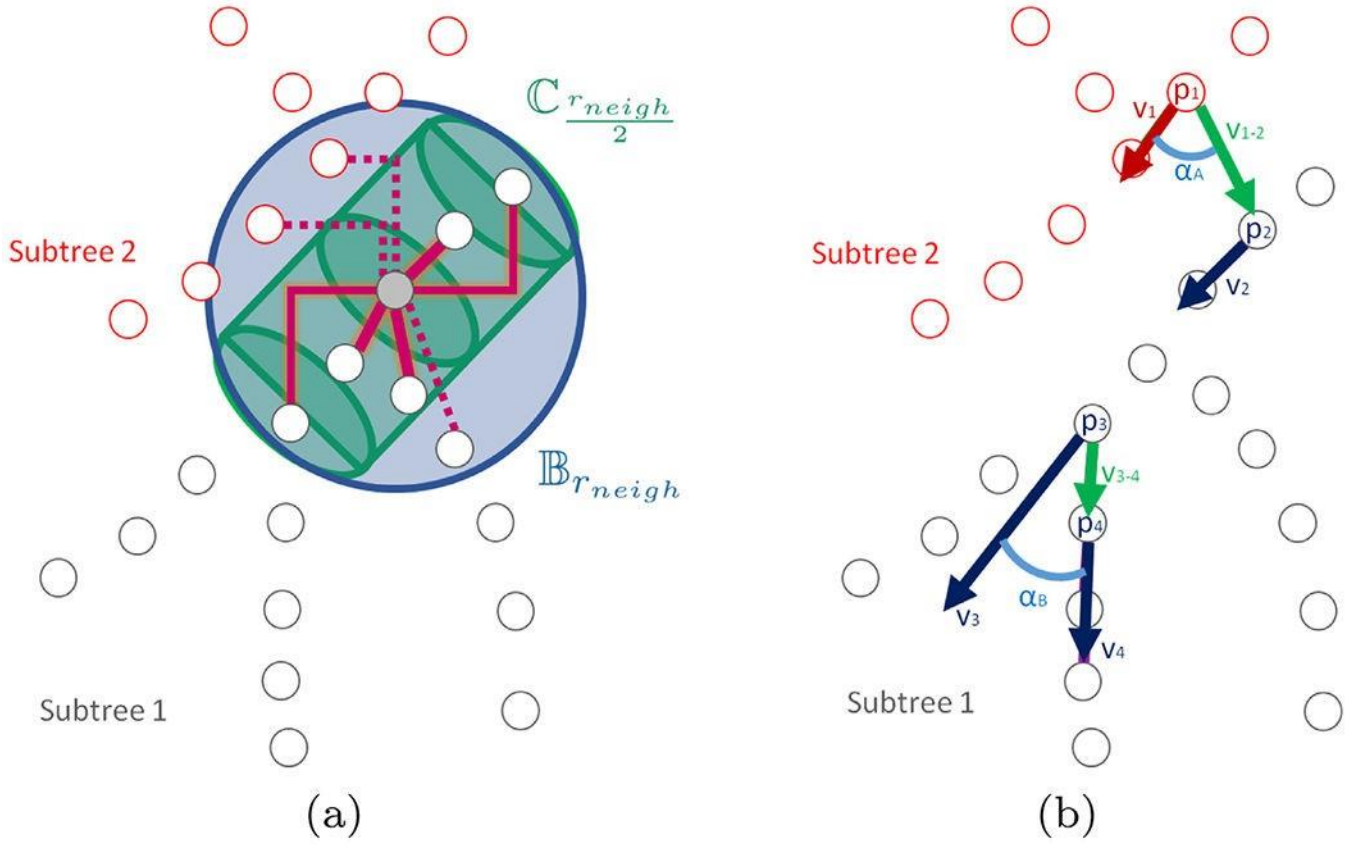

Fig. 5.

a) Anisotropic connection strategy using a cylindrical confinement (green) along the preferential local direction of vessels impede inter-irrigation connections (dashed lines) between vessel particles (circles) that could be connected when using an isotropic strategy (blue area). b) Connectivity vector between particles can improve the definition of edge weights based on parallelism between preferential directions. Intertwined structures could report unreal parallelism between preferential directions (v_1 and v_2) of particles belonging to different irrigations (p_1 and p_2) that can be minimized using the direction of connectivity vector (v_{1-2}). The instability in the definition of preferential directions (v_3) in bifurcations (p_3) may mask real coorientations (p_3 and p_4) that can be enhanced using the connectivity vector (v_{3-4} with v_4). For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

