Hepatic vessel extraction using a 3D marked point process model and region growing

Olivér Benis[1,2] and Csaba Benedek[1,2]

[1] Institute for Computer Science and Control (SZTAKI), Eötvös Loránd Research Network (ELKH), Budapest, Kende utca 13-17, Hungary
[2] Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Práter utca 50/A Hungary
benis.oliver@sztaki.hu, benedek.csaba@sztaki.hu

Abstract. Locating the blood vessels in the liver is a crucial task for surgery planning. Doing this process manually is time consuming and prone to errors, therefore in the recent years both traditional and deep learning based methods were proposed for automatic vessel extraction. For this purpose, we propose in this paper a new method that is based on a marked point process (MPP) model and region growing algorithm. As a key advantage, the model parameters can be determined fully automatically using a low number of training images annotated with ground truth information. The method was tested on the 3Dircadb dataset and it was compared to other works. Based on our experiments, our technique is showing promising results: while it is easily trainable it also comes near to the state-of-the-art performance.

1 Introduction

One of the main functions of the liver is the detoxification of the blood. To carry out this the organ has a dense vascular network. The hepatic artery and the portal vein are the inputs, the hepatic veins are the outputs of this and their branches interlace the whole organ. Based on the branches the liver can be divided into eight segments and one can be removed without blocking the blood flow of the others. The location of the blood vessels is important for hepatic tumor surgeries to maintain proper circulation after the removal of a part of the organ.

Contrast enhanced computed tomography (CT) images are suitable for locating the hepatic vascular network but manual extraction is time consuming and interoperator variability can impair the result. Published methods aiming to achieve this task automatically include both bottom up and top down segmentation algorithms.

The technique proposed in [1] divides the solution into two parts after an initial noise filtering. The parts are aiming to segment blood vessels with different thickness. Both tasks consist of two steps, first performing only an initial detection and then based on this result the segmentation itself is carried out. For thin vessels, Gaussian filter and region growing are used, the thick ones are

segmented roughly with K-means clustering and this result is refined using active contour algorithm. The method shows that between the different segments of the vascular network there is a high variation in size, which fact implies that the same technique cannot always be applied effectively for all vessels.

In the work of [2] the authors use a Hessian matrix-based filter to enhance vessels. This so called vesselness filter is often used on the vasculature of other organs like lung or the retina but in case of the liver it would also enhance the contour of the organ and the parenchyma, therefore they improved the filter to avoid this phenomenon. On the enhanced images they have used a fuzzy connectedness algorithm to segment the vessels. They were able to improve the algorithm that way that it needs only a single seed which is generated automatically.

Neural network based top down methods are also effective for the task. In [3] a 3D U-Net architecture is applied on preprocessed CT images where the loss function is based on the Dice coefficient. The neural network can be trained with incomplete labels where some blood vessel segments may be missing. The dataset was divided into training and testing samples by hand so that it yields the best result. Another U-Net based method was proposed in [4] which directly uses the Dice coefficient in the loss function. Two important modifications were made here in the architecture. First, with their proposed attention-guided concatenation (AGC) module the method is able to select the more relevant low-level features using the high-level ones during training. Second, to improve the performance they also included in the architecture a multi-scale feature fusion block which is used to capture multiscale data instead of applying convolutions. The method proposed in [5] is also learning-based but instead of a U-Net architecture they implemented a simpler CNN model. The training samples are noisy, rough blood vessel labels, therefore only a small number of manually created labels are required for this method.

Our approach is based on a marked point process (MPP) model [6]. The MPP approach is used in the field of image processing among others for detecting unknown number of objects which have similar shape. A marked object is described by the coordinates of its reference point (e.g. center), and marker values describing various geometric parameters. A potential is calculated based on how well single marked points (MP) fit the 2D/3D image data, and it also depends on geometric interaction constraints such as the relative object positions with respect to the others. During an optimization process this potential is minimized to achieve a better overall fit. MPP-based segmentation is often used for finding independent, similar objects in images. It is effective for example to detect fibers on microscopy images [7] or to locate archeological tumuli on LiDAR recordings [6]. However, with also taking connection constraints between neighboring objects into consideration, it is also possible to extract a network of objects this way. Related examples from the literature include detecting roads in aerial images [8], and extracting blood vessels from retina fundus images [9]. Both of these techniques are used for 2D image data. On the contrary, we propose in this paper an MPP model working directly in the 3D space of the CT measurements.

