Automatic Pulmonary Artery-Vein Separation and Classification in Computed Tomography Using Tree Partitioning and Peripheral Vessel Matching

Jean-Paul Charbonnier*, Monique Brink, Francesco Ciompi, Ernst T. Scholten, Cornelia M. Schaefer-Prokop, and Eva M. van Rikxoort

***Abstract*—We present a method for automatic separation and classification of pulmonary arteries and veins in computed tomography. Our method takes advantage of local information to separate segmented vessels, and global information to perform the artery-vein classification. Given a vessel segmentation, a geometric graph is constructed that represents both the topology and the spatial distribution of the vessels. All nodes in the geometric graph where arteries and veins are potentially merged are identified based on graph pruning and individual branching patterns. At the identified nodes, the graph is split into subgraphs that each contain only arteries or veins. Based on the anatomical information that arteries and veins approach a common alveolar sag, an arterial subgraph is expected to be intertwined with a venous subgraph in the periphery of the lung. This relationship is quantified using periphery matching and is used to group subgraphs of the same artery-vein class. Artery-vein classification is performed on these grouped subgraphs based on the volumetric difference between arteries and veins. A quantitative evaluation was performed on 55 publicly available non-contrast CT scans. In all scans, two observers manually annotated randomly selected vessels as artery or vein. Our method was able to separate and classify arteries and veins with a median accuracy of 89%, closely approximating the inter-observer agreement. All CT scans used in this study, including all results of our system and all manual annotations, are publicly available at "http://arteryvein.grand-challenge.org".**

***Index Terms*—Artery-vein separation, automatic classification, chest computed tomography, pulmonary vessels.**

I. Introduction

PULMONARY vessels are affected in a multitude of diseases, with effects on both respiratory and cardiac function. While pulmonary embolism or arteriovenous (AV) malformations are primary vascular diseases, the pulmonary vessels are also secondarily affected by diffuse parenchymal lung diseases such as emphysema or interstitial lung diseases, which represent the cause of significant patient morbidity. Analysis of pulmonary vessels is therefore of high interest for clinical diagnosis.

Computed tomography (CT) is the most sensitive way for in-vivo imaging of the thorax and is the modality of choice for analysis of the lungs. CT provides images with near-isotropic sub-millimeter resolution, typically consisting of over 400 slices, which enables a detailed analysis of the pulmonary vessels. However, this large amount of data makes clinical review by radiologists time consuming and visually demanding, which might therefore be prone to diagnostic mistakes. Automatic and semi-automatic methods to segment and analyze vascular structures are potentially important tools to assess image information more accurately and effectively.

Pulmonary blood circulation consists of two vascular trees: an arterial tree that supplies carbon dioxide rich blood to the lung, and a venous tree that channels the oxygen rich blood towards the heart. The arteries and veins that are visible on CT are a morphologically intertwined vascular network. They are connected by means of capillary beds surrounding numerous pulmonary alveoli, at which level the actual exchange between oxygen-rich and oxygen-poor blood takes place. These capillary beds however, are indistinguishable on CT images since their size is below the image resolution.

The goal of this study is to develop an algorithm to automatically separate and classify pulmonary vessels into arterial and venous on CT. This is relevant because pulmonary diseases specifically affect either the pulmonary arteries or veins, or both but in a different manner. Automatic classification of the vascular tree is potentially important for both the assessment of focal disease, as well as quantification of diffuse changes of the vascular tree. For example, pulmonary embolism is characterized by thrombi in the pulmonary artery. Previous evaluations of computer aided detection algorithms of pulmonary thrombi were hampered by high numbers of false positive findings in pulmonary veins [1]. Accurate separation of veins from arteries will automatically prevent these false positives and improve the accuracy of such systems. Another potential clinical application includes assessment of the muscular arteries in pulmonary arterial hypertension, or the assessment of arterial alterations associated with chronic obstructive pulmonary diseases (COPD) [2]. These more diffuse alterations are specifically difficult to

Manuscript received September 09, 2015; revised November 03, 2015; accepted November 03, 2015. Date of publication November 12, 2015; date of current version March 01, 2016. This work was supported by the Netherlands Organisation for Scientific Research (NWO, project 612.001.204). *Asterisk indicates corresponding author.*

*J.-P. Charbonnier is with the Diagnostic Image Analysis Group, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands (e-mail: jean-paul.charbonnier@radboudumc.nl).

M. Brink is with the Department of Radiology and Nuclear Medicine, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands.

F. Ciompi, E. T. Scholten, and E. M. van Rikxoort are with the Diagnostic Image Analysis Group, Radboud University Medical Center, 6525 GA, Nijmegen, The Netherlands.

C. M. Schaefer-Prokop is with the Department of Radiology and Nuclear Medicine, Radboud University Medical Center, 6525 GA Nijmegen, The Netherlands, and also with the Department of Radiology, Meander Medisch Centrum, 3813 TZ Amersfoort, The Netherlands.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2015.2500279

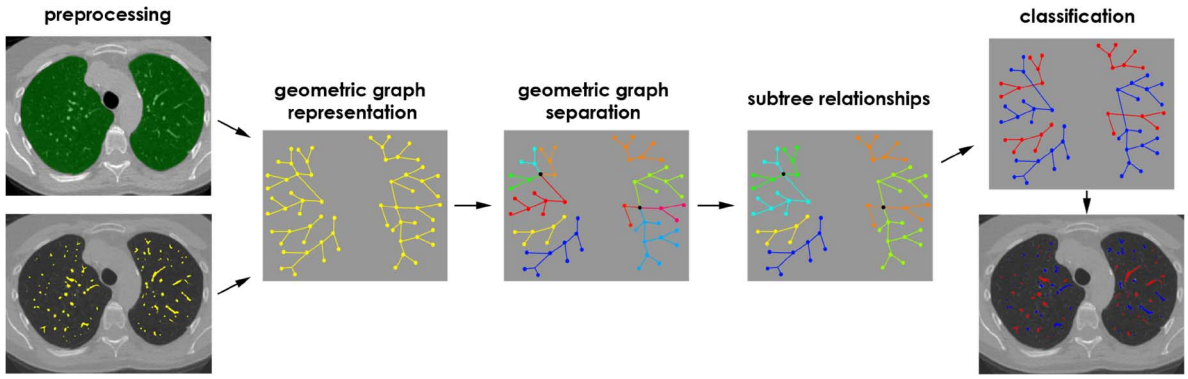

Fig. 1. Schematic overview of the presented method.

quantify for human experts without the aid of (semi-)automated software.

Arteries and veins locally appear to touch each other at multiple locations throughout a CT scan due to the relatively limited scan resolution and the high spatial density of pulmonary vessels. Partial volume effects further aggravate this by decreasing the spatial resolution. In the past, this has proven to complicate the automatic separation of arteries and veins [3]–[5]. The large inter-subject variation in vascular branching patterns additionally complicates automatic classification of arteries and veins because a large amount of bifurcations and trifurcations need to be considered, with different branching angles and diameters.

A substantial amount of research has been devoted to pulmonary vessel segmentation methods as described in a review by [6]. All available vessel segmentation algorithms attach arteries and veins at locations where arteries and veins locally appear to touch each other within the resolution limits of the CT image. These algorithms are not designed to separately extract arterial and venous trees. Only a small group of authors attempted to tackle the challenge of actually separating pulmonary arteries from veins [3]–[5], [7]. Several authors focused on methods that separate these trees explicitly by attempting to determine which branches represent the right continuation at these locations [3]–[5]. In [5], the authors argued that in the vessel segmentation, attached arteries and veins can be locally untangled by comparing the diameters of the attached vessels. In [4], a method was proposed based on a morphological multiscale opening approach to separate locations of attached arteries and veins. Recently, [3] presented a method to reconstruct vascular trees from an extracted set of vascular points, based on connection strengths between these points. In [7], a method was presented that attempted to immediately classify arteries and veins, without first separating the trees. The authors identified arteries based on co-orientation with the bronchial tree, since only the arterial tree accompanies the airways.

The methods described in [3]–[5] require a substantial amount of manual interactions to separate the vasculature. In addition, the actual classification of the separated trees into arteries and veins was done manually. The algorithm described in [7] automatically classified vessels into arteries or veins. However the anatomical information used for this classification (i.e. co-orientation of arteries with the bronchial tree), is only reliably available in the more central parts of the lung where the airways are still visible.

