DEEP LEARNING FOR SEGMENTING CAPILLARY BLOOD VESSELS CONTAINING REMAINING BLOOD CELLS

Grothausmann R., Knudsen L., Ochs M., Mühlfeld C.

Functional and Applied Anatomy, Hannover Medical School

Abstract

The gas exchange in the lung between airspace and blood vessels takes place in the alveolar capillary network (ACN). Characterizing its structure and its change during growth and disease is important to understand the sources of dysfunction in order to develop specific treatments. The various scales in lung tissue make this a challenge and demand a dataset with a large extent as well as a small voxel size.
A series of light microscopy (LM) images were aligned to form a 3D dataset. The segmentation into air, blood and tissue can be achieved with conventional techniques. Remaining erythrocytes represent a major problem concerning the complete segmentation of the blood vessels up to the details of the ACN because they have the same contrast as tissue and can only be recognized by context (at the given resolution).
Their constant volume even during deformation makes this problem an ideal candidate for applying 3D deep learning (DL) to improve the segmentation. The results from preliminary tests with specifically trained 3D U-Nets are promising but also point to general problems when the structures become larger than the receptive field (RF). We therefore investigate adjustments to the network architecture to take segmentation of nearby regions into account in order to preserve consistency between neighbouring regions.

The dataset

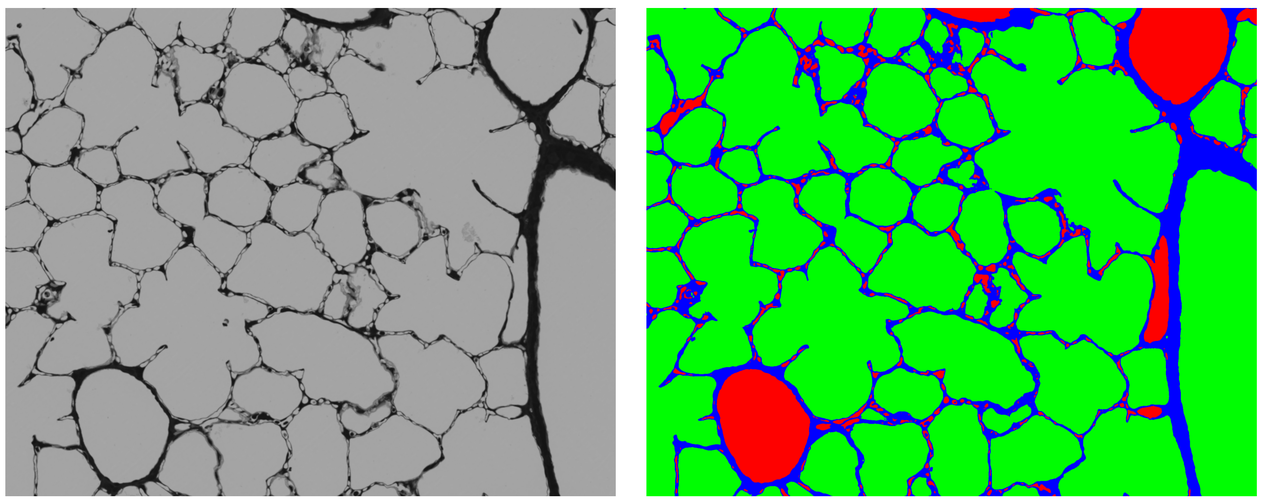

Figure 1: 2D slice from the dataset

Slice of a sub region (2000x1600) containing characteristic structures from the "raw" data and the segmentation achieved with "conventional" image processing, blood: red, air: green, tissue: blue.

Some of the various scales in lung tissue are visible in Fig. 1, 2. For example the gas diffusion distance between air and blood can be as thin as a few 100 nm, the capillary diameter is often oval, partially collapsed and mostly smaller than the diameter of erythrocytes (ca. 6 µm). Alveoli are around 70 µm in diameter whereas conduction airways start around 140 µm, which is similar to the thinnest arteries and veins (ca. 100 µm) while arterioles and venules connecting the ACN are about 30 µm. Arteries generally are adjacent to airways whereas veins tend to be as far away as possible, in our example [2] around 1 mm. Therefore, a suitable dataset must have a voxel size of around 300 nm and an extent of 3000 voxel, resulting in about 30 G voxel.

