MixLacune: Segmentation of lacunes of presumed vascular origin

Denis Kutnar[1], Bas H.M. van der Velden[1], Marta Girones Sanguesa[1], Mirjam I. Geerlings[2], J. Matthijs Biesbroek[3], and Hugo J. Kuijf[1]

[1]Image Sciences Institute, UMC Utrecht
[2]Julius Center for Health Sciences and Primary Care, UMC Utrecht
[3]Department of Neurology and Neurosurgery, UMC Utrecht

5 August 2021

1 Introduction

Lacunes of presumed vascular origin are fluid-filled cavities of between 3 15 mm in diameter [1, 2]. They are visualized as a hypointense cavity on Fluid-attenuated inversion recovery (FLAIR) and T1-weighted imaging, usually with a hyperintense rim on FLAIR imaging [1, 3]. Quantification of lacunes relies on manual annotation or semi-automatic / interactive approaches; and almost no automatic methods exist for this task [3, 4]. Initial work by Ghafoorian *et al.* [5] presented a method for detection of lacunes with a deep multi-scale location-aware 3D convolutional neural network (CNN). Preliminary results by Ooms [6] suggest that segmentation of lacunes is feasible with a U-Net [7] CNN.

In this work, we present a two-stage approach to segment lacunes of presumed vascular origin: (1) detection with Mask R-CNN [8] followed by (2) segmentation with a U-Net CNN.

2 Material and methods

2.1 Data

Data originates from Task 3 of the "Where is VALDO?" challenge (`https://valdo.grand-challenge.org/`) and consists of 40 training subjects. For each subject, a T1-weighted, T2-weighted, and FLAIR image are provided; including manual annotations of lacunes made by two different observers.

2.1.1 Lacune prevalence map

Lacunes do not occur randomly throughout the brain, but have preferential locations. They occur mostly in the basal ganglia and the white matter of the internal capsule and pons; and rarely at other locations [9]. To exploit this prior knowledge, a lacune prevalence map was created that shows the expected distribution of lacunes throughout the brain; to be used as a mask on the final results.

From the SMART-MR study [10], we included 98 subjects for which manual segmentation of lacunes were available; and combined this with the 40 training subjects of the VALDO challenge. All images were registered to the MNI-152 [11] atlas to create a lacune prevalence map. The resulting map (Figure 1) was made symmetric by mirroring, dilated by 7 mm in all directions, and the ventricles and CSF were removed.

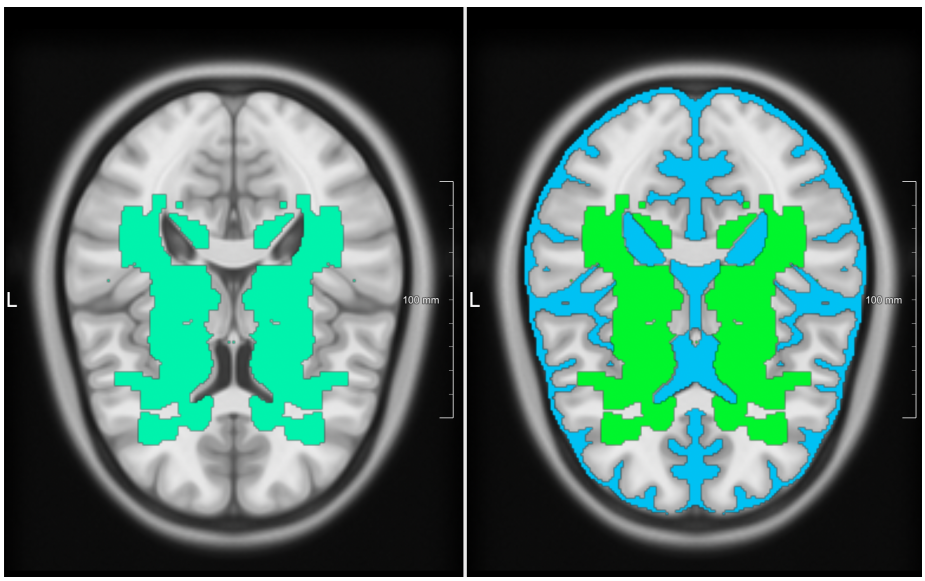

Figure 1: The final lacune prevalence map in MNI152-space is shown on the left. Ventricles and extracerebral CSF were removed from the map with a CSF mask (light-blue on the right).