All currently published artery-vein classification methods lack an extensive evaluation of their method. The two methods described in [5], [7] did not include any evaluation on CT scans, while the two methods in [3], [4] only provide a limited quantitative evaluation on two CT scans. This makes it very difficult to draw any meaningful conclusions with respect to their performance on CT scan and impossible to compare algorithm of different authors.

We propose a fully automatic artery-vein separation and classification method in which we use local information to separate the vasculature and global information to classify the arteries and veins. A quantitative evaluation was performed by assessing agreement between human observers and the results of the method in 55 non-contrast volumetric low dose thoracic CT scans. To the best of our knowledge, this is the most extensive evaluation of pulmonary artery-vein classification to date. To allow for comparison with other methods, all scans, results, and expert annotations are made publicly available at http://arteryvein.grand-challenge.org.

II. Methods

The outline of our method is shown in Fig. 1. We first segment the lungs and vessels from a thoracic CT image (Section II-A) and convert the vessel segmentation into a geometric graph representation (Section II-B). Local information is used to detect attached arteries and veins in the geometric graph. Based on these detected locations the geometric graph is separated into a set of small graphs that each represents a subtree belonging to a single artery-vein class (Section II-C). The task of artery-vein classification is thus reduced to classifying subtrees. Global information is used to extract subtree class relationships, in order to link subtrees of the same, yet unknown, artery-vein class (Section II-D). These class relationships are established by using the anatomical knowledge that arteries and veins approach a common alveolar sag. Linked subtrees are classified into artery or vein by using the volume difference as a discriminative feature (Section II-F).

A. Preprocessing

The proposed method requires a binary lung and vessel segmentation. The lung segmentation L is extracted as described by

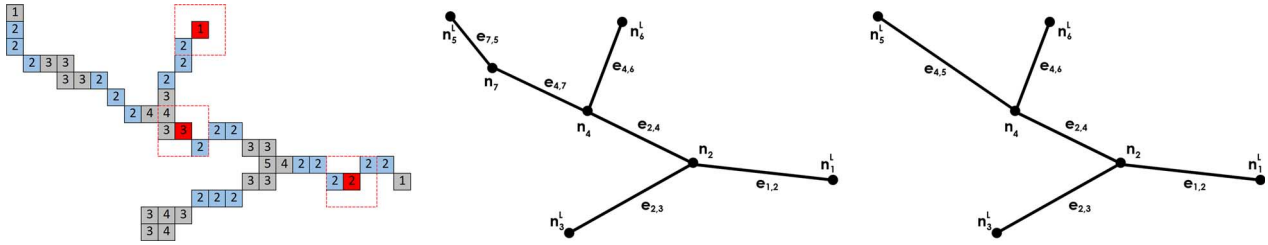

Fig. 2. A schematic 2D example of **(a)** centerline voxels $X = \{x_k\}$, **(b)** the corresponding initial geometric graph G^I, and **(c)** the refined geometric graph G, as described in Section II-B. **(a)** Each square indicates a centerline voxel x_k and includes the number of neighbors $\Omega_v(x_k)$. The voxels colored in blue represent the voxels for which $\Omega_v(x_k) = 2$ and the voxels colored in gray represent voxels for which $\Omega_v(x_k) \neq 2$. The dash red square indicates the neighborhood of the red voxel inside the square. **(b)** The corresponding initial geometric graph G^I as defined by Eq. (1), containing nodes n_i, leaf nodes ${n_i}^L$ and edges $e_{i,j}$. Note that n_i corresponds to a single, or a group of, centerline voxels in **(a)** for which $\Omega_v(x_k) \neq 2$ (colored in gray). **(c)** The refined geometric graph G in which nodes that only have two neighboring nodes (i.e. $\Omega_n(n_i) = 2$) are included in the attached edges.

[8] and the vessel segmentation V is constructed by using the method of [9]. The segmented vessels V are reduced to a centerline representation $X = \{x_k\}$ [10], where x_k represents a centerline voxel. We indicate the number of voxels in the 26-neighborhood of x_k as $\Omega_v(x_k)$. An example of a centerline representation is shown in Fig. 2(a).

B. Geometric Graph Representation

In a vessel segmentation, we are dealing with many different morphologies that are difficult for a centerline algorithm to deal with, e.g. bifurcations, trifurcations, artery-vein attachments, and false positives included in the segmentation (such as diseases). For these structures, the performance of a centerline algorithm is difficult to predict. For example, a bifurcation would ideally be a single voxel connected to three neighboring voxels, but is often represented as a group of multiple connected centerline voxels for which $\Omega_v(x_k) > 2$ (as is illustrated in Fig. 2(a)). Similarly, when a vessel segmentation is locally cluttered by noise or false positives, a centerline algorithm will most likely construct a connected structure of centerline voxels for which $\Omega_v(x_k) > 2$. Vascular branches on the other hand, have a much simpler tubular shape for which the behaviour of a centerline algorithm is more constant and easier to predict. Therefore, based on the centerline voxels and their neighbors, we can make a rough separation between voxels that are likely to belong to a vascular branch $(\Omega_v(x_k) = 2)$ and voxels from which we cannot assume anything $(\Omega_v(x_k) \neq 2)$.

Based on this rough separation, X is redefined as an initial geometric graph $G^I = (N, \mathcal{E})$ consisting of nodes $N = \{n_i\}$ and edges $\mathcal{E} = \{e_{i,j}\}$ as schematically shown in Fig. 2(a) and Fig. 2(b). This initial geometric graph represents both the topology and the spatial distribution of the vessels in V. Nodes n_i and edges $e_{i,j}$ are defined as:

$$\begin{aligned} n_i &= \{x_k : \Omega_v(x_k) \neq 2, x_k \text{ are connected}\} \\ e_{i,j} &= \{x_k : \Omega_v(x_k) = 2, x_k \text{ are connected}\}, \end{aligned} \tag{1}$$

where connectivity is defined within the 26-neighborhood and $e_{i,j}$ is the edge connecting n_i to n_j. Note that an edge only represents voxels for which $\Omega_v(x_k) = 2$, meaning that edges are expected to represent the simple tubular structures. As no assumptions are made regarding the underlying structures of the nodes, nodes can represent bifurcations, trifurcations, attached arteries and veins, end points, or false positive structures. A refined geometric graph G is extracted from G^I in which nodes with a degree $\Omega_n(n_i)$ of 2 (i.e. a node with only two neighboring nodes) are included in the attached edges as schematically shown in Fig. 2(c).

Since the geometric graph G represents both the topology and spatial distribution of the vessels in V, root nodes $n_i^R \in N$ are identified based on the vessel segmentation. In the vessel segmentation, roots are defined as the large vessels entering from the lung hilum into the lung segmentation. These vessels are therefore attached to the lung border L_b and are extracted as $V_b = V \cap L_b$. In the geometric graph, root nodes have per definition a degree of one (i.e. connected to only one neighboring node) and are therefore defined as $N^R = \{n_i^R : \Omega_n(n_i) = 1, L_b \cap n_i \neq \emptyset\}$. Leaf nodes also have a degree of one but do not intersect with the lung border, and are therefore defined as $N^L = \{n_i^L : \Omega_n(n_i) = 1, n_i \notin N^R\}$. Fig. 3(a) shows a schematic example of root and leaf nodes. By defining the root nodes as relative to the lung border (instead of relative to the actual roots at the heart), we restrict our method to the segmented vessels in the given lung segmentation. This is done to avoid vessel segmentation in the lung hilum (such as the segmentation of the pulmonary trunk and the left and right pulmonary veins) which is an extremely difficult task for both computers and medical experts. Note that the artery-vein class of the root nodes remain unknown at this point.

C. Geometric Graph Separation

The extracted geometric graph is a network of attached arterial and venous trees because it contains the artery-vein attachments that are present in V. The nodes in G that are likely to represent these artery-vein attachments are indicated as $N^* = \{{n_i}^*\}$. The goal of geometric graph separation is to construct a set $\mathcal{S}$ of subtrees Ψ_i that do not contain artery-vein attachments $\mathcal{S} = \{\Psi_j : {n_i}^* \notin \Psi_j, \Psi_j \subset G\}$. Based on the definition that a subtree does not contain ${n_i}^*$, a subtree is assumed to be of either the arterial or venous class. In this section we describe how to identify the set of nodes N^* by means of *leaf node pruning*, and how to separate G into $\mathcal{S}$. Geometric graph separation is schematically outlined in Fig. 3.