2 Proposed approach

The segmentation workflow consists of four steps. *First* the CT image is modified to our needs. *Second* the blood vessels are extracted approximately with a 3D tube MPP model. The result of this step is corrected in the *third* step which ensures correct orientation and connections between the tubes. In the *fourth*, post processing step, based on the tubes the end result is refined with a region growing algorithm.

2.1 Preprocessing

The intensity range of the liver voxels in a CT image is much narrower compared to the whole body. For this reason we can select the relevant histogram window, and work only with these voxel intensities. The center of the window is determined by the mean intensity of the liver voxels and the width of it is selected so that the root mean square contrast is the same for every image. Based on the width of the window using 8 bit images is sufficient. The preprocessed images are reduced to half of the original size and stretched based on the slice thickness.

2.2 Marked point process model

The models proposed in [8] and [9] use ellipses and rectangles as representations of tubes in 2D. Instead of using that 2D approach, our model works with 3D cylinders to cover blood vessel segments. The MPP potential function is built up from the data term and the overlap and connection potentials. The optimization is based on the multiple birth and death algorithm [6], which is applied multiple times while decreasing tube radius ranges.

3D tube mark A marked point describing a cylinder object is shown in Figure 1. The position of the object is defined by the three spatial cordinates of the cylinder's center point. In addition, the object has four geometric parameters (i.e. markers): its height, the radius and two angles which determine the orientation in 3D. For the height and radius parameters there is a range of possible values which is based on the properties of the blood vessels on CT images. The space of the markers is the following: $M = [h_{min}, h_{max}] \times [r_{min}, r_{max}] \times [0, \pi] \times [0, \pi]$, here h_{min} is the minimum, h_{max} is the maximum height, r_{min} is the minimum and r_{max} is the maximum radius value.

Potentials The configuration energy function is composed of three main potential components:

$$v_x = \zeta v_{obj_x} + \eta \sum_{\substack{i=0 \\ i \neq x}}^{n} v_{ol_{xi}} + \theta \sum_{\substack{i=0 \\ i \neq x}}^{n} v_{conn_{xi}} \tag{1}$$

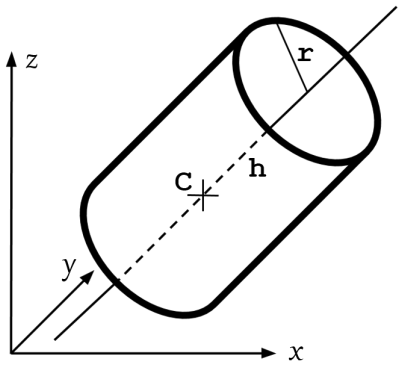

Fig. 1: Description of a cylinder object

where ζ, η and θ are weight parameters.

The first part of the potential is the data term (v_{obj_x}) which is derived from the voxel intensities in and around the cylinders. Instead of the one introduced in [10] and used in [8] and [9] we propose a new data term here. The outside region is enclosed by the object's cylinder surface and an enclosing concentric cylinder which has a 3-voxel larger radius. The end region consists of two cylinders at the bases of the original one which are the elongations of it and are 2-voxel long. We consider three measures here: the average voxel intensities inside the cylinder, outside it and at its bases in the end region.

On the contrast enhanced CT images the blood vessels have higher intensity voxels than their surrounding. The goal is to ensure that the inner average intensity is higher than the mean intensity outside the cylinder, but it should be similar with the average at the bases because both regions represent blood vessel parts. Based on these soft constraints, the data term can be described as follows:

$$v_{obj} = \frac{\alpha|mean_{in} - mean_{end}| + \beta|maxIntens - mean_{in}|}{\gamma(mean_{in} - mean_{out})^2} \tag{2}$$

where $mean_{in}$, $mean_{out}$ and $mean_{end}$ are the inside, outside and end mean intensities respectively, α, β and γ are weights and $maxIntens$ is the largest possible voxel intensity which is 255 for 8 bit images.