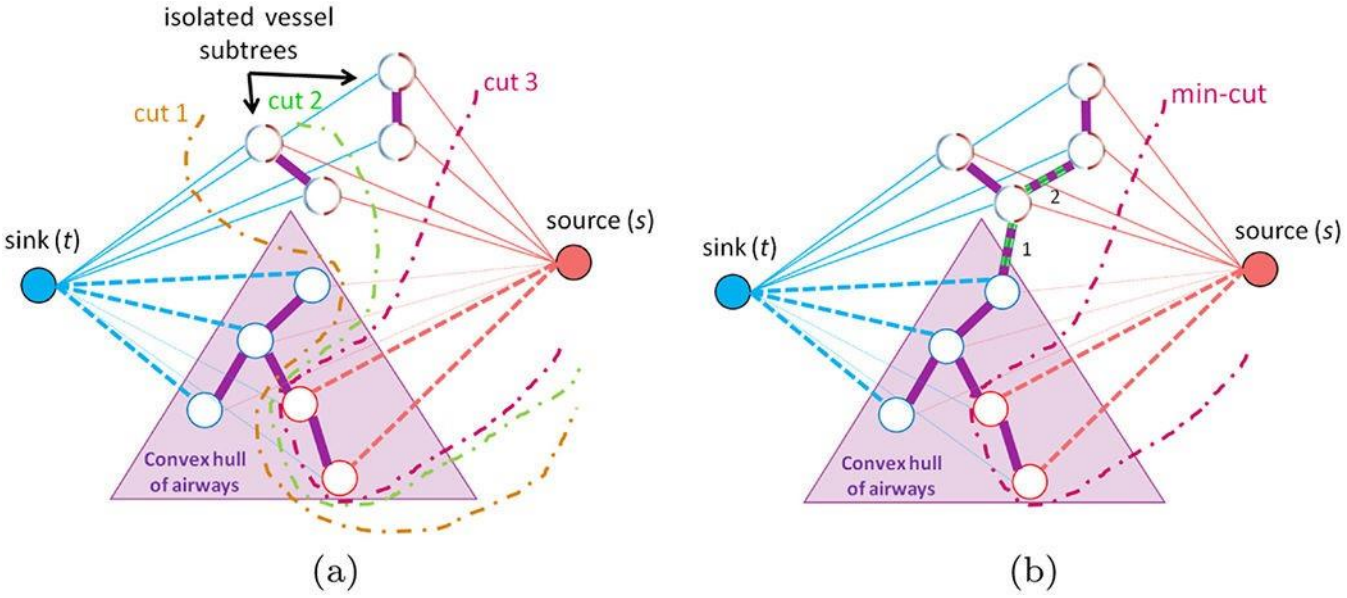

Fig. 6.

a) Example of a forest with three connected components showing two isolated subtrees located outside the convex hull of airways (purple area). Consequently, the three cuts displayed need the same energy to be performed, rendering their selection as min-cut equally probable. b) Refinement of the graph results on the addition of new edges 1 and 2 (purplegreen dashed lines) and the uniqueness of min-cut. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