1) Detection of Artery-Vein Attachments: Leaf node pruning is an algorithm that iteratively removes leaf nodes from a graph

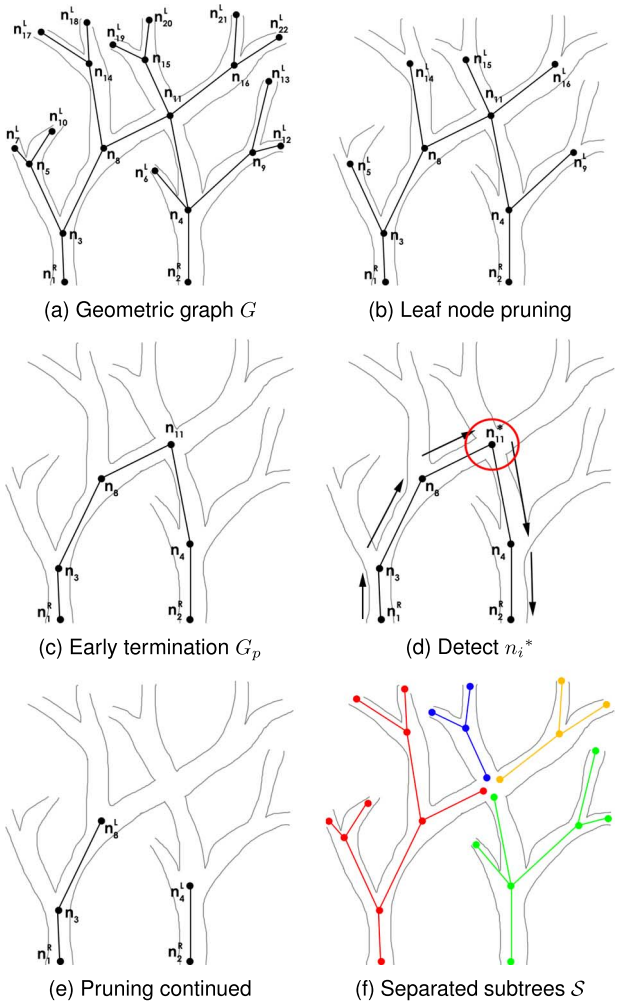

Fig. 3. An illustration of separating a geometric graph G into a set of subtrees $\mathcal{S}$ as outlined in Section II-C. **(a)** Leaf nodes n_i^L and root nodes n_i^R are detected in G. **(b)** and **(c)** Leaf node pruning is performed iteratively until all leaf nodes are removed from the graph, resulting in a pruned graph G_p. **(d)** A path is constructed between root nodes in G_p (illustrated by arrows). By analyzing the smoothness of the transition between two edges at each node in this path, nodes n_i^* are identified that are likely to represent an attached artery and vein. An example of the detected node n_{11}^* is indicated by the red circle. **(e)** By removing n_{11}^* from G_p, new leaf nodes are constructed which enables the pruning process to continue. **(f)** After splitting the detected nodes n_i^* from the original graph G and thereby disconnecting all vessels at those locations, G has been partitioned into a set of multiple subtrees $\mathcal{S}$. (a) Geometric graph G; (b) Leaf node pruning; (c) Early termination G_p; (d) Detect $n_i{}^*$; (e) Pruning continued; (f) Separated subtrees $\mathcal{S}$.

as illustrated in Fig. 3(a), Fig. 3(b) and Fig. 3(c). Leaf node pruning terminates when all leaf nodes are iteratively removed from the graph, which typically occurs when the root node of a graph is reached. However, a graph that contains more than one root node cannot be completely pruned as illustrated in Fig. 3(c), resulting in early termination of the pruning process. Early termination therefore provides a pruned graph G_p that contains at least one $n_i{}^*$, where G_p serves as a reduced search space to identify these attachments.

To identify $n_i{}^*$ in G_p, all possible paths between two root nodes $P^{R\rightarrow R} = (n_a, n_{a+1}, \ldots, n_{b-1}, n_b)$ in G_p are constructed for which only $\{n_a, n_b\} \in N^R$, as shown in Fig. 3(d). The direction of the edge between two subsequent nodes n_i and n_{i+1} in a path is denoted by $\vec{e}_{i,i+1}$. Given that the geometric graph represents the spatial distribution of the vessels, a smooth transition between subsequent edges of the same artery-vein class is expected. Smoothness of the transition at node n_i is expressed by the parameter $\beta_i = \vec{e}_{i-1,i} \cdot \vec{e}_{i,i+1}$, where $\beta = 1$ represents a transition angle of 0° and $\beta = -1$ a transition angle of 180°. Nodes n_i that correspond to a sharp transition are identified as a potential location of artery-vein attachment, where sharp transitions are defined by a minimum smoothness threshold α. Given the large anatomical variations among patients, α must be estimated per-patient, rather than using a fixed threshold. Therefore, we consider the distribution of transition angles in normal vascular branchings ($n_i \notin G_p$) to estimate a patient-specific threshold α. For each $n_i \notin G_p$, the transition angle between two directly connected edges is calculated with respect to the pruning direction. Histogram analysis is performed on this set of transition angles, defining α as the mth-percentile of the histogram. All nodes in $P^{R\rightarrow R}$ for which $\beta_i < \alpha$, are potential locations of artery-vein attachment and are therefore indicated as $n_i^* \in N^*$. By removing the detected nodes n_i^* from G_p, new leaf nodes are created and pruning is continued, as shown in Fig. 3(e). Note that indicating normal branching nodes as n_i^*, and consequently removing them from G_p, is not an issue for the proposed method since class linking (Section II-D-2) is able to regroup subtrees of the same class.

2) Subtree Construction: By splitting all the detected nodes N^* in G, a set of multiple smaller graphs $\mathcal{S} = \{\Psi_j\}$ is constructed as illustrated in Fig. 3(f). Each Ψ_j is assumed to contain no nodes $n_i{}^*$ that attach an artery to a vein and is therefore assumed to be a subtree of either the arterial or venous tree.

D. Subtree Class Relationships

The geometric graph has been separated into a set of subtrees $\mathcal{S}$ that each belong to either the arterial or venous class. In this section, subtrees of the same artery-vein class are linked based on subtree class relationships. The concept of *periphery matching* is introduced which establishes a direct class relationship between two subtrees. These relationships are extracted by analyzing the peripheral vessel of a subtree and are used to link subtree of the same artery-vein class.

1) Periphery Matching: The rationale behind periphery matching comes from the anatomical knowledge that arteries and veins meet at the alveolar sacs. Since this is far below the resolution of the CT images, arteries and veins merely approach each other in the vessel segmentation without actually connecting. In terms of the subtrees in $\mathcal{S}$, an arterial leaf node is therefore expected to be spatially close to a venous leaf node, and vice versa. This one-on-one relationship can simply be determined without knowing the actual artery-vein class, by matching a leaf node in a subtree Ψ_i to the spatially closest leaf node of a different subtree Ψ_j. Periphery matching quantifies the inter-subtree class relationship by analyzing each leaf node in a subtree and finding the closest leaf node in another subtree within a maximum Euclidean distance D_{max}. The number of leaf nodes that match between two subtrees Ψ_i and Ψ_j determines the inter-subtree matching strength $\mathrm{IMS}_{i,j}$. A high $\mathrm{IMS}_{i,j}$ indicates that Ψ_i and Ψ_j are likely to belong to a different vascular class. To reduce the possibility of establishing a relationship by chance, a minimum of two

matching vessels between subtrees (i.e. $\text{IMS}_{i,j} \geq 2$) is required for a valid relationship.