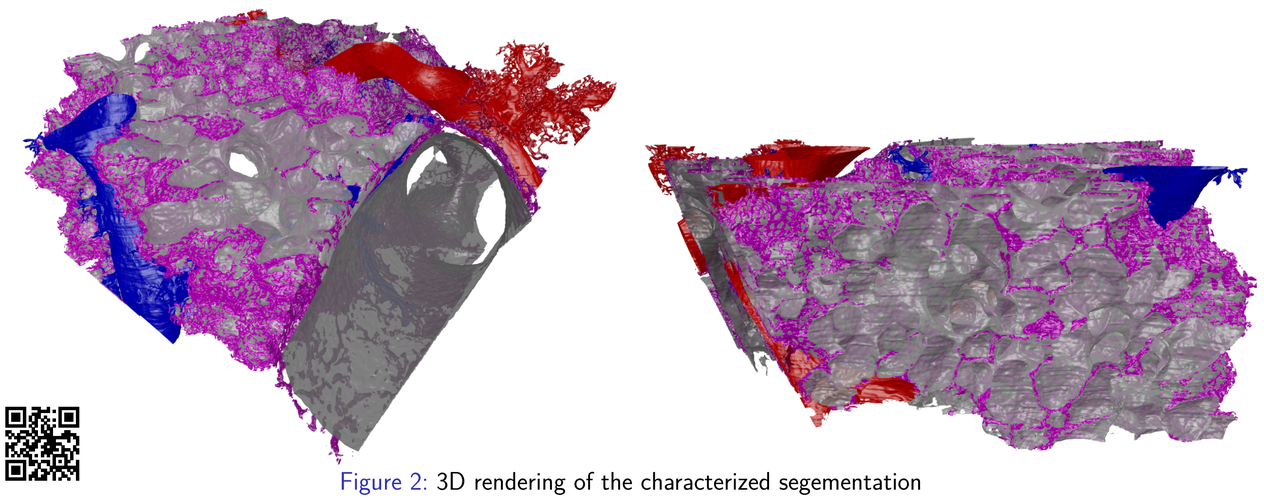

Figure 2: 3D rendering of the characterized segementation

Surface rendered images visualizing the main objects of interest: arteries (red), veins (blue), terminal bronchiolus (Bt, dark grey); largest acinar region contained within the dataset (Aci, light grey); ACN surrounding the two selected air lumina Bt (dark magenta) and Aci (bright magenta) (see also Video at `http://osf.io/gbsns/`).
left: viewed along the bronchiolus and partially along an open duct; right: viewed along widest entry to Aci

Due to the high signal-to-noise ration ($SNR \approx 15$) of light microscopy (LM) images the segmentation into air, blood and tissue can be achieved with conventional techniques combining distance maps, 2D and 3D watershed, fill holes (Reproduction package: `http://github.com/romangrothausmann/h3d`). Apart from a few cutting and staining artefacts (folds, dirt), remaining erythrocytes (Fig. 3) represent a major problem concerning the complete segmentation of the blood vessels up to the details of the ACN because they have the same contrast as tissue and can only be recognized by context (at the given resolution). Since they often block a whole segment of the ACN, their proper segmentation is essential for quantifications such as the Euler number (Euler-Poincaré characteristic, χ) of the ACN. [5]

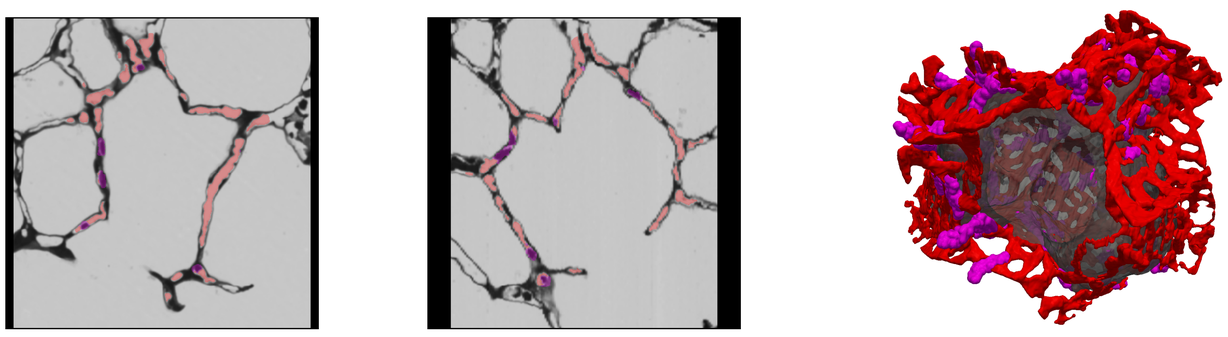

Figure 3: ACN of a single alveolar cluster

Alveolar cluster with surrounding ACN, blood: red, erythrocytes: magenta.