2.2 Method

Lacunes are hard to find, because of their small size, versatile shape, and sparse occurrence in the brain [1, 3, 4]. For this reason, a two-stage segmentation approach was chosen. In the first step, Mask R-CNN [8] localizes possible lacunes. In the second step, a U-Net [7] segments lacunes based on the locations suggested by the Mask RCNN.

2.2.1 Full prediction pipeline

The full prediction pipeline is shown in Figure 2. The various steps of the pipeline are: z-score intensity normalization, selecting slices and patches for Mask R-CNN, up-sampling the 64×64 patches to 256×256, applying Mask R-CNN, down-sampling back to 64×64 and reconstructing the full 3D image, applying the lacune prevalence map to remove false positive detections, creating 32×32 patches for the remaining detections, segmenting them with the U-Net, and finally reconstructing the final segmentation map. The following sections describe this pipeline in detail.

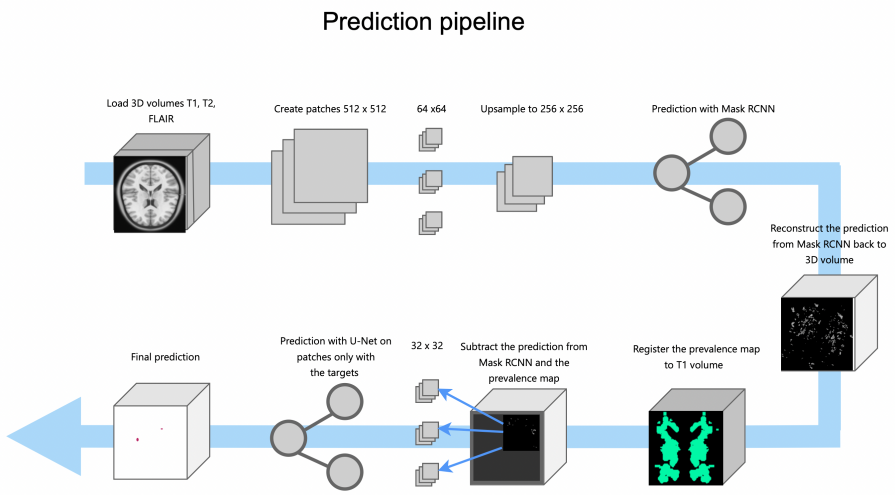

Figure 2: The full prediction pipeline showing the pre-processing steps, the two stages (Mask R-CNN and U-Net) and the application of the lesion prevalence map; to generate the final lacune segmentation for a patient.

2.2.2 Pre-processing

All image data was scaled using z-score normalization (zero-mean-unit-variance) on the patient level.

2.2.3 Stage 1: Mask R-CNN

Mask R-CNN was trained on 2D patches of size 64×64 with 50% overlap, extracted from all slices with lacunes in the dataset. A pretrained Mask R-CNN with a Resnet50 backbone trained on ImageNet was selected. This network is more familiar with objects of medium and large size; and empirical results showed that the current implementation cannot detect the lacunes, because of

their small size. In order to detect lacunes with this network, patches were up-sampled with nearest neighbour interpolation by a factor four to a size of 256×256. This creates object sizes in which the target is larger and more similar to the pre-trained network.

Additional parameters of the Mask R-CNN include the anchor size and aspect ratios. The chosen anchor size ranged from 4 to 64 and the aspect ratios were 0.02, 0.25, 1.0, 2.0, and 2.75. The T1, T2, and FLAIR images were provided as three separate channels to the network with a batch size of six. Data augmentation consisted of random horizontal flips. The model was trained for 20 epochs, after which there was no apparent improvement.

2.2.4 Stage 2: U-Net

A U-Net was employed to segment lacunes at every location detected by Mask R-CNN in stage one. The U-Net was trained on 2D patches of 32×32 with 50% overlap, selected from the locations within the lesion prevalence map in 10/90 ratio for lacune/background (selected empirically on validation data). The model was trained for 30 epochs, and visual inspection showed that after 25 epochs there was no improvement on the training and validation losses (see Results Figure 3).

