DS6: Deformation-aware learning for small vessel segmentation with small, imperfectly labeled dataset

Soumick Chatterjee⋆1,2,3, Kartik Prabhu⋆1, Mahantesh Pattadkal⋆1, Gerda Bortsova4, Florian Dubost4, Hendrik Mattern3, Marleen de Bruijne4,5, Oliver Speck3,6,7,8, and Andreas Nünberger1,2

1 Faculty of Computer Science, Otto von Guericke University Magdeburg, Germany
2 Data and Knowledge Engineering Group, Otto von Guericke University, Magdeburg, Germany
3 Biomedical Magnetic Resonance, Otto von Guericke University Magdeburg, Germany
4 Biomedical Imaging Group Rotterdam, Department of Radiology & Nuclear Medicine, Erasmus MC, The Netherlands
5 Department of Computer Science, University of Copenhagen, Denmark
6 German Center for Neurodegenerative Disease, Magdeburg, Germany
7 Center for Behavioral Brain Sciences, Magdeburg, Germany
8 Leibniz Institute for Neurobiology, Magdeburg, Germany
soumick.chatterjee@ovgu.de

Abstract. Originating from the initial segment of the middle cerebral artery of the human brain, Lenticulostriate Arteries (LSA) are a collection of perforating vessels that supply blood to the basal ganglia region. With the advancement of 7 Tesla scanner, we are able to detect these LSA which are linked to Small Vessel Diseases (SVD) and potentially a cause for neurodegenerative diseases. Segmentation of LSA with traditional approaches like Frangi or semi-automated/manual techniques can depict medium to large vessels but fail to depict the small vessels. Also, semi-automated/manual approaches are time-consuming. In this paper, we put forth a study that incorporates deep learning techniques to automatically segment these LSA using 3D 7 Tesla Time-of-flight Magnetic Resonance Angiogram images. The algorithm is trained and evaluated on a small dataset of 11 volumes. Deep learning models based on Multi-Scale Supervision U-Net accompanied by elastic deformations to give equivariance to the model, were utilized for the vessel segmentation using semi-automated labeled images. We make a qualitative analysis of the output with the original image and also on imperfect semi-manual labels to confirm the presence and continuity of small vessels.

Keywords: Lenticulostriate Arteries · Small vessel segmentation · Deep learning · MR Angiograms · 7 Tesla MRI · TOF-MRA · High Resolution MRI · Imperfect ground-truth

⋆ S. Chatterjee, K Prabhu and M Pattadkal have Equal Contribution

DS6 = Deformation-aware Segmentation, trained on 6 MR Volumes. The name pays homage to Star Trek: Deep Space Nine (in short: DS9)
1 Introduction

Lenticulostriate Arteries (LSA) are terminal branches supplying blood to large portions of the basal ganglia\cite{9} \cite{8} with minimal overlap of perfusion territories. Recently, 7 Tesla Time-of-flight Magnetic resonance angiogram was able to depict LSA non-invasively\cite{11} \cite{13}. LSA is associated with aging, dementia and Alzheimer \cite{21} \cite{24}. Further, the detection of LSA is critical to study Small Vessel Diseases(SVD) non-invasively in detail. In this paper, as small vessels, we will consider an apparent diameter of 1-2 voxels.

The manual segmentation of these small vessels is a reliable method, but the process is time-consuming and laborious. Traditional approaches\cite{10} \cite{3} \cite{12} \cite{6} which are capable of extracting vessels based on structural properties and semi-automated (involves manual intervention in annotating vessels and machine learning techniques to extracts features based on the annotations of observer and classify the non-annotated pixels of the image ) techniques can capture medium to large vessels but fail in detecting LSA. Also, a critical problem on the semi-automatic segmentation \cite{19} of vessels is generating spurious points in the segmentation result\cite{6}.