2) Class Linking: Periphery matching establishes a direct class relationship between subtrees without knowing the actual artery-vein class. After applying periphery matching to all subtrees in $\mathcal{S}$, indirect class relationships can be deducted from a combination of these direct relationships. For example, consider a set of three subtrees $\mathcal{S} = \{{\Psi_1}^A, {\Psi_2}^B, {\Psi_3}^C\}$, with unknown artery-vein classes A, B, and C. When both ${\Psi_1}^A$ and ${\Psi_2}^B$ have a high IMS with ${\Psi_3}^C$, the direct class relationship dictates that both A and B are of an opposite class compared to C. Combining these direct class relationships with the fact that there can only be two classes (i.e. artery or vein) the following indirect class relationship is implied:

$$A \neq C \wedge B \neq C \Longrightarrow A = B \tag{2}$$

Eq. (2) provides a link between the classes of subtrees ${\Psi_1}^A$ and ${\Psi_2}^B$, thus allowing the construction of a new set of linked subtrees $\mathcal{S}_L = \{{\Psi_1}^A, {\Psi_2}^A, {\Psi_3}^{\bar{A}}\}$, where $\bar{A}$ is the opposite class of A. This logic reasoning is iteratively used in $\mathcal{S}$ to link subtrees of the same class, starting with the subtrees with the highest $\text{IMS}_{i,j}$ (i.e. the subtrees that are most likely to belong to a different artery-vein class). Note that one single subtree in $\mathcal{S}$ will not be interlinked to all other subtrees due to the fact that a periphery match is unlikely to be established across lungs and lobes.

E. Inherent Separation Quality Estimation

Periphery matching provides an automatic separation quality estimation for the subtrees in $\mathcal{S}$, which is based on the idea that every leaf node should find a periphery match in another subtree. If more than one leaf nodes within a subtree Ψ_i remains unmatched, it is likely that these periphery matches are within Ψ_i itself. These subtrees are automatically flagged and may be subjected to further analysis. This information is valuable since it allows for both guiding-user interaction and automatic re-separation of a flagged subtree.

For the purpose of this study, we propose a simple approach to automatically reanalyse a flagged subtree Ψ_i^{flagged}. For this, periphery matching is performed between the previously unmatched leaf nodes in Ψ_i^{flagged}, allowing a periphery match to be established within Ψ_i^{flagged}. The path between two matched leaf nodes in Ψ_i^{flagged} is given by $P^{L \to L} = (n_a, n_{a+1}, \ldots, n_{b-1}, n_b)$, where only $\{n_a, n_b\} \in N^L$. Since the matched leaf nodes are each assumed to be of a different artery-vein class, $P^{L \to L}$ is expected to contain the previously unfound ${n_i}^*$. Since this ${n_i}^*$ was not detected in the earlier subtree separation step, a more sensitive criterion than the smoothness threshold α needs to be employed. For this, a sign change in intensity differences along the edges in $P^{L \to L}$ is a sensitive, but less specific, criterion to detect ${n_i}^*$. This criterion is guaranteed to detect more nodes than the actual artery-vein attachment, which is a design choice since the method is able to handle this by linking subtrees in Section II-D-2. After re-separating the subtrees at ${n_i}^*$, the separated subtrees are included in the subtree grouping step as described in Section II-D-2.

F. Classification

Subtrees in $\mathcal{S}_L$ are now linked by means of their class, both to subtrees of the same class and to subtrees of the opposite class. The final task is to classify these subtrees into arterial or venous. Artery-vein classification is performed by automatically analyzing the linked subtrees in $\mathcal{S}_L$, where the difference in artery-vein volume is used as a classification feature. This feature is based on the physiological knowledge that the blood pressure in arteries is higher than in veins. Since the total blood flow remains equal, the venous tree has to have a higher total volume compared to the total volume of the arterial tree. This can be exploited since $\mathcal{S}_L$ contains not only class information of subtrees that belong to the same class, but also information of subtrees that belong to the opposite class. The total volume of subtrees with the same class in $\mathcal{S}_L$ is extracted from the vessel segmentation and normalized by the total length of the corresponding edges. Normalization using the length is needed since subtrees may vary substantially in length, which otherwise means that shorter subtrees are more likely to have a lower volume. Classification is performed by comparing the total normalized volume of subtrees of the opposite class in $\mathcal{S}_L$, where the class corresponding to the highest volume is classified as vein, and the other class as artery.

The vast majority of subtrees have been classified at this stage, however the resulting artery-vein labeling is not guaranteed to give a set of continuous labeled trees. For example, subtrees with only a few leaf nodes might have insufficient peripheral information to establish reliable inter-subtree matches (i.e. $\text{IMS}_{i,j} < 2$). These subtrees can therefore not be linked to any of the other subtrees, which means that no label is given in the classification stage. To ensure the continuation of vascular labels, each leaf node in G is connected to a root node via a path $P^{L \to R} = (n_a, n_{a+1}, \ldots, n_{b-1}, n_b)$, where only $n_a \in N^L$ and $n_b \in N^R$. These paths are constructed by iteratively tracking from node to node until a root node is reached. At each node the next direction is given by maximizing the branching angle β and minimizing the difference in vascular diameter between the previous and next edge. Since an unlabeled node or edge is likely to belong to more than one path, majority voting of the already classified edges in all these path is used for classification. Note that using class information proximal and distal to an unlabeled edge can only provide a reliable classification since the vast majority of the vessels were already classified.

III. Data

A. Data Description

We evaluated the proposed method in 55 non-contrast volumetric low dose (30 mAs at 120–140 kV) thoracic CT scans taken from the Dutch-Belgium lung cancer screening study (NELSON), which showed emphysema and interstitial lung disease with variable severity. Scans were acquired on a 16 or 64-slice CT scanner (Philips Medical systems, Cleveland, OH), at full inspiration. Axial images were reconstructed to 512 $\times$ 512 matrices with a 1.0 mm thickness at 0.7 mm increment, by means of a soft reconstruction kernel (Philips B). More information on the acquisition of the CT scans can be found in [11]. All 55 CT scans are publicly available through the ANODE

challenge [12], with permission from the organizers to use the data for the purpose of this study. The expert annotations, as described in Section III-B, and the resulting classification of our system can be found at http://arteryvein.grand-challenge.org.

B. Manual Annotations

Since annotating all vessels in each scan is too labor-intensive and redundant, we propose an evaluation based on two sets of manually annotated arteries and veins. The first set consisted of 50 randomly selected vessels that were annotated for all 55 scans. The second set consisted of full annotation of all vessels in a subset of ten randomly selected scans.

1) Randomly Selected Vessels: We randomly selected 50 vessels per scan which were each manually annotated as artery or vein by two human observers independently. These vessels were each presented to the observers as an overlay on the original CT scan. The observers classified each presented vessel as an artery or a vein having the 3D data available and the option to window the CT data according to their preferences. As a third option, the observers could indicate that a reliable vessel annotation was not possible. However, observers were specifically instructed to use this option only in exceptionally difficult cases. A fourth option referred to the fact that a presented vessel was not an actual vessel, which could happen when false positives were included in the vessel segmentation.

The total set of annotations for all 55 patients consisted of 2750 vessels, all read by two observers independently. The set of observer 1 was considered to be the reference standard, since this observer had the most experience in this specific task. Observer 1 annotated 26 vessels as too difficult to classify and 48 as not part of a vessel. The reference set therefore consisted of 2676 annotated vessels. Observer 2 annotated 24 vessels as not part of a vessel and 130 vessels as too difficult to classify. A consensus set was constructed from annotations for which both observer 1 and 2 agreed. From the total set of 2750 annotations, 56 annotations were scored as not part of a vessel, and 143 annotations as too difficult to classify by at least one of the observers. In the remaining set, observers disagreed on 186 vessels resulting in a consensus set of 2365 annotations. A total time of 52 hours was spent by the observers for annotating the vessels.

2) Full Vascular Annotations: In addition to the random selection of vessels, a full evaluation was performed on a randomly selected subset of ten scans. In order to provide a full reference standard for these cases, all vessels in these scans were annotated by observer 1, by interactively correcting the subtrees in $\mathcal{S}_L$. Linked subtrees were consecutively presented to the observer as an overlay on the original CT scan in which the observer first determined the artery-vein class of the majority of vessels. Vessels that were not part of the determined class were manually selected and reclassified. The observer spent a total time of 15.5 hours for fully annotating 10 cases.

IV. Experiments and Results

Geometric graph separation resulted on average in 444 subtrees per scan. Subtree class relationships were used to link these 444 subtrees into an average of 50 sets. An average of 10 subtrees per scan were detected and automatically reanalyzed due to poor separation quality (Section II-E).