2.2.5 Applying the lesion prevalence map

The lesion prevalence map was registered to the T1-weighted images using elastix [12], using an affine transformation followed by a bspline transformation. The lesion prevalence map was used to mask out false positive segmentations.

2.2.6 Uncertainty

The VALDO challenge requires an additional uncertainty map to be submitted. Limited time and resources prohibited us from computing the actual epistemic or aleatoric uncertainty. Prior experience suggests that methods are most uncertain at the boundaries of objects. Visual inspection of our results suggest that the final method tends to under-segment lacunes. Therefore we decided to create a 1-pixel in-plane border outside of the segmented lacunes, as a pseudo-uncertainty.

2.3 Evaluation

2.3.1 Stage 1: Mask R-CNN

The goal of the Mask R-CNN was to localize all the lacunes (high sensitivity) with a trade-off of potential low specificity. The model produced two outputs that were taken into account, the bounding-box and segmentation predictions; which were both optimized. To evaluate these predictions, we calculated three measures between the bounding-box predictions and the bounding-box of the

ground truth; and also between the segmentation predictions and the segmentation ground truths. The measures were as follows: Intersection over Union (IoU), Average Precision, and Average Recall. Performance of the Mask R-CNN is defined with these three measures, calculated separately for areas of small, medium, and large size, and all combined. In other words, the model gives a clear evaluation for different areas based on their size. For example, the model with both high Average Precision and Average Recall on small areas, can accurately localize small lacunes.

2.3.2 Stage 2: U-Net

The goal of the U-Net was to segment lacunes in the ROIs given by the Mask R-CNN. To evaluate the U-Net, we calculated the DICE between the segmentations and the ground truth. We optimized the threshold for the posterior probability based on the DICE of the validation set.

3 Results

3.1 Stage 1: Mask R-CNN

The best results of the Mask R-CNN were achieved after 20 epochs, both for the bounding box and segmentation predictions. The highest value for the Average Precision and Average Recall was measured on areas of medium size; both for the bounding box and segmentation predictions. The Average Precision for the medium-size bounding box was 0.02 and for segmentation was 0.06. The Average Recall for the medium-size bounding box was 0.23 and for segmentation 0.27. This suggests that the model might be best in localizing medium-sized lacunes, than small or large lacunes. Visual inspection showed that the majority of lacunes were correctly localized, although at the cost of a large number of FPs.

3.2 Stage 2: U-Net

The U-Net model was trained and evaluated independently from the Mask R-CNN in three different dataset splits. The first case was on a 50/50 split for lacune/background, resulting in a mean DICE on the validation set of 0.87. However, this model predicted many false positives (FPs). For example, the model predicted parts with low-intensity values, like the CSF, to be a lacune. For this reason, more slices without lacunes were included in the training process, with a new split of 25/75 for lacune/background. The given ratio improved the results by having a visual inspection. Despite that still, many FPs remained. Lastly, the dataset with a split of 10/90 for the lacune/background can recognize the lacunes and produced only a minimal number of FPs, which were removed with the prevalence map. We report the mean DICE on the training set of 0.83 and on the validation set of 0.84. The training loss got down to 0.19145 after 30 epochs and the evaluation metrics during training can be seen in Figure 3.

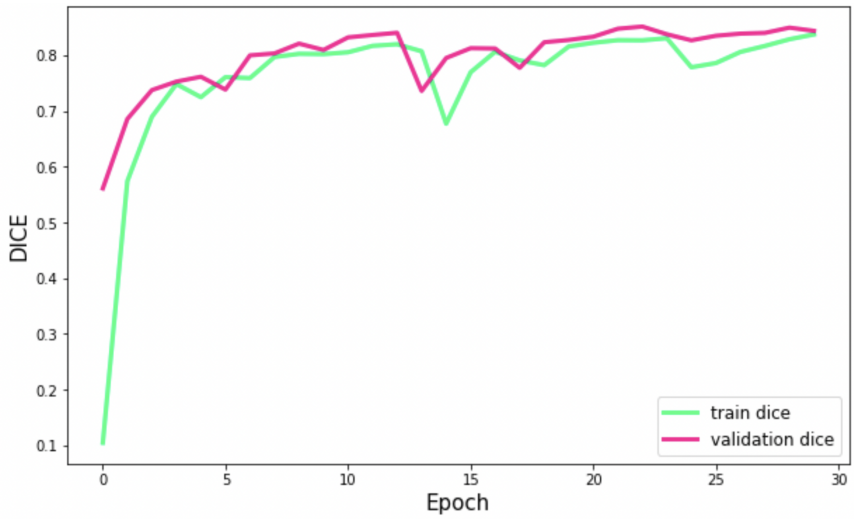

Figure 3: Progression of the Dice metric during training, reported on both the training and validation set.