We propose instead to segment LSA using machine learning approaches based on convolution neural networks. U-Net\cite{18} and Attention U-Net\cite{16} are neural networks architectures designed for segmentation. They have proven to be promising techniques in biomedical image segmentation. Thereby we use both the above architectures as baselines for our LSA segmentation approach. We propose a network combining Multiscale Supervision\cite{23} in U-Net with elastic deformations, that learns on a small dataset of 6 volumes and is consistent under elastic transformations.

1.1 Related work

Vessel segmentation with manual filtering: C.Hsu et al \cite{12}, attempted using the Cardano formula to assign each voxel to binary value 0 or 1 as non-vessel or vessel region respectively. This technique requires manual feedback for filter parameters such as sphericity, tubeness, background suppression, and scale parameter. The output response shows the technique has improved the contrast of small vessels and also delineated them well. Machine learning is proposed to be a follow-up method of this technique.W.Liao et al \cite{14}, suggested a two-step approach for the detection of LSA. The first step is to detect all the major vessels and most parts of thin vessels. Then the second step involves detecting the vessel gaps and regions with low-contrast or noisy regions where it is difficult to detect LSA. This step involves two minimal path approaches, one being probabilistic mapping and other being fast marching. However minimal path approach involves the manual intervention of the observer for keypoint detection. Hence the accuracy of segmentation depends on the information provided by the observer and also it is time-consuming.
**Traditional automatic techniques:** Bernier et al \cite{3} used Frangi filter \cite{3} to segment cerebral vessels from susceptibility weighting imaging (SWI) and time-of-flight angiography (TOFMRA) dataset. Initially, the images are denoised, and intensity-thresholding is done to classify voxels as a vessel or non-vessel region. Then the score map is subjected to the multiscale filtering approach by Frangi et al\cite{2}. Finally, the vessel enhancement diffusion (VED) filter was used to enhance the vesselness output by combining the images obtained from multiple scales.

**Deep Learning techniques:** U-Net is used in many biomedical image segmentation tasks \cite{17}\cite{4}\cite{17} and the main advantage of U-Net model is that it can be trained end-to-end with few training samples. O.Oktay et al \cite{16}, implemented Attention U-Net for pancreas segmentation. It is observed that the latter improves the segmentation accuracy but with less computational effort. In \cite{5}, an architecture was proposed for chest X-ray segmentation, with a dataset consisting of both labeled and unlabeled 2D images. The model is trained on supervised loss from labeled data and unsupervised consistency loss between an elastically transformed output of a Siamese network made up of U-Nets. Since this paper dealt with a small dataset, we considered this as our reference and leverage this strategy for our small dataset at hand.

**1.2 Contribution**

Conventionally, deep learning based segmentation requires a large set of properly labeled data. The main contribution of the paper is to achieve good segmentation results despite the following problems of i) very small data set and ii) imperfectly labeled data. To overcome these two challenges, in this paper, we propose a novel architecture for the detection of LSA, then we compare the performance of our proposed approach against the state-of-the-art approaches.

The contribution of the work can be summarised as: i) comparison of our proposed model against non-Deep Learning\cite{3},\cite{3} and Deep Learning\cite{18},\cite{16} based methods for biomedical image segmentation, on our imperfectly labeled small dataset. ii) Better segmentation performance for small vessels with high variance in shape and volume iii) Improved generalization of the model and making it equivariant to a certain set of elastic transformations by leveraging the consistency loss in the network.

**2 Methods**

**2.1 Dataset and Label Creation**

The dataset comprises TOF MRIs of 11 individuals, scanned at a 7T MR scanner with an isotropic resolution of 300 µm. The dataset was divided into train, validation and test in the ratio 6:2:3 randomly.
Prospective motion correction (PMC) technique was utilized on all the subjects to improve the resolution. PMC is proven to improve the edge strength of vessels. The labels for these MRA images were annotated using Ilastik by a computer scientist and not a neurologist trained in the field of the cerebral vasculature. The label generated is still imperfect because of the limitations of the tool, as it contains gaps and does not capture some small vessels. However, the semi-automated method is not a bad technique, in Fig 3 it can be observed qualitatively, labeled image and proposed deep learning model gave similar segmentation results. Fig 1 shows the MIP of the original volume and the segmentation result achieved using Ilastik.