Two parameters need to be set for the proposed method. We empirically determined the transition smoothness threshold α (Section II-C) to be the 15th-percentile of the histogram (bin size of 1°), providing a good trade-off between true positive and false positive detections on a separate training set of 10 independent CT scans. This automatically implies that by taking the 15th-percentile, some normal branchings will be misidentified as unnaturally sharp transitions. This is a specific design choice since the method is able to handle this when subtree class linking is performed. This makes misidentification of a normal branching preferred over missing an attached artery and vein. We observed that a typical distance between a periphery match was around 5 mm, based on the segmentations of a separate training set of 10 independent CT scans. This distance may vary between scans since it depends on how far the vessels are still visible in the periphery, and how well the vessel segmentation performed on the peripheral vessels. To incorporate this variations, we set the maximal Euclidean distance D_{max} for a periphery match to 30.0 mm, assuming that a periphery match can not be further apart (Section II-D). Periphery matches found beyond this maximal distance are flagged since it is very unlikely that these vessels go towards the same capillary bed.

To evaluate the performance of the method, three experiments were performed. All results will be presented in terms of mean accuracy, median accuracy, and Cohen's κ, including their 95% confidence interval (CI). These statistics were computed by comparing the class of the annotated vessels against the results of the method, the number of correctly classified vessels was used to compute the accuracies (correctly classified vessels/all annotated vessels) and Cohens κ [13](IBM SPSS Statistics, version 20).

A. Experiment 1

In the first experiment the performance of the proposed automatic method was evaluated on all 55 cases by using the annotations of randomly selected vessels as described in Section III-B-1. An overview of these results is presented in Fig. 4 and Table I. The automatic system was first compared to the reference standard. This resulted in a mean accuracy of 88%, a median accuracy of 89%, and a Cohen's κ of 0.76. By considering the arteries as positive class and the veins as negative class, the mean sensitivity and specificity were 88% and 89%, respectively. A few examples of classified artery-vein attachments are presented in Fig. 5. The inter-observer agreement between the reference standard and the second observer yielded a mean accuracy of 93%, a median accuracy of 94%, and a Cohen's κ of 0.84. In addition, the automatic system was compared to the consensus set, which resulted in a mean accuracy of 89%, a median accuracy of 89%, and a Cohen's κ of 0.77. The mean sensitivity and specificity for this experiment were 89% and 90%, respectively. All confidence intervals are reported in Table I.

B. Experiment 2

In the second experiment, the performance of the automatic method was assessed on the scans with a fully annotated vasculature (Section III-B-2). The performance of the automatic method compared to these manual annotations yielded a mean

TABLE I
AN OVERVIEW OF THE RESULTS OF EXPERIMENT 1 AND 3. THE RESULTS ARE PRESENTED AS THE MEAN ACCURACY, MEDIAN ACCURACY AND COHEN'S κ. THE NUMBER OF MANUAL INTERACTIONS INDICATE THE AVERAGE NUMBER OF HUMAN INTERACTIONS THAT WERE NEEDED TO PERFORM THE MANUAL ARTERY-VEIN CLASSIFICATION APPROACH

	Mean accuracy (%) [95%-CI]	Median accuracy (%) [95%-CI]	Cohen's κ [95%-CI]	Manual interactions
Automatic method vs. Reference Standard	88 [86,90]	89 [87,91]	0.76 [0.73,0.79]	-
Automatic method vs. Consensus Set	89 [87,91]	89 [87,93]	0.77 [0.74,0.80]	-
Inter-observer agreement	93 [91,94]	94 [91,96]	0.84 [0.82,0.86]	-
Manual subtree classification method vs. Reference Standard	90 [89,92]	92 [89,93]	0.80 [0.77,0.83]	25
Manual subtree classification method vs. Consensus Set	91 [89,92]	92 [89,94]	0.81 [0.78,0.84]	25

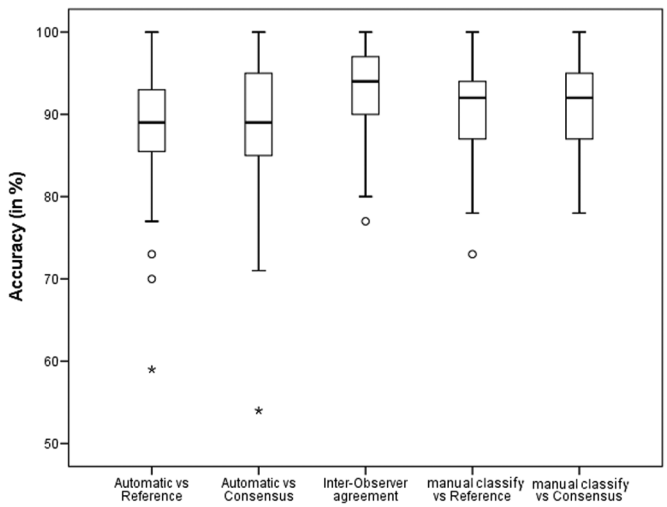

Fig. 4. A graphical overview of the results of experiment 1 and 2. From left to right: The first and second box plot show the results from the evaluation of the automatic method with respectively the reference standard and the consensus set. The case indicated by the star is the only case out of 55 scans that had an accuracy below 70%. In the third box plot, the agreement between the first and second observer is shown. The last two box plots show the results of the method where the subtrees were manually classified, compared with the reference standard (fourth box plot) and the consensus set (fifth box plot). Note that the case resulting in a low accuracy in the first and second box plot (indicated by the star) had an increased accuracy when applying the manual subtree classification.

and median accuracy of 92% (95%-CI [88,95]) and 94% (95%-CI [84,96]), respectively. For two cases, 3D renderings of the full annotations and the results of the automatic method are shown in Fig. 6 and Fig. 7. To compare the two annotation methods (i.e. randomly selected vessel versus all vessels), we evaluated our algorithm using the annotations of the randomly selected vessels on the same subset of ten scans. This resulted in a mean accuracy of 92% (95%-CI [88,95]) and a median accuracy of 93% (95%-CI [84,96]).

C. Experiment 3

The goal of the last experiment was to investigate whether improving the automatic classification step (as explained in Section II-F) could further increase the accuracy of our method. For this purpose, we considered observer 1 to be the best available option for classifying the constructed subtree. Observer 1 therefore provided the best suitable class for the linked subtrees in $\mathcal{S}_L$ in all 55 cases of test set 1, without correcting the constructed subtrees. An average of 25 manual interactions per scan were needed to perform this manual classification. To do this, subtrees of the same class in $\mathcal{S}_L$ were consecutively presented to the observer as an overlay on the original CT scan, and the best suitable artery-vein class was selected. A total time of 7.5 hours was spent for manually labeling the subtrees in all 55 cases. The evaluation was done by comparing these labels to both the reference standard and the consensus set. The results of this experiment are presented in Fig. 4 and Table I. The comparison of this method against the reference standard resulted in a mean accuracy of 90%, a median accuracy of 92%, and a Cohen's κ of 0.80. Comparing this method against the consensus set yielded a mean accuracy of 91%, a median accuracy of 92%, and a Cohen's κ of 0.81.

V. DISCUSSION

We have presented a method to automatically separate and classify pulmonary arteries and veins in non-contrast thoracic CT scans, using a geometric graph separation scheme and anatomical characteristics of the pulmonary vasculature. Since the proposed method uses a binary vessel segmentation as input and not the original CT, the presence of contrast material in the vessels is not used in this method to separate arteries from veins. If the timing of the scan is such that the arteries have a higher attenuation compared to the veins, this could be used as an additional feature to improve the classification. This however, is dependent on the used scan technique and can therefore be a relatively unreliable feature which we did not include in our algorithm. We thus choose to evaluate our system on 55 more challenging non-contrast scans.

The results from the automatic method show that in only five of the 55 scans, the accuracy was lower than 80%, from which only one scan had an accuracy below 70%. Careful inspection of these five cases revealed that the primary reasons for these lower accuracies were errors in the vessel segmentation resulting from a diseased lung. Two examples of scans with false positives in the vessel segmentation are shown in Fig. 8. We observed that mainly false positives in the periphery of the lungs caused problems for the matching procedure, since merging the peripheral vessels of subtrees alters the matching information. Improving the vessel segmentation will be helpful to overcome this issue by reducing these false positives. However, since segmentation errors are often localized to specific parts of the lungs, they do not severely affect the classification in other areas as can be seen from the misclassified vessels in Fig. 6 and Fig. 7.