3.3 Qualitative evaluation

Visual inspection of the results showed that most of the lacunes were correctly detected. An example is shown in Figure 4. However, the predicted binary mask does not precisely match with the ground truth. In some cases, the prediction was a bit smaller compared to its ground truth. A small number of false positive detections remain.

Figure 5 shows the positive and negative effects of the lacune prevalence mask. Most of the times, the lesion prevalence map correctly removes false positive detections; as can be seen in the bottom row of Figure 5. Unfortunately, sometimes the lesion prevalence map also removes a true positive; as can be seen in the top row of Figure 5. This is most likely caused by an image registration error, where the close proximity of the lacune to the ventricles caused a misalignment.

4 Discussion

In this work, we presented a two-stage approach to segment lacunes of presumed vascular origin on brain MR images. The results suggest that this method can accurately detect most of the lacunes on our training/validation dataset. As illustrated in Figure 2, the U-Net decides whether the input from the Mask R-CNN is the true lacune or not and produces the final binary segmentation.

The Mask R-CNN currently produces a substantial number of false positive detections. On the one hand, a high sensitivity is essential in this first stage, probably at the cost of a moderate specificity. On the other hand, a reduction of

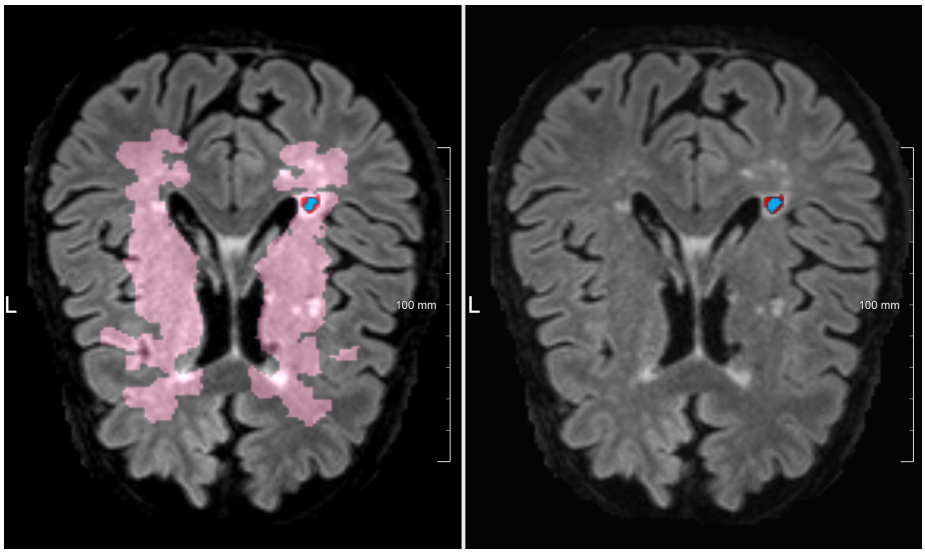

Figure 4: Example of a correctly detected but slightly under-segmented lacune. Pink is the lacune prevalence map, red is the ground truth segmentation, and blue is the predicted segmentation.

the number of false positives in this first stage might help the overall prediction pipeline. Future work could look into thresholding the Mask R-CNN results, to provide a better selection of lacune candidates to stage two.

A limitation of our U-Net model is the use of 2D patches. In some cases, the lacune was correctly segmented on one slice, but missed on the next slice. Using a 3D U-Net might solve this issue, because it can consider the full 3D extent of the lacunes.