The overlap between the cylinders should be avoided, which constraint can be ensured with the overlap potential v_{ol}. This term is calculated as in [8] from the overlap ratio of the cylinders and if it is over a threshold, the marked point will be deleted with $p = 1$ probability:

$$R = \frac{V_{tube_i} \cap V_{tube_j}}{min(V_{tube_i}, V_{tube_j})}$$

$$v_{ol} = \begin{cases} R, if R < T_{ol} \\ \infty, otherwise \end{cases} \tag{3}$$

To cover the whole region of the blood vessels with the cylinders we need many of them connected together at their ends. This way the segments can follow the curvatures and branches. The connections are promoted with the connection potential. With extending the connection circles in [8] to 3D we create spheres at both ends of each cylinder. The overlap between spheres of two cylinders means that they are connected, and the higher the ratio of overlap the stronger the connection. For a single marked point the connection potential contains the overlap ratios of both of its spheres with all the other marked points. Since the connected tubes should have similar thickness there is a term included in the potential v_{conn} which penalizes large radius differences:

$$R_{ij} = \frac{V_{sph_i} \cap V_{sph_j}}{min(V_{sph_i}, V_{sph_j})}$$

$$v_{conn} = -\delta(\sqrt[3]{R_{Fij}} + \sqrt[3]{R_{Bik}}) + \varepsilon(|r_i - r_j| + |r_i - r_k|) \tag{4}$$

Here δ and ε are the weights of the two terms.

2.3 Parameter tuning

The weights of the different potential terms influence the result greatly because they determine the importance of the different energy components relatively to each other. The values of the weights could be determined by trial and error but this choice can be time consuming and it is easy to miss the best setting for our purposes.

The data term parameters, including the α, β and γ coefficients and the ζ weight parameter are determined automatically. Since the remaining parameters of the model do not depend on the image data, they are selected beforehand and are fixed.

For the data term's parameter tuning, five CT volumes were used from the 3Dircadb dataset along with their ground truth data. First, cylinder objects are generated randomly. Because of the high variation in the size of the blood vessels the weights should be different for larger and smaller vessels, respectively. This requirement is fulfilled by a multilevel operation, which we discuss later. The tuning is done for each level separately, which yields that the accepted size range of the generated cylinders are based on the current level.

The Sørensen–Dice coefficient (SDC) is calculated for each generated marked point but since with this measure the whole vascular network cannot be compared to a single cylinder, only a part of the ground truth data derived from the neighborhood of the cylinder is considered. However, uncontrolled MP generation may result in almost entirely low precision cylinders, spoiling the training process that may result in poor weights. To avoid this situation the possible

range $[0, 1]$ is divided into five equal subranges and 25 marked points are selected with Dice coefficients in each of these ranges. This way a more uniform distribution of SDC values can be achieved.

For one level, 625 marked points are generated alltogether, which are used to determine the weights that give the best result. The $mean_{in}$, $mean_{out}$ and $mean_{end}$ values are calculated and the possible weight distributions are searched exhaustively: with every set of weights the object potentials are calculated and analysed. The goal is to ensure that the possible largest number of cylinders with high SDC values have low potentials, while at the same time the ones with lower coefficients have large potentials. SDC over 0.8 was considered high here.

2.4 Optimization

Finding the correct configuration of cylinder objects is an iterative process. In every iteration first a birth and a death step are executed to create points stochastically and remove the ones which cannot be accepted based on their potentials. Next, in the perturbation step the parameters of the marks are changed to achieve a better fit.

Birth and death step The creation and deletion of the marked points is achieved according to the multiple birth and death algorithm [6].

At the birth step a marked point can be created with a predefined probability at every voxel. The parameters of the mark are selected randomly.

After the birth step all of the earlier generated marked points and also the ones which were created in the current iteration are checked. A death rate is calculated from their potential v and they are deleted or left alive according to this value:

$$deathrate = \frac{1}{1 + \kappa e^{-\lambda v}} \tag{5}$$

The κ parameter is decreased, λ is increased in each iteration. This step is also stochastic which means that the death rate is compared to a random value and this comparison determines whether the marked point has to be deleted.