The results from the method with manual subtree classification (experiment 3) show that only four of the 55 scans resulted in an accuracy below 80%. Interestingly, the case that scored

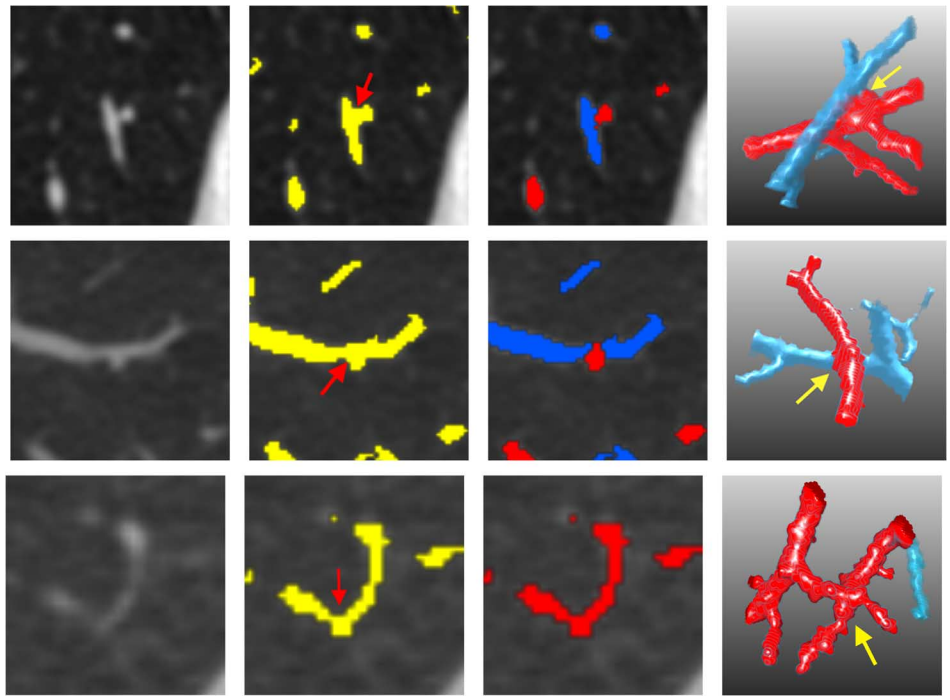

Fig. 5. Three examples of artery-vein attachments and classification results of the automatic method. Each of the rows represent a different artery-vein attachment. In the first frame of each row, an enlarged part of an axial slice of the original scan is presented at a location where an artery and a vein appear to touch each other. The second frame shows the same area overlaying the results of the vessel segmentation, where the red arrows indicate locations where arteries and veins are attached in the vessel segmentation. The third frame shows the results from the proposed automatic artery-vein classification method, where the blue vessels represent the arteries and the red vessels the veins. A 3D rendering of the same results is shown in the fourth frame, indicating that the method correctly separated the artery-vein attachments. The last row shows an example in which the method failed at this location.

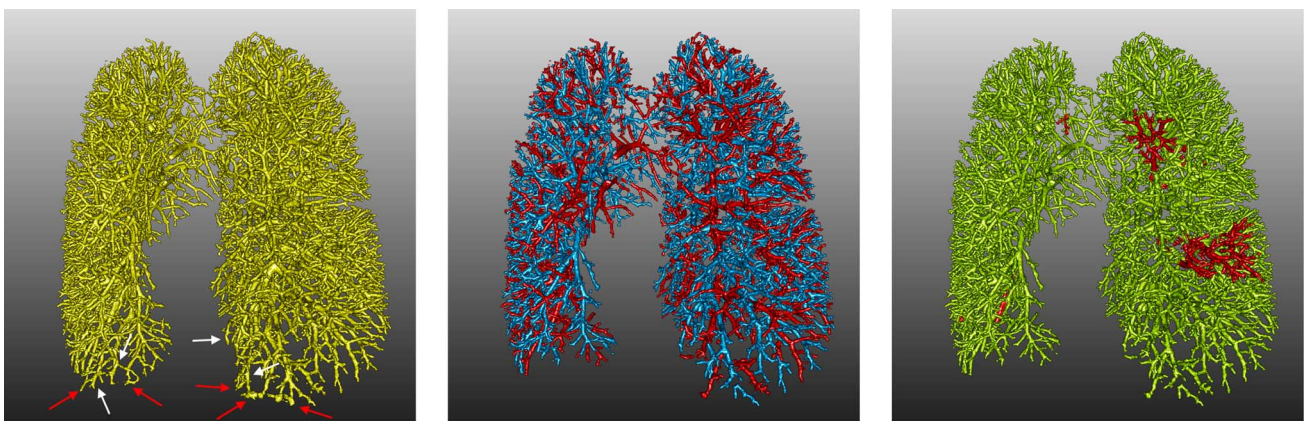

Fig. 6. 3D renderings of example case 1, with in (**a**) the vessel segmentation, (**b**) the result from the proposed automatic method, and in (**c**) the differences (in red) between our result and the full reference standard. The evaluation of this case, as described in experiment 2, yielded an accuracy of 95%. The red arrows in (**a**) indicate some false positives in the vessel segmentation due to diseases in the lungs. These errors merged the ending of vessels together, leaving no information in these subtree to perform periphery matching. The white arrows in (**a**) indicate a few subtrees that are attached to the indicated non-vascular structures. Both the non-vascular structures and their attached subtrees are automatically excluded from the results. Note that this is not affecting the classification of the surrounding vessels.

the lowest accuracy when using the automatic method, now resulted in an accuracy of 84%. When comparing the median performance of this method against the automatic method, an increase in accuracy of 3% is observed. This means that for the used test set, a slight increase in performance can be achieved by improving the automatic classification step as presented in Section II-F. In addition, a diseased lung may affect the difference in total arterial and venous volumes in the lung, making our classification feature less pronounced. Including additional information such as bronchi features (as for example used in [7]) or artery-vein labels of the root vessel in the lung hilum (e.g. identification of the pulmonary trunk), will therefore help to improve the performance and robustness of the automatic classification step. Experiment 3 also suggests that the errors in our method come from the way the subtrees are constructed. For example, determination of the transition smoothness to detect n^* may be inaccurate when calculated between very short edges, which may lead to false positives or false negatives. Improving the detection of artery-vein attachments can therefore also increase the performance of our method.

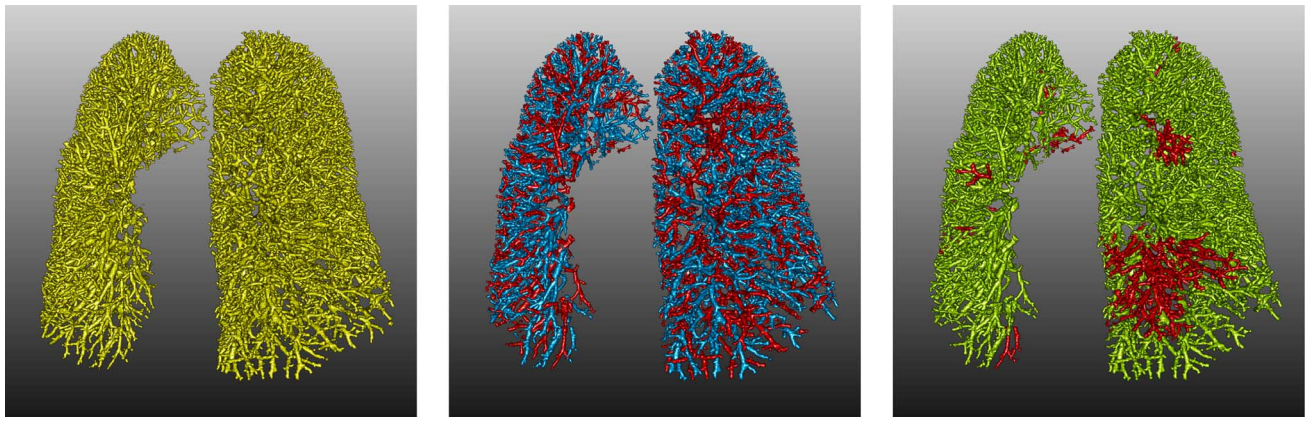

Fig. 7. 3D renderings of example case 2, with in (**a**) the vessel segmentation, (**b**) the result from the proposed automatic method, and in (**c**) the differences (in red) between our result and the full reference standard. The evaluation of this case, as described in experiment 2, yielded an accuracy of 90%.