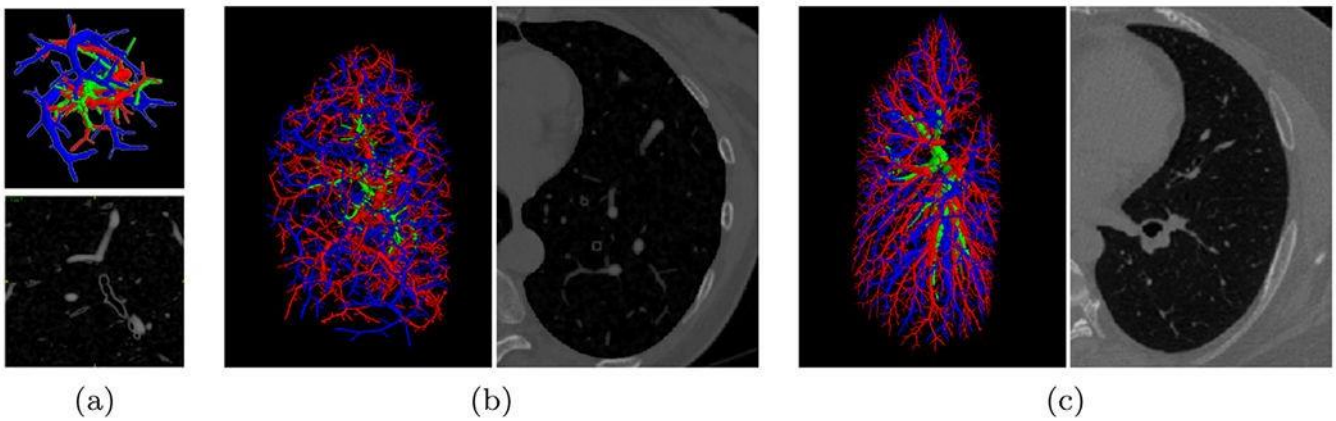

Fig. 7.

Examples of reference standards and CT images of bronchopulmonary phantom (a), anthropomorphic pulmonary CT phantom (b), and clinical noncontrast CT case (c). Arteries, veins, and airways are labelled in red, blue, and green, respectively. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