The use of a lacune prevalence map seems to improve the method results, by removing a substantial amount of false positive detections; at the cost of also removing some true positives. For future applications, the use of a lacune prevalence map has some limitations. Even though lacunes occur at preferential locations, they can occur throughout the full brain. Using a map to mask out detections, might introduce false negatives at rare locations.

In conclusion, a two-stage deep learning method shows promise for facilitating detection and segmentation lacunes of presumed vascular origin.

5 More information

Source code is available at: `https://github.com/hjkuijf/MixLacune`. The docker container hjkuijf/mixlacune can be pulled from `https://hub.docker.com/r/hjkuijf/mixlacune`.

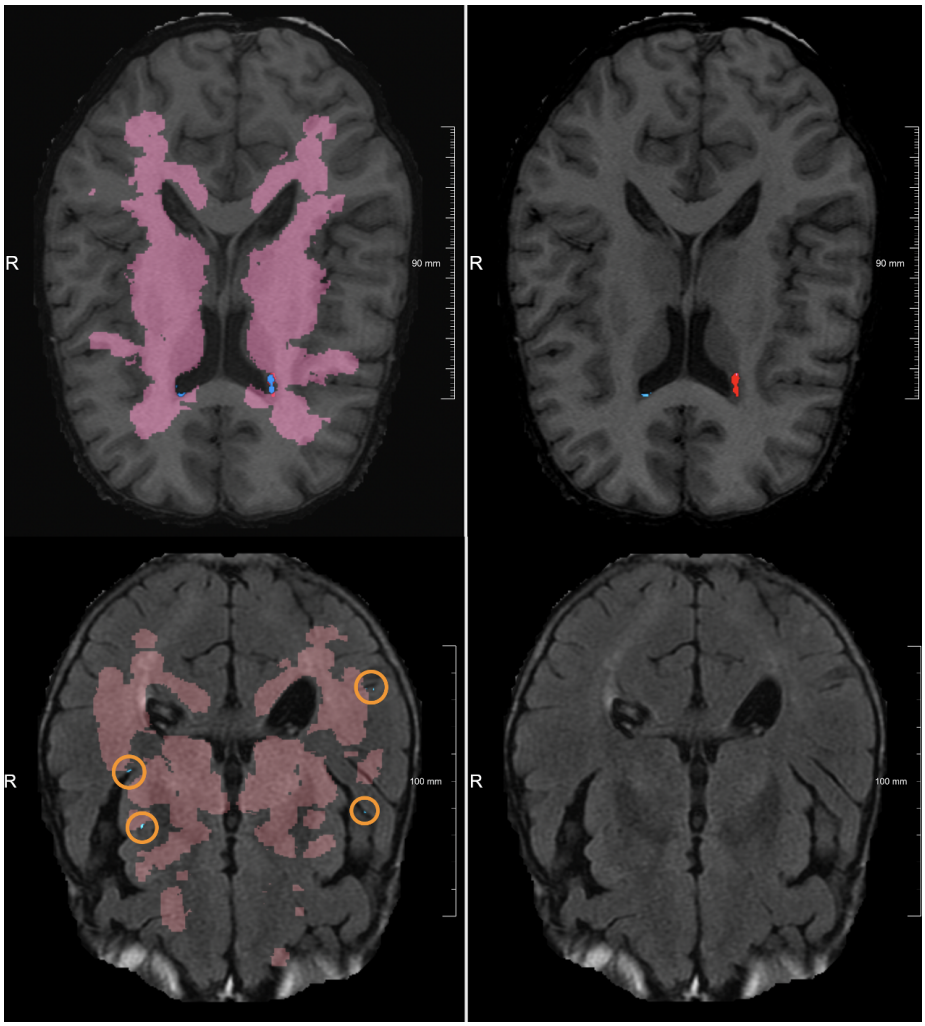

Figure 5: The first pair of images (top row) shows a correctly segmented lacune which was later removed by the prevalence map. The second pair (bottom row) shows how false positives are removed by the prevalence map.