![Fig. 1: Axial MIPs of input volume and Semi-automated label](image)

### 2.2 Proposed Methodology

The proposed methodology is based on the combination of Multi-Scale Supervision (MSS) in U-Net with elastic deformations. The U-Net is modified to Multi-Scale Supervision (MSS) by computing the overall loss as the sum of losses at each up-sampling scale of the U-Net in its expansion path. The output at each scale is interpolated to the size of the original segmentation mask for loss calculation. Backpropagating these losses, allows the architecture to improve performance at each level and thereby making the gradient flow to earlier stages which enhances the learning of the network in comparison to U-Net. Eq.1 represents the loss function of U-Net MSS. The $m$ refers to the total up-sampling scales in the U-Net, $\ell_{scale}$ is the loss at each up-sampling scale, $\alpha_i$ is the weight assigned to loss at a specific up-sampling level and network parameter $\theta$

$$L_{MSS}(\theta) = \frac{1}{m} \sum_{i=1}^{m} \alpha_i \ell_{scale;i}(\theta)$$  \hspace{1cm} (1)
Let $\mathcal{X}$ be the set of input volumes while $\mathcal{Y}$ is the set of corresponding labels and $\mathcal{T}$ be the set of elastic transformations. The proposed network is based on usage of Siamese architecture, it has two identical branches and the first branch is fed with the tuple $(x, y)$ where $(x \sim \mathcal{X}, y \sim \mathcal{Y})$, while the second branch is fed with the elastically transformed volume and label $(t(x), t(y))$ where $t \sim \mathcal{T}$. These tuples are passed through the U-Net MSS to derive segmentation outputs $\hat{y}_1$ and $\hat{y}_2$ respectively. These outputs are compared with the corresponding labels to derive the Supervised loss depicted in Eq.2. Further the $\hat{y}_1$ is elastically transformed by $t(\hat{y}_1)$ to $\bar{y}$. Now the $\bar{y}$ is compared with $\hat{y}_2$ for computing the consistency loss shown in Eq.3. The network is trained to find optimal value of $\theta$ that minimizes the overall loss defined as the sum of Supervised and the Consistency loss.

$$
L_{\text{Sup}}(\theta) = \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} L_{\text{MSS}}(x, y, \theta) + L_{\text{MSS}}(t(x), t(y), \theta) \quad (2)
$$

$$
L_{\text{Cons}}(\theta) = \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} E_{t \sim \mathcal{T}} [C(t(f(x); \theta), f(t(x); \theta))] \quad (3)
$$

As the training dataset is small, this network learns consistency under elastic transformations and assists in terms of data augmentation. The Supervised loss $L_{\text{Sup}}(\theta)$ trains the network to lower the segmentation loss in comparison to ground truth and the Consistency loss $L_{\text{Cons}}(\theta)$ ensures that the network is equivariant to the elastic transformation.

### 2.3 Implementation

The Siamese network consists of U-Net MSS as the base. Every volume is converted to 3D patches of size 64X64X64 with a stride of 32,32,16 across the 3D space, to get overlapped samples and allow the network to learn continuity. Each of the 50 epochs randomly selects 8000 patches for training and validation. The focal-tversky loss was used for gradient back-propagation and Adam optimizer was used for both training. The deformation fields for elastic deformations were created by randomly sampling two-dimensional displacement maps for each slice with scale parameters chosen randomly in $(1,5)$ and smoothing them with a Gaussian filter of size 15 pixels with the standard deviation of 100 pixels. Nearest neighbor interpolation was applied to input images, labels, and predictions.