Perturbation step After the birth and death steps there is a perturbation step in every iteration. In this process the parameters of all existing objects are modified randomly. Based on the difference between the object's original potential and the one derived using the newly generated parameters, the modification can be accepted or rejected. If the potential decrease is high enough the probability for accepting the change will also be higher.

2.5 Multilevel operation

As it is mentioned earlier there is a range of possible height and radius values. We have observed that if marked points with any radius in this range can be

generated the smaller ones have a tendency to occupy the place where a larger should be.

This problem can be avoided if not all the possible marked points are created at the same time, but the extraction works on multiple levels. First only larger objects can be generated and the smaller ones are born only on later levels. If the iterations on the current level stop, the marked points are added permanently to the result. On a higher level the lower level objects cannot be changed and cannot be deleted anymore but they are considered when the potentials are calculated. This way the overlap can be avoided and the connections with them are promoted.

The multilevel operation also gives us the possibility that the parameters of the algorithm can be different for each level which can be advantageous.

2.6 Tube refinement

Although the connections are promoted during the MPP optimization there are usually cylinders which cover the vessels precisely but they are not connected to each other. These gaps can be eliminated with slightly increasing the length of the segments, to check if there is another cylinder close by.

After that the segments which have connections at each end are rotated to maximize the overlap between the end spheres. At most 30° rotation is possible. The orientation is considered the best and will be the new one where the sum of the overlaps of the spheres will be the highest. It is also enforced that existing connections cannot be broken up.

2.7 Region growing based on the MPP result

In some cases, for example at curvatures, the cylinders cannot follow the blood vessels correctly and longer ones can have parts which do not cover vessel regions even if the overall fit is good for the MP. This phenomenon implies that smaller units would be better instead of the cylinders, which is possible with using 3D region growing algorithm with seed points based on the tubes.

First the algorithm selects a seed point for each cross section circle of the cylinders. This process is done by finding the maximum intensity voxel in a small region on the plane of the circle. This region is enclosed by the circle which is concentric with the original and its radius is twice as big. A threshold is applied for these voxels, region growing will not be started from low intensity ones. The region growing is performed for every seed point only inside a sphere which has a radius 1.5 times larger than the one of the cross section circle. With executing the algorithm in 3D instead of just in the cross section plane we can avoid leaving out voxels because of rounding errors of the coordinate values.

The intensity threshold for the region growing algorithm is different for tubes in each radius range of the MPP model and was determined with the 5 CT volumes which were also used for parameter tuning.

The maximum intensity voxels used as seed points have more similar directions with the blood vessels than the cylinders had. With fitting a line onto

multiple consecutive seed points we can get the new direction vector. This vector is used to find nearby regions at disconnections. If in this direction other positive voxels can be found at a distance of at most 10 voxels, the gap will be filled between them with region growing.

3 Results

The 3Dircadb[1] dataset was used for the qualitative and quantitative evaluation of the method. This benchmark consists of 20 contrast enhanced CT volumes from different patients. The majority of them have liver tumours. The slice thickness ranges from 1 mm to 4 mm. As it was mentioned earlier the dataset contains ground truth segmentation of the blood vessels done by experts and five of the volumes were used for tuning the parameters so the remaining 15 pairs were included in the tests obtainig the presented results.

We have noticed some imperfections of the ground truth data from the 3Dircadb dataset, especially several missing vessel segments, which issue was also discussed in [4] and [3]. However, in some cases it is hard to decide where do the vessels end exactly, and the delineation was done by experts regarding the original database. For this reason we decided not to modify the ground truth this way, only the blood vessels outside the liver were removed.

As the dimensions of the images were modified during preprocessing, the results are resized to the original size before evaluation.