Segmentation with Deep Learning

Even tough erythrocytes deform drastically during the transition through the ACN their volume stays constant. This fact makes the problem an ideal candidate for applying 3D deep learning (DL) to improve the segmentation and ideally do the whole segmentation. About 1000 erythrocytes were segmented by hand in $1/5$ of the whole dataset (0.06 mm^3). The results from preliminary tests with specifically trained 3D U-Nets [1] are promising but also point to general problems when the structures become larger than the receptive field (RF), see Fig. 4, 5 and Tab. 1.

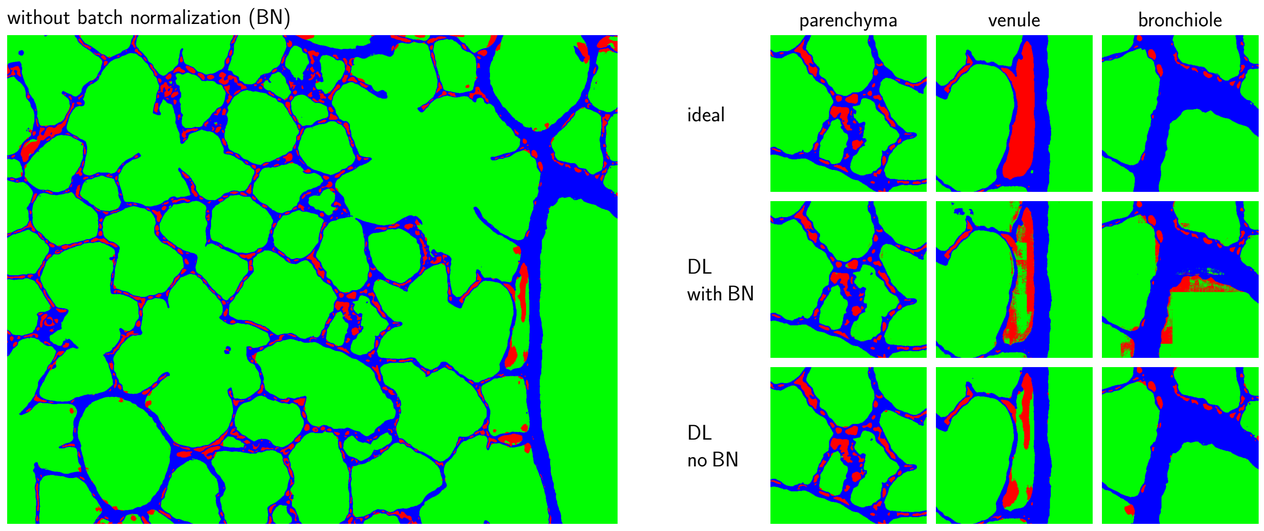

Figure 4: Infer results

left: Slice from the infer result of a 3D U-Net without batch normalization (BN) after 100k training iterations. In comparison to Fig. 1, it becomes obvious that the net is good at segmenting capillaries but has problems with wider blood vessels like arteries and veins that exceed the RF (88x88x88).
right: Three particular regions (400x400). The upper row shows the "ideal" segmentation, the middle row the result with batch normalization (pt, see also Tab.1) and the lower row the results without batch normalization (*pt_noBN*, see also Tab.1). This shows that the net can distinguish blood capillaries from alveolar airspace (parenchyma) but has difficulties with larger blood vessels (venule) and conducting airways (bronchiole).
The sudden changes within square regions originate from the iterative inference when using batch normalization, i.e. tile-by-tile for the maximal 3D U-Net input tile size of 132x132x116 resulting in a segmentation output of 44x44x28. [1]
The size of the receptive field (RF) of the unmodified 3D U-Net architecture (3-stage 3D U-Net as published) is only 88x88x88, which is one possible reason for the discrepancies to the ideal segmentation. The RF can be increased by changes to the architecture (like the V-Net [4] 551x551x551 for binary segmentation) but due to hardware and software limitations (VRAM, 31-bit indexing) this comes with the loss of capability to distinguish multiple classes especially based on context.
Another factor is that conducting airways also resemble tubular structures (in contrast to alveoli) similar to larger blood vessels. This problem could be solved by enforcing consistency to adjacent segmentations, starting the tile-by-tile inference at an "easy" position like parenchyma, where class assignment is obvious.