The network was trained on Nvidia Tesla V100-SXM2-32GB for 50 epoch which took 40 hours. We observed a significant decrease in memory utilization with mixed precision training. Mixed precision decreases the memory consumption for the model by using half-precision values (16-bit floating point) instead of regular full precision values (32-bit floating point). The training setup could hold only 8 3D patches (64X64X64) with FP32, while the mixed-precision allowed the model to train a batch of 20 patches (64X64X64) thereby training more samples per batch, effectively reducing the time needed for each epoch to finish.
3 Results & Discussion

We have evaluated the performance of non Deep Learning as well as deep learning models on the test set that consists of 3 volumes that were chosen randomly from the dataset. In this section, we present and discuss quantitative and qualitative evaluations. As a pre-processing step, the bias field corrected[20] input was provided to the non deep learning methods. Though for the deep learning models (baseline and also our proposed model), no such pre-processing was performed.

3.1 Quantitative evaluation

The quantitative evaluation is based on the dice coefficient and IOU. A detailed comparison of the non-Deep Learning and Deep Learning methods is shown in Table 1. The U-Net MSS with deformations outperforms all the other models in terms of dice coefficient and has a comparable variance.

Table 1: Comparison of metrics for Non-Deep Learning and Deep Learning methods

| Type  | Method                        | Dice Co-efficient | IOU    |
|-------|-------------------------------|-------------------|--------|
| Non-DL| Frangi Filter[10]             | 5.92± 0.66        | 6.10± 0.70 |
| Non-DL| Vessel Enhancement Diffusion filter[3] | 22.92± 0.40       | 25.89± 0.51 |
| DL    | U-Net                         | 76.69± 0.33       | 62.20± 0.44 |
| DL    | Attention U-Net               | 74.92±0.23        | 59.90±0.30 |
| DL    | Proposed Model                | 79.45±0.91        | 65.89±1.25 |

3.2 Qualitative evaluation

As part of the qualitative analysis, we looked for the regions where the small vessels were detected successfully and analyzed the segmentation performance of the model on the entire volume.

Region of Interest  As seen in Fig[7], the highlighted region of interest in yellow, the non-deep learning method fails to even segment larger vessels, whereas U-Net and Attention U-Net fails to keep the continuity in the small vessels. The proposed model U-Net with MSS and deformation is not only capable of detecting the small vessels but also successful in maintaining the continuity of shapes of these small vessels. In Fig. [4], we can see that the model has detected the vessels which are visible in the MIP of the input volume and not detected in the labeled volume as well as the non-DL method.
Fig. 2: Comparison of a region of interest containing a small vessel for different models. The green and red color corresponds to over and under segmentation respectively in comparison to the labels

4 Conclusion, Limitations & Future Work

The proposed network uses U-Net Multiscale Supervision architecture with deformations which makes it equivariant to elastic deformations and thereby acts as data augmentation for the small training dataset. The evaluation of this network on small vessel segmentation with MR angiograms dataset showed that it outperforms the traditional methods as well as current Deep learning methods like U-Net and Attention U-Net. The model though has few limitations; it is predicting some noise as vessels. Prior knowledge of vessels being cylindrical in nature is disrupted with deformation to a certain extent. Also since the dice coefficient is calculated with respect to an imperfect labeled image, some of the inherently detected vessels in the proposed model are considered as over-segmentation. In future work, we would like to explore the behavior of the proposed architecture.
using different scales as deformations. Additionally, we would like to consider the Maximum Intensity Projection loss to improve the continuity in the vessel segmentation. Further, the balance between over and under segmentation can be tuned by adjusting parameters in Focal-Tversky loss. Currently, any kind of post-processing hasn’t been used in our approach. Using Conditional Random Fields (CRF) as post-processing might further improve the results, which can be explored in the future.