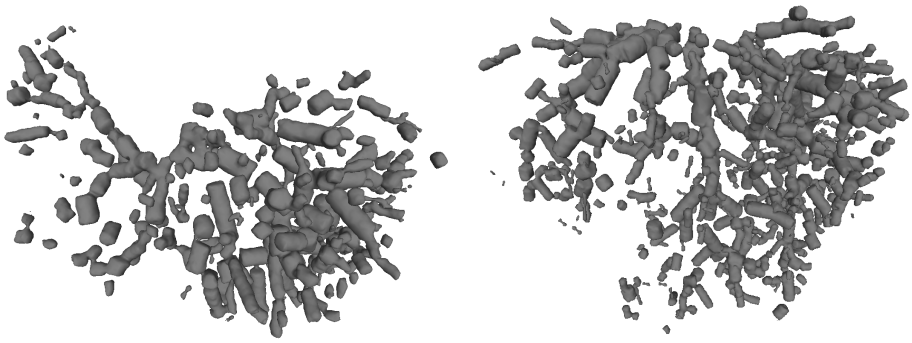

Fig. 2: 3D mesh result of the tube MPP model without refinement

First we have examined the results of the 3D tube MPP model without refinement and region growing. Examples of these can be seen on Figure 2. These initial results are still low precision, as the individual tubes are clearly visible and there are many disconnections between them. It is also visible that many cylinders are very short. This phenomenon is intentional because the gaps

[1] http://www.ircad.fr/research/3dircadb

and the small incorrectly oriented segments can be corrected effectively in the later steps and these shorter tubes can be fitted onto the blood vessels easier.

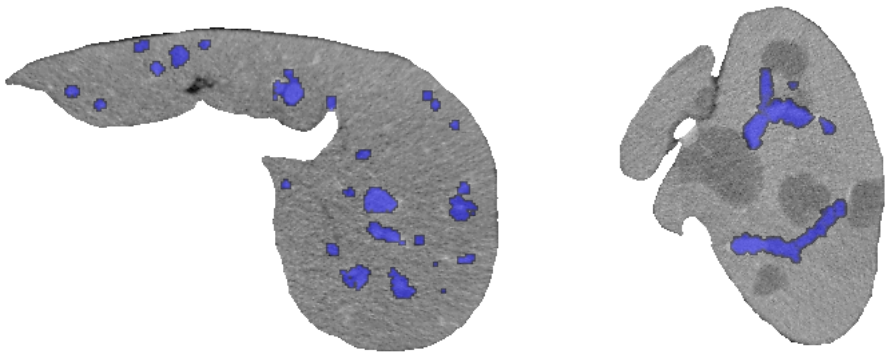

Fig. 3: Two axial slices of CT volumes with the segmentation result (blue)

Inspecting the refined results slice-by-slice, it can be seen that the edges of the blood vessels are correctly identified, and with the refinement short segments and curves are also accurately delineated. Smaller vessels belonging to the last level of the radius ranges with a diameter of not more than 2 voxels are left out the most often. The region growing algorithm causes the merging of close vessels in rare occasions but the parts of the liver parenchyma are almost never segmented together with the vasculature, only when even the MP cylinder is outside the vessel. Tumors inside the liver do not affect the accuracy of the extraction as it is visible on the right side of Figure 3.

To visualize the results we have created three dimensional meshes of the extracted blood vessels (Figure 4). Comparing these meshes with the tube-based result (Figure 2) we can see that the gaps are mostly filled here, and the end result follows the blood vessels much better than the cylinders.

The Sørensen–Dice coefficient, the sensitivity, the specificity and the accuracy were used for quantitative performance evaluation on the 15 CT volumes of the 3Dircadb dataset. The results of our method are given in Table 1.

We compared the result with other approaches discussed in Section 1 where the method was tested also on the 3Dircadb dataset. Although several ones reported slighlty higher Dice rates (0.67-0.69) than our approach, the training and test circumstances were different in each case making direct comparison less relevant. The first technique [2] uses vesselness filter and fuzzy connectedness algorithm, which are considered traditional methods but multiple parameters were set based on the same images which are used for evaluation. The other three considered methods are deep learning techniques. The first one [3] is based on a 3D U-Net architecture. In this work the authors have refined the ground truth (GT) data of the 3Dircadb dataset, but as we have used the original GT, the