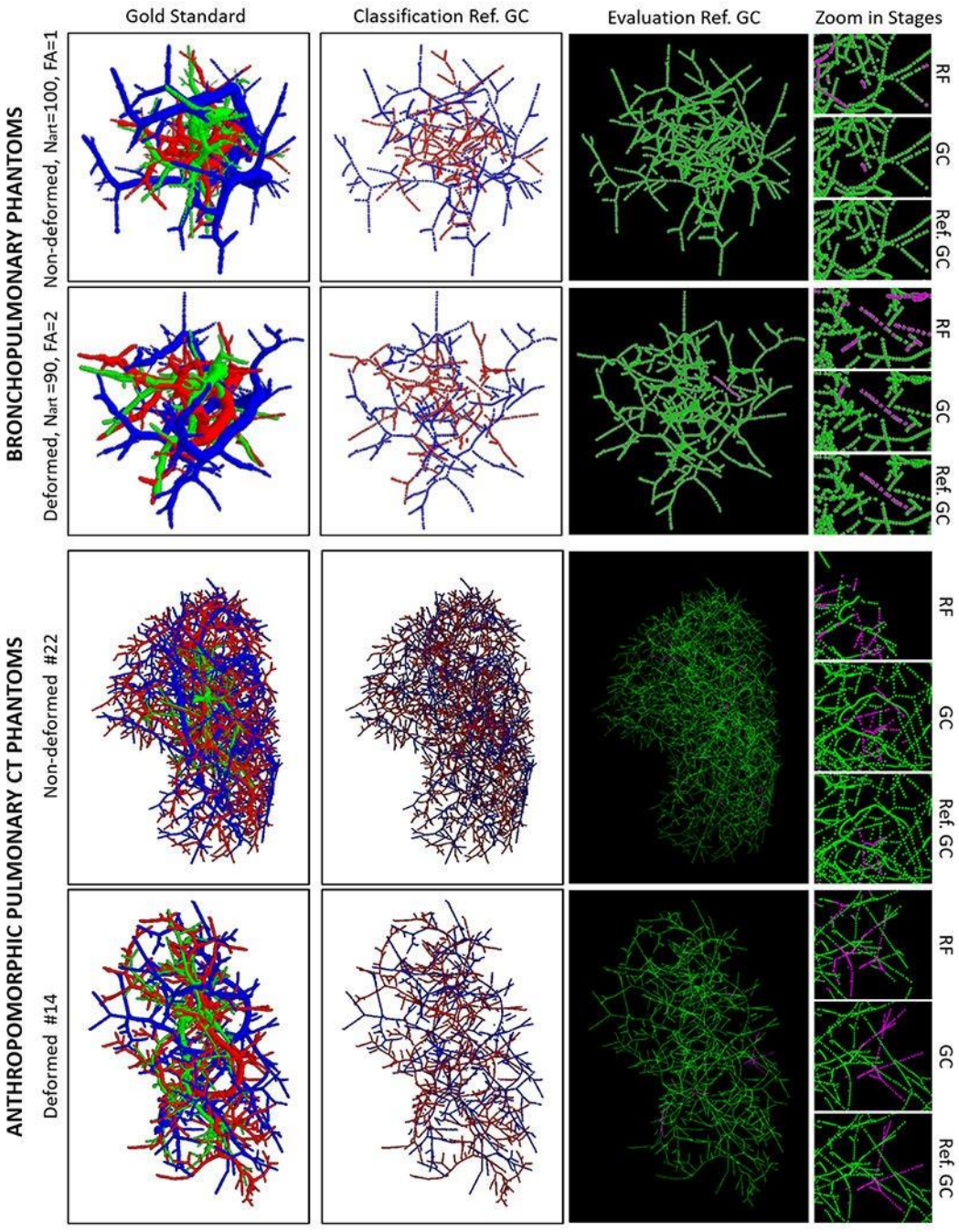

Fig. 8.

Example of AV segmentations in synthetic cases. Each row displays from left to right: a volumetric representation of the reference standard (showing scale of particles), the final particle-based AV classification, its evaluation against the gold standard, and a zoom on a specific area showing some improvements of results in subsequent stages of the algorithm. In the first two columns, arteries, veins and airways are shown in red, blue, and green respectively. Red: arteries, blue: veins, green: correct classification (airways in first column), pink: misclassification. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