Name	all	B	A	T	!T
pt/	84	30	90	92	
pt_dS@1µm/	93	67	96	89	
pt_noBN/	84	28	90	91	
pt_PDF@T+1/	84	31	90	89	
pt_T-noT/	97			98	90

Table 1: Experiments and achieved IoU

Measurements in % for "Intersection over Union" (IoU) for all and each individual label (B: blood, A: air, T: tissue) at 100k training iterations. Naming scheme: *pt*: "plain train", *dS*: "down scaled" (from 0.33µm to 1µm voxel size), *noBN*: "no Batch Normalization", *ps*: "pre-segmentation", *ePS*: "empty" pre-segmentation. *PDF* stands for Probability Distribution Function Experiment *dS* demonstrates that scores improve significantly especially for blood if the RF "sees" more structure. Experiment *noBN* avoids tile-artefacts (Fig. 4 row *pt_noBN*) but does not increase the overall result score. Experiment *T-noT* only distinguishes tissue (label T) and non-tissue (!T, consisting of B+A), and hints the general ability to achieve high scores when disregarding distinction between air and blood. *dS* is meant as a preliminary test towards a multi-scale (*ms*) approach, which can be regarded as an indirect way to "increase" the RF by shrinking more structure into the available RF.

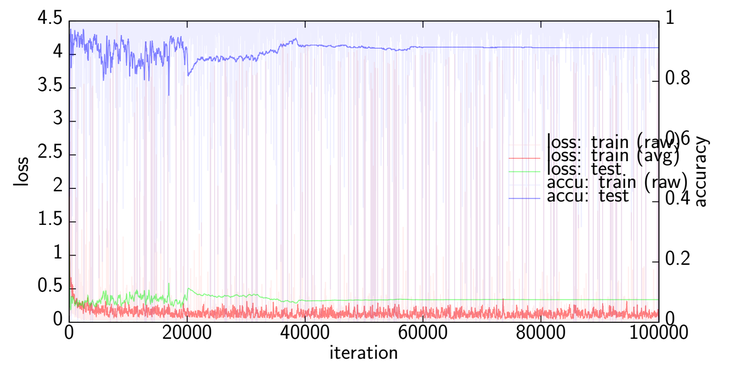

Figure 5: Plot of loss and accuracy for experiment pt

Loss (red, green) and accuracy (blue) plotted for train-phase (red) and test-phase (green) over 100 k iterations.
The training data is generated on the fly (random positions, rotations, deformations, crop size 132x132x116, as described in Ref. [1]), therefore there is no definite notion of epoch. The test dataset of which the actual net-input is drawn from has a size of 2000x1600x573, which corresponds to about the amount of data of 900 net-inputs.
The plots of the test-phase (and averaged train-loss) exhibit far less noise than the raw train-phase plots. This is due to the random cropping positions falling into regions with vast blood vessels or airways, which the network has problems with (as described in Fig. 4). The difference between the loss of test and avg. train (green, red) can be explained by some "very difficult" cases in the validation data set, see Fig. 4.

While blob size limitations of caffe [3] restrict the extension of the RF (88x88x88), limits of VRAM will also make this a problem with other frameworks should the 3D-RF be as large as the whole dataset. We therefore investigate adjustments to the network architecture to take segmentation of nearby regions into account in order to preserve consistency between neighbouring regions when inferring the data tile by tile. In a further step, even letting the net decide where to start and where to proceed. Ideally, the net would infer regions multiple times improving with the data/segmentation seen near by (indirectly extending the RF) with every iteration, visiting difficult regions more often.

References

[1] Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. Ed. by Ourselin, S., Joskowicz, L., Sabuncu, M. R., Unal, G., and Wells, W. Cham: Springer International Publishing, 2016, pp. 424–432. DOI: 10.1007/978-3-319-46723-8_49.

[2] Grothausmann, R., Knudsen, L., Ochs, M., and Mühlfeld, C. "Digital 3D reconstructions using histological serial sections of lung tissue including the alveolar capillary network". In: *American Journal of Physiology - Lung Cellular and Molecular Physiology* 312.2 (2017), pp. L243–L257. DOI: 10.1152/ajplung.00326.2016.

[3] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., and Darrell, T. "Caffe: Convolutional Architecture for Fast Feature Embedding". In: *Proceedings of the 22Nd ACM International Conference on Multimedia*. MM '14. Orlando, Florida, USA: ACM, 2014, pp. 675–678. DOI: 10.1145/2647868.2654889.

[4] Milletari, F., Navab, N., and Ahmadi, S. "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation". In: *Computer Vision and Pattern Recognition*. Vol. abs/1606.04797. June 2016. URL: http://arxiv.org/abs/1606.04797.

[5] Willführ, A., Brandenberger, C., Piatkowski, T., Grothausmann, R., Nyengaard, J. R., Ochs, M., and Mühlfeld, C. "Estimation of the number of alveolar capillaries by the Euler number (Euler-Poincaré characteristic)". In: *American Journal of Physiology - Lung Cellular and Molecular Physiology* 309.11 (Dec. 1, 2015), pp. L1286–L1293. DOI: 10.1152/ajplung.00410.2014.