References

[1] J. M. Wardlaw, E. E. Smith, G. J. Biessels, C. Cordonnier, F. Fazekas, R. Frayne, R. I. Lindley, J. T. O'Brien, F. Barkhof, O. R. Benavente, S. E. Black, C. Brayne, M. Breteler, H. Chabriat, C. Decarli, F.-E. de Leeuw, F. Doubal, M. Duering, N. C. Fox, S. Greenberg, V. Hachinski, I. Kilimann, V. Mok, R. van Oostenbrugge, L. Pantoni, O. Speck, B. C. M. Stephan, S. Teipel, A. Viswanathan, D. Werring, C. Chen, C. Smith, M. van Buchem, B. Norrving, P. B. Gorelick, and M. Dichgans, "Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration.," *The Lancet. Neurology*, vol. 12, no. 8, pp. 822–38, 2013.

[2] L. Pantoni, "Cerebral small vessel disease: from pathogenesis and clinical characteristics to therapeutic challenges," *The Lancet Neurology*, vol. 9, no. 7, pp. 689–701, 2010.

[3] F. De Guio, E. Jouvent, G. J. Biessels, S. E. Black, C. Brayne, C. Chen, C. Cordonnier, F. E. De Leeuw, M. Dichgans, F. Doubal, M. Duering, C. Dufouil, E. Duzel, F. Fazekas, V. Hachinski, M. A. Ikram, J. Linn, P. M. Matthews, B. Mazoyer, V. Mok, B. Norrving, J. T. O'Brien, L. Pantoni, S. Ropele, P. Sachdev, R. Schmidt, S. Seshadri, E. E. Smith, L. A. Sposato, B. Stephan, R. H. Swartz, C. Tzourio, M. Van Buchem, A. Van Der Lugt, R. Van Oostenbrugge, M. W. Vernooij, A. Viswanathan, D. Werring, F. Wollenweber, J. M. Wardlaw, and H. Chabriat, "Reproducibility and variability of quantitative magnetic resonance imaging markers in cerebral small vessel disease," *Journal of Cerebral Blood Flow and Metabolism*, vol. 36, no. 8, pp. 1319–1337, 2016.

[4] L. Zhao, A. Lee, Y. H. Fan, V. C. Mok, and L. Shi, "Magnetic resonance imaging manifestations of cerebral small vessel disease: Automated quantification and clinical application," *Chinese Medical Journal*, vol. 134, pp. 151–160, jan 2021.

[5] M. Ghafoorian, N. Karssemeijer, T. Heskes, I. W. M. van Uden, C. I. Sanchez, G. Litjens, F.-E. de Leeuw, B. van Ginneken, E. Marchiori, and B. Platel, "Location Sensitive Deep Convolutional Neural Networks for Segmentation of White Matter Hyperintensities," *Scientific Reports*, vol. 7, no. 1, p. 5110, 2017.

[6] E. Ooms, "Automated segmentation of lacunes of presumed vascular origin in brain MRI scans," tech. rep., 2021.

[7] O. Ronneberger, Philipp Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention MICCAI 2015. Lecture Notes in Computer Science*, vol. 9351, pp. 234–241, Springer, Cham, 2015.

[8] Z. He, H. Liu, Y. Wang, and J. Hu, "Generative adversarial networks-based semi-supervised learning for hyperspectral image classification," *Remote Sensing*, vol. 9, no. 10, 2017.

[9] J. L. Marti-Vilalta, A. Arboix, and J. P. Mohr, "Lacunes," in *Stroke: Pathophysiology, Diagnosis, and Management*, ch. 11, pp. 275–299, Churchill Livingstone, jan 2004.

[10] J. M. Biesbroek, H. J. Kuijf, Y. van der Graaf, K. L. Vincken, A. Postma, W. P. T. M. Mali, G. J. Biessels, M. I. Geerlings, and o. b. o. t. S. S. Group, "Association between Subcortical Vascular Lesion Location and Cognition: A Voxel-Based and Tract-Based Lesion-Symptom Mapping Study. The SMART-MR Study," *PLoS ONE*, vol. 8, p. e60541, apr 2013.

[11] V. Fonov, A. C. Evans, K. Botteron, C. R. Almli, R. C. McKinstry, and D. L. Collins, "Unbiased average age-appropriate atlases for pediatric studies.," *NeuroImage*, vol. 54, no. 1, pp. 313–27, 2011.

[12] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, "Elastix: a Toolbox for Intensity-Based Medical Image Registration.," *IEEE transactions on medical imaging*, vol. 29, no. 1, pp. 196–205, 2010.