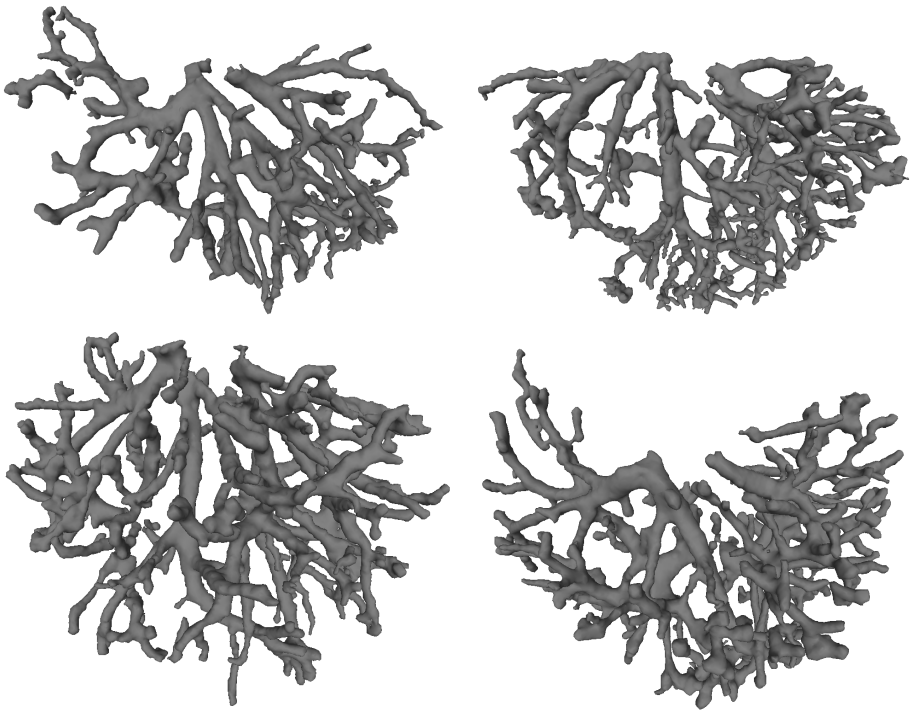

Fig. 4: 3D meshes of results on the 3Dircadb dataset

comparison was done on this non-refined data. [5] utilized a convolutional neural network and performed somewhat better than [2] and [3]. The last method [4], which is also U-Net based, achieved the best quantitative results by far reporting a Dice score around 0.9. However we found multiple concerns regarding the presentation of the results: for example the method proposed in [2] was included here as reference technique with significantly better results than in the original paper. This observation suggests that either the ground truth data was modified or only a small subset of the CT images were used for testing. Also, three quarters of the samples were used for training, and only 25% for the tests, which provides very limited information about the generalization ability of the network. The specificity of our method is higher than the one in [2] and this is also true if we compare the average of the best 10 out of the 15 results with [3] and which measure similar to the result of [5].

As shown in Table 1, between the sensitivity of the best and worst results by our approach there is a large difference. In this case only the average of the 5 best results can surpass [2] and [3]. The lower sensitivity values also decrease the SDC and the accuracy but with the improvement of the worst performing CT images the traditional and the first two deep learning method can be surpassed in all four measures.

	SDC	Sensitivity	Specificity	Accuracy
Minimum	0.506	0.453	0.947	0.922
Maximum	0.682	0.825	0.996	0.970
Average	**0.602**	**0.627**	**0.976**	**0.953**

Table 1: Results on the 3Dircadb dataset

We have also tested the proposed method on the Visceral[2] and Sliver07[3] datasets but since ground truth data for blood vessels are not included in these benchmarks, quantitative measurements could not be done here. With these datasets we could test how well the algorithm performs with parameters based on CT volumes from the 3Dircadb on other databases. Two example results from these datasets can be seen in Figure 5.