Acknowledgment

This work was in part conducted within the context of the International Graduate School MEMoRIAL at Otto von Guericke University (OVGU) Magdeburg, Germany, kindly supported by the European Structural and Investment Funds (ESF) under the programme Sachsen-Anhalt WISSENSCHAFT Internationalisierung (project no. ZS/2016/08/80646). The authors would also like to acknowledge the contribution of Harish Kumar Harihara Subramanian of the Faculty of Computer Science, Otto von Guericke University Magdeburg, Germany; for his efforts with data handling and processing.
Fig. 4: Comparison of a region of interest containing a small vessel that is present in MIP of input image (a), not captured in label (b) or non-DL method (Bernier)(c), but detected by proposed model (d)
References

1. Abraham, N., Khan, N.M.: A novel focal tversky loss function with improved attention u-net for lesion segmentation. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). pp. 683–687 (April 2019). https://doi.org/10.1109/ISBI.2019.8759329

2. Berg, S., Kutra, D., Kroeger, T., Strachle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J.I., Xu, B., Beuttenmueller, F., Wolny, A., Zhang, C., Koethe, U., Hamprecht, F.A., Kreshuk, A.: ilastik: interactive machine learning for (bio)image analysis. Nature Methods 16(12), 1226–1232 (2019). https://doi.org/10.1038/s41592-019-0582-9

3. Bernier, M., Cunnane, S.C., Whittingstall, K.: The morphology of the human cerebrovascular system. Human Brain Mapping 39(12), 4962–4975 (2018). https://doi.org/10.1002/hbm.24337

4. Blanc-Durand, P., Van Der Gucht, A., Schaefer, N., Itti, E., Prior, J.O.: Automatic lesion detection and segmentation of 18F-FET PET in gliomas: A full 3D U-Net convolutional neural network study. PLoS ONE 13(4), 1–11 (2018). https://doi.org/10.1371/journal.pone.0195798

5. Bortsova, G., Dubost, F., Hogeweg, L., Katramados, I., de Bruijne, M.: Semi-supervised Medical Image Segmentation via Learning Consistency Under Transformations. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 11769 LNCS, 810–818 (2019). https://doi.org/10.1007/978-3-030-32226-7_90

6. Chen, L., Mossa-Basha, M., Balu, N., Canton, G., Sun, J., Pimentel, K., Hatsukami, T.S., Hwang, J.N., Yuan, C.: Development of a quantitative intracranial vascular features extraction tool on 3D MRA using semi-automated open-curve active contour vessel tracing. Magnetic Resonance in Medicine 79(6), 3229–3238 (2018). https://doi.org/10.1002/mrm.26961

7. Fabijan, A.: Segmentation of corneal endothelium images using a U-Net-based convolutional neural network. Artificial Intelligence in Medicine 88, 1–13 (2018). https://doi.org/10.1016/j.artmed.2018.04.004

8. Feeckes, J., Cassell, M.: The vascular supply of the functional compartments of the human striatum. Brain: a Journal of Neurology 120(1), 2189–201 (2006)

9. Feeckes, J., Hsu, S., Chaloupka, J., Cassell, M.: Tertiary microvascular territories define lacunar infarcts in the basal ganglia. Annals of Neurology 58(1), 18–30 (2005)

10. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 1496 (February), 130–137 (1998). https://doi.org/10.1007/bfb0056195

11. Hendrikse, J., Zwanenburg, J.J., Visser, F., Takahara, T., Luijten, P.: Non-invasive depiction of the lenticulostriate arteries with time-of-flight MR angiography at 7.0 T. Cerebrovascular Diseases 26(6), 624–629 (2008). https://doi.org/10.1159/000168638