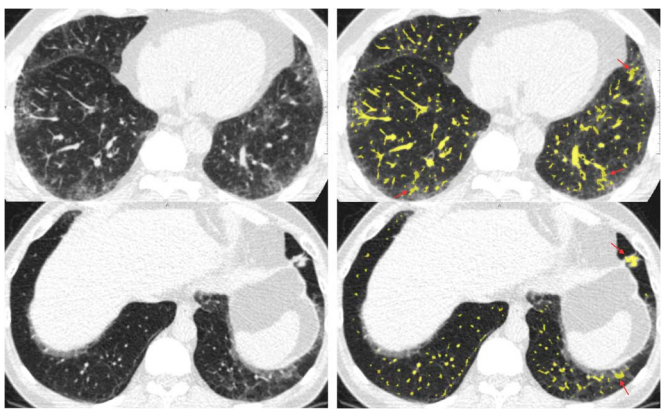

Fig. 8. Examples of some errors in the vessel segmentation in two different cases, where the first frame of each row shows an axial slice of the original CT scan and the second frame shows the vessel segmentation as an overlay on the original scan. The red arrows in the second frame indicate some examples of non-vascular parts in the vessel segmentation.

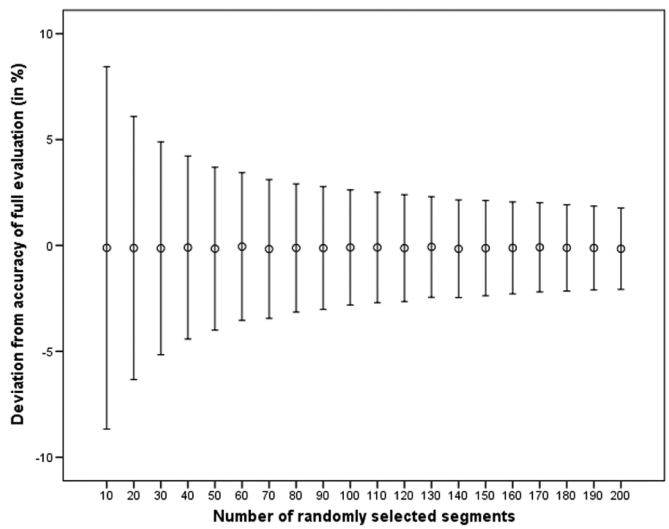

Fig. 9. A simulation experiment to investigate the influence of randomly selecting vessels on the resulting accuracy. For the cases with a full reference (i.e. annotations of all the vessels), we know the true accuracies of our method (experiment 2). In this simulation experiment, $n = 10, 20, 30, \ldots, 200$ vessels were randomly selected from these scans and the artery-vein labels were taken from the full reference. For each n, an approximated accuracy was calculated by comparing the annotation labels to the result of our automatic method. By repeating this simulation a 1000 times for each n and for each of the ten scans, a mean difference and standard deviation was calculated between the approximated accuracies and the true accuracies. Selecting 50 vessels for evaluation provided a good approximation of the true accuracy, with a standard deviation of only 3.8%.

Full annotation of vascular trees is extremely time-consuming and tedious, usually limiting evaluation in previous work [3]–[5], [7]. In experiment 2, we performed an evaluation on a subset of ten scans for which the full reference annotations were available (test set 2). The accuracy that was calculated from each scan with this full reference was considered to be the true accuracy of our method, since each vessel was used for the evaluation. In experiment 2, we additionally performed the evaluation of the same set of scans using the annotations of 50 randomly selected vessels (as in Section III-B-1). This accuracy was considered to be an approximation of the true accuracy. We showed that both the true accuracies and the approximated accuracies were very similar, which indicates that randomly selecting a smaller subset of vessels is a valid evaluation method. To support this statement, we performed an additional experiment in which we randomly select $n = 10, 20, 30, \ldots, 200$ vessels from each scan in test set 2, and simulated the annotation process of these selected vessels using the full reference annotations. For these n vessels, we get the artery-vein labels from the full reference annotations, and compare these labels to the results of our method. This gives an accuracy that approximates the true accuracy of our method in this scan. The difference between this approximated accuracy and the true accuracy, indicates whether selecting n vessels provides a good approximation of the true accuracy. When we repeat this experiment 1000 times, we can calculate the mean difference and standard deviation between the approximated accuracies and true accuracy. Fig. 9 shows the results of this experiment which was performed in each of the ten cases for which the full reference annotations were available. The results indicate that randomly selecting 50 vessels for evaluation is almost equal to selecting more vessels, with a standard deviation of 3.8% for 50 vessels compared to a standard deviation of 1.3% for 200 vessels. This minor decrease in the variability of the accuracy estimate was considered to be not significant enough compared to the increasing annotation time.

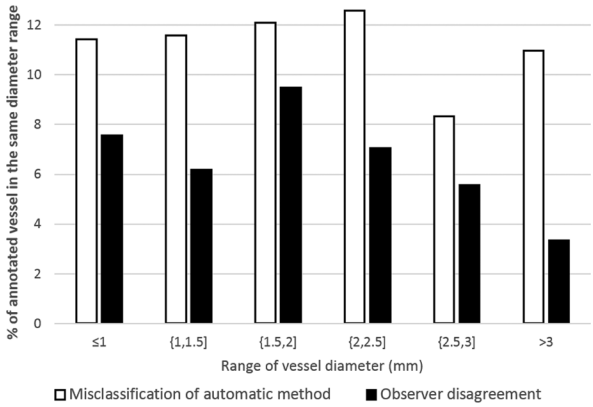

Fig. 10. An additional evaluation on the relation between the misclassifications of our method and the vascular size (open bars), and the relation between observer disagreement and the vascular size (solid bars). The x-axis shows the range of vessel diameters (in mm), whereas the y-axis shows the percentage of vessels in that diameter range that were misclassified by the proposed method, or disagreed upon by the observers. From the open bars, it can be observed that there are only minor differences between the percentages of misclassification per diameter range, which indicates that the performance of our method is size independent. The disagreement between observers seems to become slightly less, suggesting that annotation of larger vessels is easier.

When the consensus set was used for evaluation, no changes were observed in the median accuracies of both the automatic method and the method with manual subtree classification as compared to the evaluation with the reference set. This suggests that the disagreement between the method and observers was different than the disagreement between the observers. A possible explanation is that humans rely on different sources of information. Humans are, for instance, very capable of solving the continuation of vessels at unexpectedly difficult vascular structures such as artery-vein attachments, which is an extremely difficult task for a computer algorithm.

An additional evaluation was done in which we investigated the relation between the performance of our method and the vascular size. Based on the reference set (as defined in Section III-B-1), the mean radius of the correctly classified vessels was 1.29 mm whereas the mean radius of the misclassified vessels was 1.32 mm. This difference was found not to be significant (based on a two-sided Students t-test, $p > 0.05$). In Fig. 10 the percentage of misclassified vessels are shown with respect to their size, where only minor differences between the misclassification per diameter range indicate that the performance of our method is size independent. This is expected since subtrees consist of both smaller and larger vessels, and the proposed method determines only a single class for an entire subtree. We performed a similar analysis for the relation between the vascular size and the observer disagreement, which is also shown in Fig. 10. The mean radius of the vessels for which there was agreement was 1.30 mm and for vessels for which there was disagreement was 1.22 mm. Although this difference was also not significant (based on a two-sided Students t-test, $p > 0.05$), it can be observed from Fig. 10 that there was less disagreement for vessels larger than 3.0 mm.

Despite the high inter-observer agreement, artery-vein classification remains a tedious and time-consuming task for human observers. From the total set of 2750 vessels, 5.2% of the vessels were too difficult to classify for at least one of the observers and the observers disagreed on 6.8% of the presented vessels. Furthermore, in our experiments, an observer needed on average 28 minutes per scan to classify 50 randomly selected vessels. Full manual classification without the aid of an interactive systems would therefore be unfeasible in a clinical setting. The computation time of our method, on the other hand, was on average 6 minutes for classifying the whole vasculature, recorded on a single core with a processor speed of 2.70 GHz. The code was implemented in C++, however it was not optimized in terms of computation time.