4 Conclusion

In this paper, a 3D tube MPP model and region growing based method for hepatic blood vessel extraction was proposed. With the MPP model a rough segmentation result could be obtained which was used for selecting the seeds and size limits, so that the region growing algorithm can accurately extract the vasculature. In comparison with the state-of-the-art methods the performance

[2] http://www.visceral.eu/closed-benchmarks/anatomy3/
[3] http://www.sliver07.org/

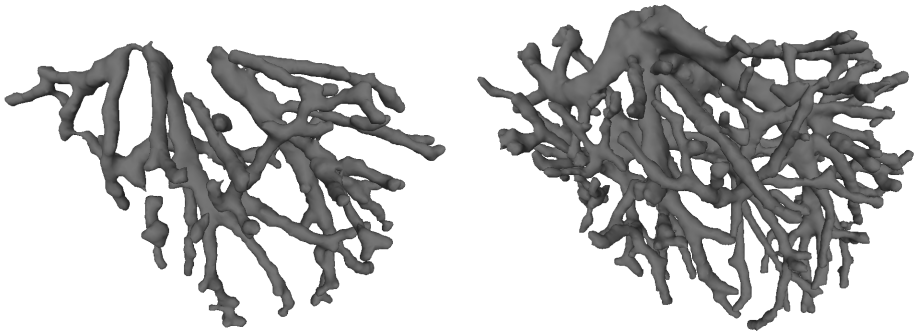

Fig. 5: 3D mesh result from the Visceral (left) and the Sliver07 (right) dataset

of our method cannot reach them in all measures but it has some advantages: automatic parameter tuning is possible independently of the test samples and unlike in case of neural networks the parameters can be modified easily by hand if needed. In future work, we expect to increase the quality of our results, especially the sensitivity and this way reaching the state-of-the-art results will be possible.

Acknowledgements

The research was supported by the Ministry of Innovation and Technology NRDI Office within the framework of the Artificial Intelligence National Laboratory (MILAB) Program, and by the NRDI OTKA grant number K-120233 and by the European Union and the Hungarian Government from the project 'Integrated program for training new generation of researchers in the disciplinary fields of computer science' under grant number EFOP-3.6.3-VEKOP-16-2017-00002.

References

1. Y.-Z. Zeng, S.-H. Liao, P. Tang, Y.-Q. Zhao, M. Liao, Y. Chen, and Y.-X. Liang, "Automatic liver vessel segmentation using 3D region growing and hybrid active contour model," Computers in Biology and Medicine, vol. 97, pp. 63–73, 2018.
2. R. Zhang, Z. Zhou, W. Wu, C.-C. Lin, P.-H. Tsui, and S. Wu, "An Improved Fuzzy Connectedness Method for Automatic Three-Dimensional Liver Vessel Segmentation in CT Images," Journal of Healthcare Engineering, vol. 2018, pp. 1–18, 2018.
3. Q. Huang, J. Sun, H. Ding, X. Wang, and G. Wang, "Robust liver vessel extraction using 3D U-Net with variant dice loss function," Computers in Biology and Medicine, vol. 101, pp. 153–162, 2018.
4. Q. Yan, B. Wang, W. Zhang, C. Luo, W. Xu, Z. Xu, Y. Zhang, Q. Shi, L. Zhang, and Z. You, "An Attention-guided Deep Neural Network with Multi-scale Feature Fusion for Liver Vessel Segmentation," IEEE Journal of Biomedical and Health Informatics, pp. 1–1, 2020.

5. M. Xu, Y. Wang, Y. Chi and X. Hua, "Training Liver Vessel Segmentation Deep Neural Networks on Noisy Labels from Contrast CT Imaging," 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, pp. 1552-1555, 2020.
6. Z. Németh and C. Benedek, "Automatic Tumuli Detection In Lidar Based Digital Elevation Maps," ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLIII-B2-2020, pp. 879–884, 2020.
7. B. Keresztes, O. Lavialle, S. Pop, and M. Borda, "Fiber segmentation on composite materials using marked point processes," Acta Technica Napocensis, 2009.
8. T. Li, M. Comer, and J. Zerubia, "A Connected-Tube MPP Model for Object Detection with Application to Materials and Remotely-Sensed Images," 2018 25th IEEE International Conference on Image Processing (ICIP), 2018.
9. T. Li, M. Comer, and J. Zerubia, "An Unsupervised Retinal Vessel Extraction and Segmentation Method Based On a Tube Marked Point Process Model," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020.
10. X. Descombes, R. Minlos and E. Zhizhina, "Object extraction using a stochastic birth-and-death dynamics in continuum", Journal of Mathematical Imaging and Vision, vol. 33, no. 3, pp. 347-359, 2009.