12. Hsu, C.Y., Ghaffari, M., Alaraj, A., Flannery, M., Zhou, X.J., Linninger, A.: Gap-free segmentation of vascular networks with automatic image processing pipeline. Computers in Biology and Medicine 82(January), 29–39 (2017). https://doi.org/10.1016/j.compbiomed.2017.01.012
13. Kang, C.K., Park, C.W., Han, J.Y., Kim, S.H., Park, C.A., Kim, K.N., Hong, S.M., Kim, Y.B., Lee, K.H., Cho, Z.H.: Imaging and analysis of lenticulostriate arteries using 7.0-Tesla magnetic resonance angiography. Magnetic Resonance in Medicine 61(1), 136–144 (2009). https://doi.org/10.1002/mrm.21786
14. Liao, W., Rohr, K., Kang, C.K., Cho, Z.H., Fellow, L., Wörz, S.: Automatic 3D segmentation and quantification of lenticulostriate arteries from high-resolution 7 tesla MRA images. IEEE Transactions on Image Processing 25(1), 400–413 (2016). https://doi.org/10.1109/TIP.2015.2499085
15. NVIDIA/apex: A pytorch extension: Tools for easy mixed precision and distributed training in pytorch. https://github.com/NVIDIA/apex. Accessed:2020-03-08.
16. Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B., Rueckert, D.: Attention U-Net: Learning Where to Look for the Pancreas (Midl) (2018), http://arxiv.org/abs/1804.03999
17. Raza, S.E.A., Cheung, L., Epstein, D., Pelergaris, S., Khan, M., B, N.M.R.: Mi-MONet: Gland Segmentation Using Neural Network. Computer Methods and Programs in Biomedicine 1(d), 698–706 (2017). https://doi.org/10.1007/978-3-319-60964-3
18. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 9351, 234–241 (2015). https://doi.org/10.1007/978-3-319-24574-4_28
19. Sommer, C., Straehle, C., Kothe, U., Hamprecht, F.A.: Ilastik: Interactive learning and segmentation toolkit. Proceedings - International Symposium on Biomedical Imaging (September 2014), 230–233 (2011). https://doi.org/10.1109/ISBI.2011.5872394
20. Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C.: N4itk: improved n3 bias correction. IEEE Transactions on Medical Imaging 29(6), 1310–1320 (2010). https://doi.org/10.1109/TMI.2010.2046908
21. Wardlaw, J.M., Smith, C., Dichgans, M.: Mechanisms of sporadic cerebral small vessel disease: Insights from neuroimaging. The Lancet Neurology 12(5), 483–497 (2013). https://doi.org/10.1016/S1474-4422(13)70060-7, http://dx.doi.org/10.1016/S1474-4422(13)70060-7
22. Wikipedia: Start trek: Deep space nine. https://en.wikipedia.org/wiki/Star_Trek:_Deep_Space_Nine/ last accessed on 16-Mar-2020
23. Zeng, G., Yang, X., Li, J., Yu, L., Heng, P.A., Zheng, G.: 3D U-net with multi-level deep supervision: Fully automatic segmentation of proximal femur in 3D MR images. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2017). https://doi.org/10.1007/978-3-319-67389-9_32
24. Zwanenburg, J.J., Van Osch, M.J.: Targeting cerebral small vessel disease with MRI. Stroke 48(11), 3175–3182 (2017). https://doi.org/10.1161/STROKEAHA.117.016996
Appendix

Fig. 5: The proposed network is based on Siamese architecture, the input is fed to the first branch while the second branch receives the elastically deformed version of the input, both the branches use the same U-Net MSS for segmentation. The segmentation output of the first branch is elastically deformed and is compared with the output of the second branch for consistency.
Fig. 6: MIP of the input volume, labeled volume & output from the proposed model (left to right). In the highlighted region of interest, the MIP tells us that the label is imperfect, but even so, the model predicts the vessels as over-segmentation (green) when compared against the label and maintains continuity.

Fig. 7: Whole volume (left to right) MIP of an input image, label, non-DL method (Bernier) and output of the proposed model. The green and red color corresponds to over and under segmentation respectively in comparison to the labels.