The proposed method inherently incorporates the possibility to perform guided user-interaction to improve the performance efficiently in cases where this is needed, as explained in Section II-E. Not only is the task of manually classifying arteries and veins reduced to manually classifying subtrees, leading to on average 25 clicks per scan, but the method also incorporates the possibility to automatically flag potential problematic areas, such as subtrees with numerous unmatched leaf nodes. Instead of an automatic refinement of these subtrees, the location where an artery and a vein are attached could be identified by a single click of the user. In the study by [4], between 25 and 40 manual seed points were required for each of the artery-vein subtrees for both the left and right lung. The total number of manual clicks per scan is therefore substantially higher than in our proposed manual subtree classification method.

Previous studies did not profoundly evaluate their systems on clinical data. Only the studies of [3] and [4] provided a quantitative evaluation of merely two scans. In the study by [4], the authors represented their results as a function of the amount of manual interactions needed to achieve the artery-vein classification. From a full evaluation of the two cases, an overall accuracy of 95% was achieved with 27 manual interactions. The study presented in [3] resulted in an accuracy of 91% for case one and 98% for case two. In addition to the quantitative analysis, they provided a visual scoring scheme to evaluate eight more scans, which resulted in a mean accuracy of 91.8%. Note that the methods in both [3] and [4] performed the artery-vein classification in a semi-automatic way. Although the results of both studies are comparable to the results of the full evaluation of our automatic method (median accuracy of 94%), a fair comparison can not be made since each method was evaluated on a different data set. To fully appreciate and compare the results of different methods in a fair way, all scans, manual annotations, and classification results that are mentioned in this study were made publicly available for future research (http://arteryvein.grand-challenge.org).

Morphological changes of the vasculature, and especially changes in arterial morphology, is an interesting topic in the COPD community [2]. We therefore performed an additional experiment in which we compared the performance of our method in patients with and without COPD. Pulmonary function tests were available in a subset of 25 patients, for which a GOLD stage was calculated. Seven patients were diagnosed with COPD ($\mathrm{GOLD} > 0$) and the remaining eighteen patients were diagnosed as not having COPD ($\mathrm{GOLD} = 0$). The evaluation of our method for the COPD group resulted in a mean

accuracy of 88.57% which did not significantly differ from the no-COPD group with a mean accuracy 88.56% (tested with a two-sided Students t-test, $p > 0.05$). This suggests that our method could potentially be used to investigate morphological changes in the arteries and veins of patients with COPD.

In the proposed method, two intuitive anatomical characteristics of the pulmonary vasculature are extracted, which provide a large amount of information to perform the automatic separation and classification. For example, by using the leaf nodes of all the vessels in a scan, a large amount of information (around 6000 leaf nodes per scan) is available to perform the periphery matching. This is in contrast to the previously proposed bronchial tree feature [7], which is unable to provide information in the more distal parts of the lung since the airways are too small to be distinguishable from lung tissue. This is a problem since there are many artery-vein attachments in these distal parts of the vessel segmentation. Furthermore, using the artery-vein volume differences for the classification of the linked subtrees has not been proposed before for pulmonary artery-vein classification. This is because all existing methods either classify individual vessels [7] or perform manual classification [3]–[5], where our algorithm compares entire sets of subtrees. These subtree sets provide enough information to perform the classification reliably, which is not the case if this feature is used to classify individual vessels.

Periphery matching requires leaf node information in order to establish relationships among subtrees. In situations where this leaf node information is not sufficient, periphery matching is unable to produce reliable inter-subtree matches. The second part of the classification step accounts for these unmatched vessels by estimating their class from the already classified vessels. However, subtrees that did not receive a label after classification, remained unlabeled and were excluded from the final segmentation. After careful inspection of the vessel segmentations, we concluded that the majority of the excluded volume was due to false positives in the vessel segmentation. Some examples of false positives are indicated by red arrows in Fig. 8. The main sources of false positives in the vessel segmentation are the presence of lung diseases, motion artifacts, and the inclusion of non-vascular structure, e.g fissures and airways.

A secondary effect of vessel segmentation errors is observed when non-vascular structures are located in the periphery of the lungs (e.g. in diseased lungs), which causes an abrupt ending of vessels and reduces peripheral information for the matching procedure. Subtrees that are attached to these non-vascular structures might in addition be excluded from the classification results. Some examples of excluded vessels that are attached to non-vascular structures are highlighted by the white arrows in Fig. 6. Improving the vessel segmentation will potentially solve the majority of these issues, however correctly extracting the vessels in scans of diseased patients is currently an unsolved problem in the literature [14].

The extraction of vascular roots, as presented in Section II-B provides a set of root nodes which are used as stopping criteria for the leaf node pruning approach. By defining these roots as the large vessels entering from the lung hilum into the lung segmentation, we can use the border of the lung segmentation to identify these vessels. However, in patients suffering from interstitial lung disease, dense reticulation may be included in the vessel segmentation and additionally be attached to the lung border, which may produce false root nodes. Although these issues did not occur in our test set, this could be avoided by adding a hilar region detection step, such as using a relative location in the lung segmentation or defining a maximum distance from the carina, to exclude the construction of false root nodes.

In conclusion, we have shown that it is possible to automatically separate and classify pulmonary arteries and veins in non-contrast CT, with an accuracy comparable to human observers.

REFERENCES

[1] R. Wittenberg *et al.*, "Acute pulmonary embolism: Effect of a computer-assisted detection prototype on diagnosis—An observer study," *Radiology*, vol. 262, pp. 305–313, 2012.

[2] R. S. J. Estépar *et al.*, "Computed tomographic measures of pulmonary vascular morphology in smokers and their clinical implications," *Am. J. Respir. Crit. Care Med.*, vol. 188, pp. 231–239, 2013.

[3] S. Park, S. Min Lee, N. Kim, J. Beom Seo, and H. Shin, "Automatic reconstruction of the arterial and venous trees on volumetric chest CT," *Med. Phys.*, vol. 40, p. 071906, 2013.

[4] P. K. Saha, Z. Gao, S. K. Alford, M. Sonka, and E. A. Hoffman, "Topomorphologic separation of fused isointensity objects via multiscale opening: Separating arteries and veins in 3-D pulmonary CT," *IEEE Trans. Med. Imag.*, vol. 29, no. 3, pp. 840–851, Mar. 2010.

[5] T. Yonekura *et al.*, "Classification algorithm of pulmonary vein and artery based on multi-slice CT image," in *Proc. SPIE Med. Imag.*, 2007, vol. 6514, pp. 65 142E-1–65 142E-8.

[6] D. Lesage, E. D. Angelini, I. Bloch, and G. Funka-Lea, "A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes," *Med. Image Anal.*, vol. 13, pp. 819–845, 2009.

[7] T. Bülow, R. Wiemker, T. Blaffert, C. Lorenz, and S. Renisch, "Automatic extraction of the pulmonary artery tree from multi-slice CT data," in *Proc. SPIE*, 2005, vol. 5746, pp. 730–740.

[8] E. M. van Rikxoort, B. de Hoop, M. A. Viergever, M. Prokop, and B. van Ginneken, "Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection," *Med. Phys.*, vol. 36, pp. 2934–2947, 2009.

[9] E. van Dongen and B. van Ginneken, "Automatic segmentation of pulmonary vasculature in thoracic CT scans with local thresholding and airway wall removal," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2010, pp. 668–671.

[10] K. Palagyi and A. Kuba, "A 3D 6-subiteration thinning algorithm for extracting medial lines," *Pattern Recognit. Lett.*, vol. 19, pp. 613–627, 1998.

[11] C. A. van Iersel *et al.*, "Risk-based selection from the general population in a screening trial: Selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON)," *Int. J. Cancer*, vol. 120, pp. 868–874, 2006.

[12] B. van Ginneken *et al.*, "Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: The ANODE09 study," *Med. Image Anal.*, vol. 14, pp. 707–722, 2010.

[13] J. Cohen, "A coefficient of agreement for nominal scales," *Edu. Psychol. Meas.*, vol. 20, pp. 37–46, 1960.

[14] R. D. Rudyanto *et al.*, "Comparing algorithms for automated vessel segmentation in computed tomography scans of the lung: The VESSEL12 study," *Med. Image Anal.*, vol. 18, pp. 1217–1232, 2014.