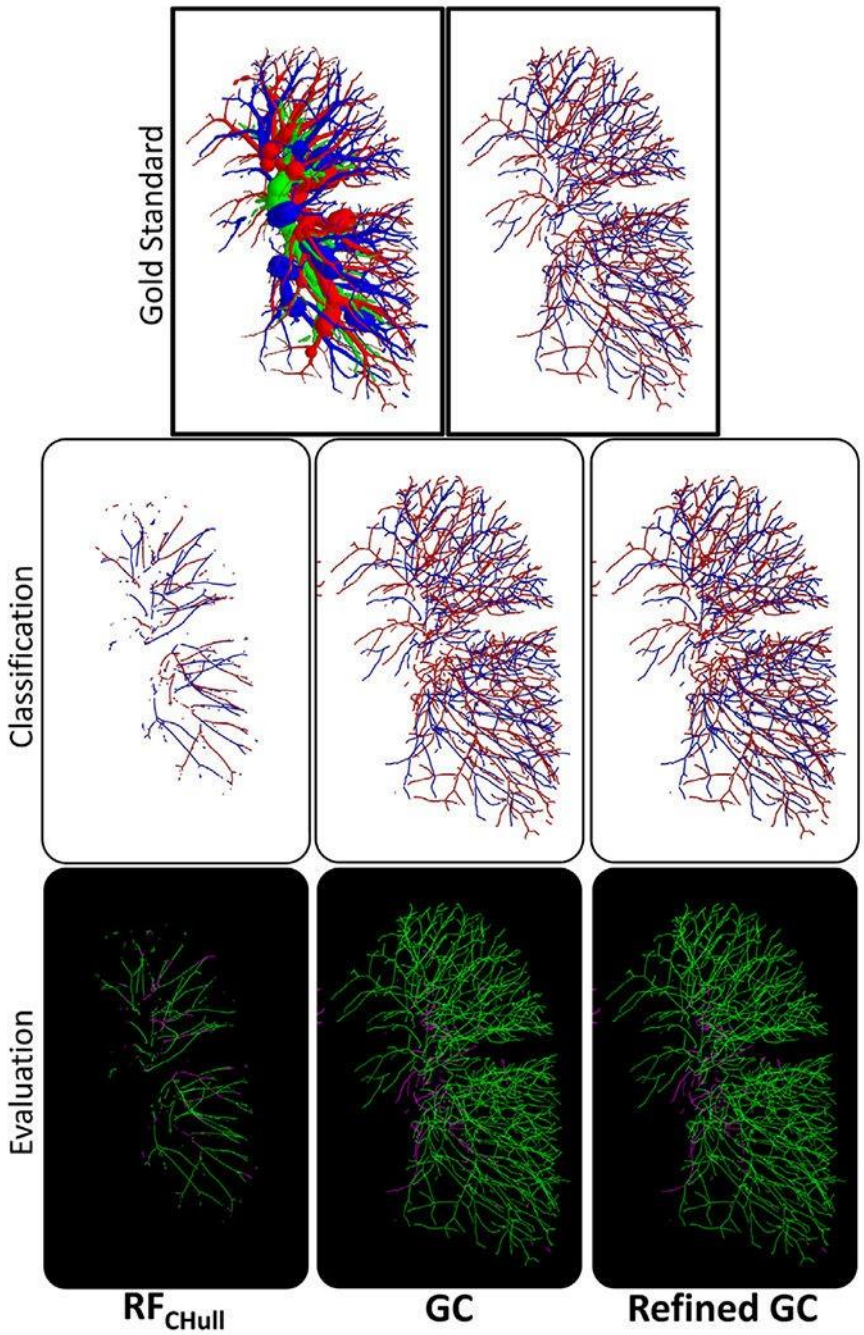

Fig. 9.

Example of AV segmentation in clinical noncontrast CT #5 (best result). First row:

volumetric and particle-based representations of the reference standard (red: arteries, blue: veins, green: airways). The second row: AV classification after each stage: RF, GC, and refined GC. Third row: evaluations (green: correctly classified, pink: errors).

Misclassifications are located in some small-calibre subtrees, and in areas close to the hilum where all air and blood flows run very closely. Moreover, our method shows a robust behaviour in the presence of segmentation inaccuracies. F1 scores: RF = 82.7%, GC =

95.5%, Refined GC = 97.1%. The electronic version contains high-resolution color images to ease visualization of results. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

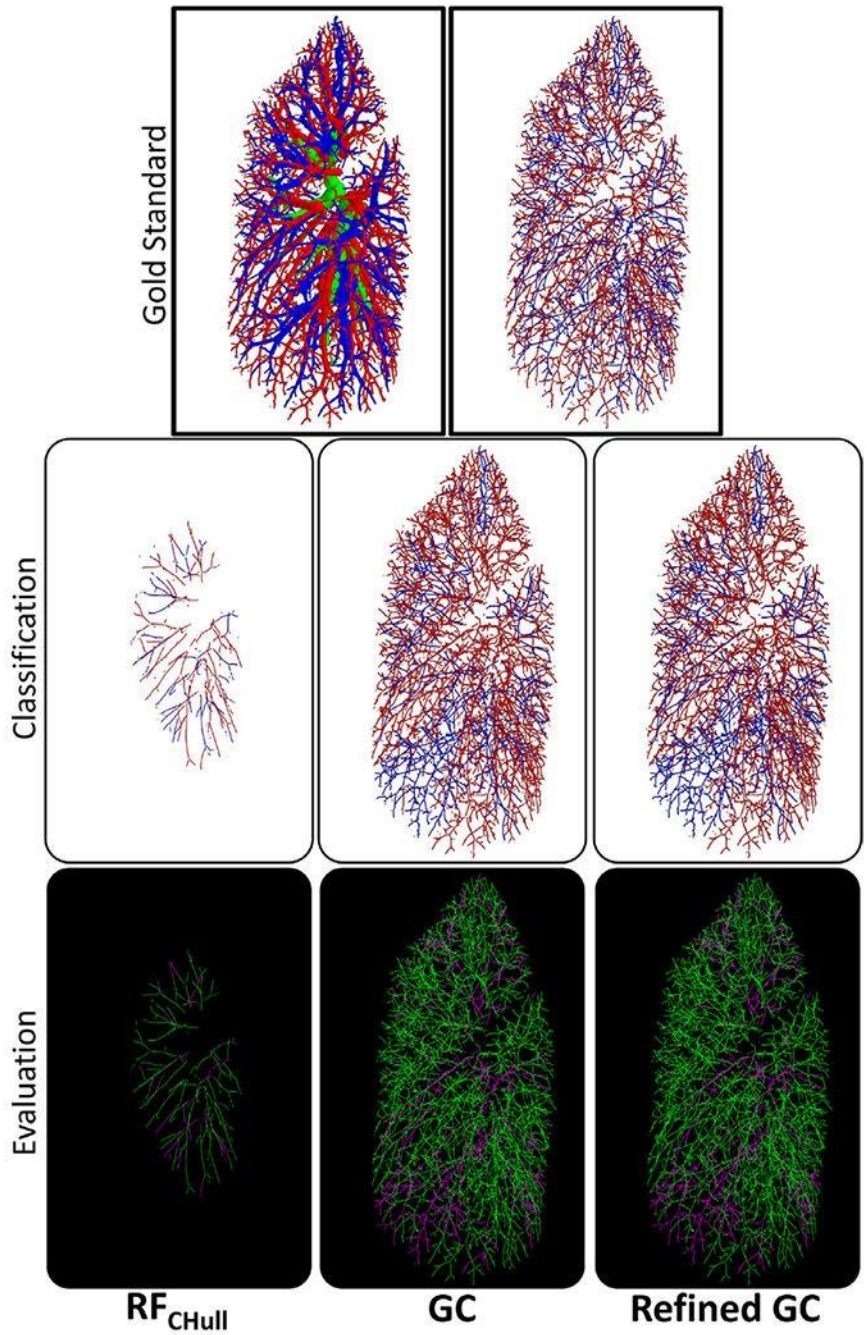

Fig. 10.

Example of AV segmentation in clinical noncontrast CT #1 (worst result). In this case, misclassifications located in important vessels in RF stage limited the correct behaviour of the GC strategy, which is not able to correct them and propagates the errors to other particles. F1 scores: RF = 74.8%, GC = 78%, Refined GC = 81%. Electronic version contains high-resolution color images to ease the visualization of the results.

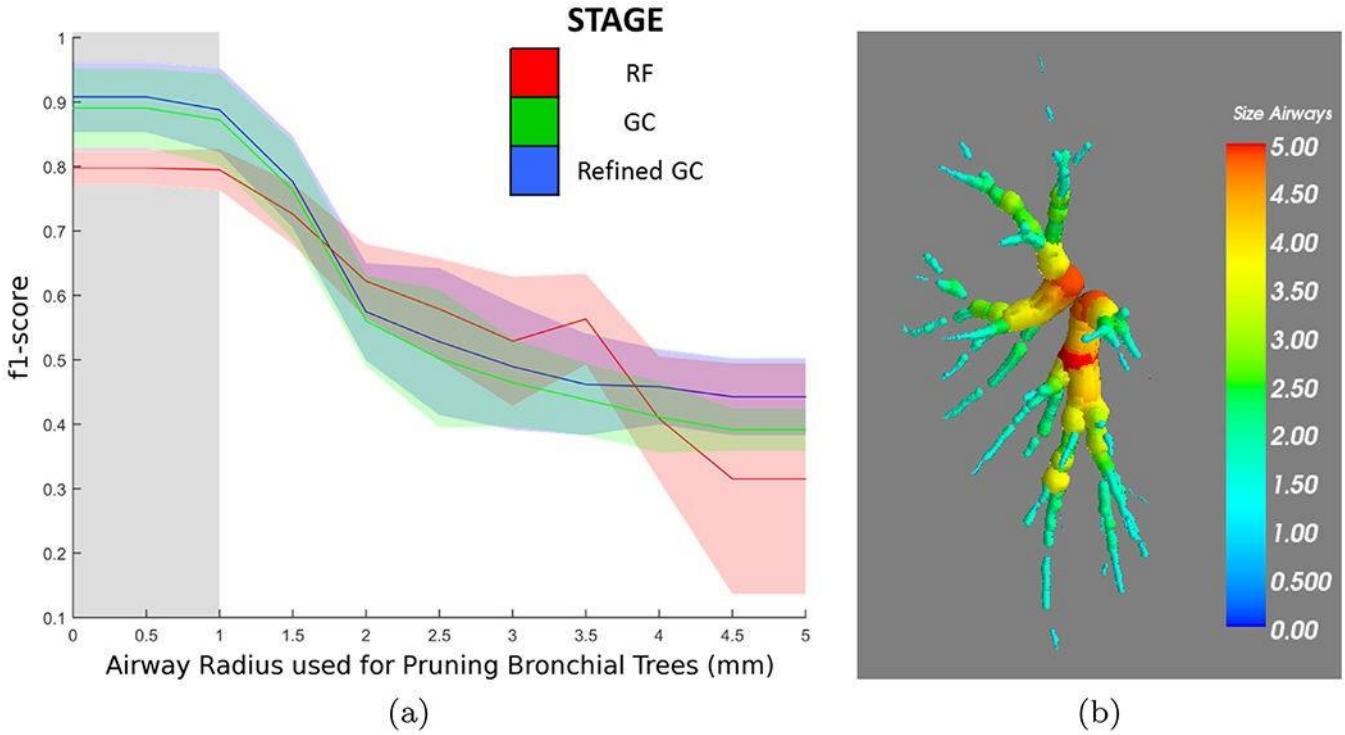

Fig. 11.

Importance of airways segmentation in AV results. a) Evolution of F1-scores when pruning airway segmentations by removing particles with scales lower than a specific value. Mean values (lines) and standard deviations (shadows) are displayed for the three evaluation stages. b) Example of an airway tree where the scale of particles at each bronchial location is represented.

Table 1

on bronchopulmonary segment phantoms. Mean F1-score values [0–100] and standard deviations are reported for both nondeformed ions of the phantoms and three densities of airways (FA), in the three evaluation stages: random forest (RF), graph-cuts (GC), and final

			Deformed		
	Ref. GC	RF	GC	Ref. GC
1.9	98.0 ± 1.9	83.3 ± 2.2	97.1 ± 2.7	97.4 ± 2.7
3.7	96.3 ± 3.8	78.2 ± 3.1	95.5 ± 4.4	95.5 ± 4.5
6.4	94.3 ± 6.6	74.5 ± 4.1	91.3 ± 6.7	92.0 ± 6.6

Table 2

ation results on anthropomorphic pulmonary CT phantoms. Mean F1-score values [0–100] and standard deviations are reported for both eformed and deformed versions using leave-one-out CV and 2-fold (2F) strategies, in the three evaluation stages

Nondeformed			Deformed		
RF	GC	Ref. GC	RF	GC	Ref. GC
73.4 ± 3.0	92.5 ± 3.5	93.4 ± 3.0	72.3 ± 3.4	90.3 ± 6.7	92.3 ± 5.9
73.6 ± 3.4	92.4 ± 3.6	94.0 ± 3.0	72.8 ± 3.7	90.9 ± 6.0	92.9 ± 4.1

Table 3

Evaluation results on clinical noncontrast CT cases. F1 scores are reported (leave-one-out and three-fold CV leaving one subject out) for the three evaluation stages: random forest (RF), graph-cuts (GC) and final refined graph-cuts (Ref. GC). The last two rows show the mean and standard deviation values.

	CV				3F			
Case	**RF_{CHull}**	**RF**	**GC**	**Ref. GC**	**RF_{CHull}**	**RF**	**GC**	**Ref. GC**
1	74.8	56.8	78.0	81.0	74.3	55.2	76.7	79.8
2	79.6	65.5	87.7	90.8	76.4	63.8	87.5	90.6
3	80.3	59.6	88.7	89.4	79.8	59.7	88.9	89.7
4	80.2	57.9	92.9	93.7	80.0	57.3	92.1	93.0
5	82.7	66.7	95.5	97.1	79.3	64.4	92.0	93.7
6	81.2	60.1	91.4	92.8	79.2	59.2	87.3	87.7
Mean	79.8	61.1	89.1	90.8	78.2	59.9	87.4	89.1
Std	2.4	4.1	5.6	5.0	2.1	3.3	5.2	4.6

Table 4

Evaluation results on the independent and partially AV-labelled dataset of clinical noncontrast CT cases. Mean F1-score values [0–100] and standard deviations are reported in the three evaluation stages categorized by the quality of the airway tree segmentation as Complete, Partial and Incomplete.

Airway segmentation	RF	RF	GC	Ref_GC
Complete (10 cases)	79.3 ± 4.8	56.1 ± 2.1	79.8 ± 6.2	83.0 ± 6.2
Partial (7 cases)	76.0 ± 6.2	53.8 ± 3.3	76.8 ± 10.3	79.1 ± 9.4
Incomplete (3 cases)	72.6 ± 10.5	53.3 ± 1.5	70.1 ± 8.9	71.4 ± 8.9
All	76.9 ± 6.6	54.6 ± 2.8	77.0 ± 8.7	79.5 ± 8.6