Anatomically-Guided Segmentation of Cerebral Microbleeds in T1-weighted and T2*-weighted MRI

Junmo Kwon[1 [0000−0002−2242−1637]], Sang Won Seo[2], and Hyunjin Park[1,3 [0000−0001−5681−8918]]

[1] Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon 16419, South Korea
skenfn1231@skku.edu
[2] Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, South Korea
sw72.seo@samsung.com
[3] Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon 16419, South Korea
hyunjinp@skku.edu

Abstract. Cerebral microbleeds (CMBs) are defined as relatively small blood depositions in the brain that serve as severity indicators of small vessel diseases, and thus accurate quantification of CMBs is clinically useful. However, manual annotation of CMBs is an extreme burden for clinicians due to their small size and the potential risk of misclassification. Moreover, the extreme class imbalance inherent in CMB segmentation tasks presents a significant challenge for training deep neural networks. In this paper, we propose to enhance CMB segmentation performance by introducing a proxy task of segmentation of supratentorial and infratentorial regions. This proxy task could leverage clinical prior knowledge in the identification of CMBs. We evaluated the proposed model using an in-house dataset comprising 335 subjects with 582 longitudinal cases and an external public dataset consisting of 72 cases. Our method performed better than other methods that did not consider proxy tasks. Quantitative results indicate that the proxy task is robust on unseen datasets and thus effective in reducing false positives. Our code is available at https://github.com/junmokwon/AnatGuidedCMBSeg.

Keywords: Cerebral microbleeds · Image segmentation · Proxy tasks

1 Introduction

Cerebral microbleeds (CMBs) are characterized as small, round accumulations of blood products, each spanning just a few millimeters in diameter [8]. These lesions can result from specific pathologies, such as cerebral amyloid angiopathy (CAA) and hypertensive microangiopathy, which could potentially damage the vascular architecture of the brain [8,10,4]. Beyond these conditions, aging is one

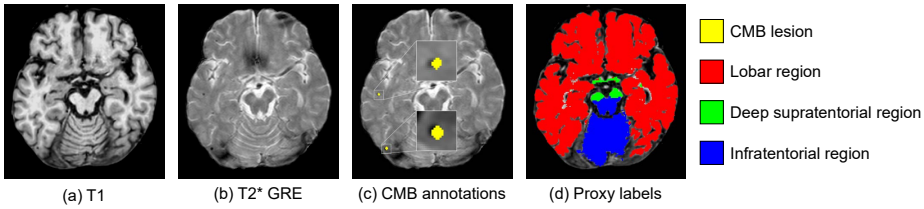

Fig. 1. An example of CMB lesions in an axial view of (a) T1-weighted MRI and (b) GRE T2*-weighted MRI with (c) corresponding manual annotation of CMBs and (d) automatically labeled lobar, deep supratentorial, and infratentorial regions.

of the important factors that correlates with an increased number of CMBs [10,8]. Thus, the accurate detection of CMBs serves as an important biomarker of small vessel disease in the brain.

Typically, CMBs are represented as hypointense lesions in the T2*-weighted gradient-recalled echo (GRE) sequence [8]. Unlike the conventional T1-weighted or T2-weighted MRI, the T2*-weighted GRE sequence is sensitive to the blooming effect caused by CMBs [8]. Despite the usefulness, a major concern arises from the relatively low specificity in the detection of CMBs. False positives of CMBs include calcium or iron deposits, blood flow in pial vessels, and melanin, which mimics the susceptibility effect of CMBs on the T2*-weighted MRI [8]. Thus, manually distinguishing CMB mimics presents a significant burden on clinicians.

Several studies have suggested automated frameworks for the segmentation and detection of CMBs to alleviate the burden on clinicians in identifying CMBs. Notably, introducing proxy labels [15], as well as single-stage [16,14] and dual-stage pipelines [5,17,23] for the identification of CMBs and false positive reduction, have been proposed. Kim et al. [15] showed the effectiveness of using proxy labels for anatomical localization in CMB detection. Lee et al. [16] proposed a single-stage CMB detection network that ensembles information from axial, coronal, and sagittal planes. Kim et al. [14] unified 3D U-Net [22] and Mask R-CNN [11] into a single network, significantly improving precision compared to conventional region proposal networks. Chen et al. [5] implemented statistical thresholding based on the intensity distribution of CMBs as a preprocessing step. Liu et al. [17] applied the 3D fast radial symmetry transform for false positive reduction. Sanguesa et al. [23] introduced Mask R-CNN [11] to detect candidate regions of CMBs. Despite the effort to eliminate false positives of CMBs, existing methods heavily rely on data-driven algorithms and often overlook the incorporation of clinical prior knowledge.

In this paper, we propose a segmentation framework for the identification of CMBs by incorporating a proxy task designed to align with Greenberg's criteria [8]. Specifically, our method extends a binary segmentation of CMBs on T2*-weighted MRI by implementing a segmentation of critically related brain regions in CMB identification: lobar, deep supratentorial, and infratentorial re-

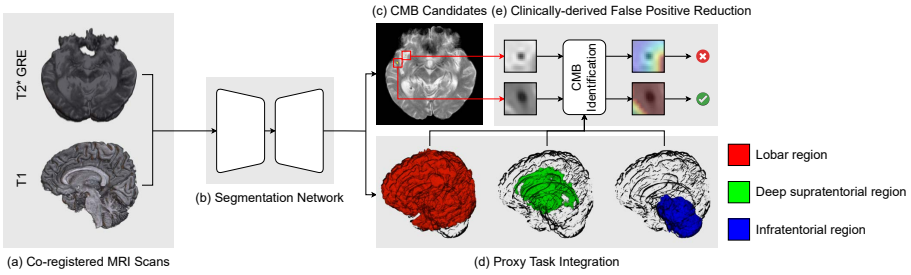

Fig. 2. An overview of our CMB segmentation framework. Our model takes (a) co-registered multi-modal MRI inputs. (b) Then, a segmentation model jointly produces (c) predicted CMB candidate regions and (d) proxy labels, followed by (e) a false positive reduction module that employs clinical insights toward CMB identification.

gions, as suggested in [9,21]. We visualized CMB annotations besides related brain regions as proxy labels in Fig. 1. Especially, identifying whether microbleeds are lobar or deep is crucial for treatment planning. Lobar microbleeds are most commonly associated with CAA [8]. On the other hand, deep microbleeds are primarily related to chronic hypertension and hypertensive arteriopathy [8]. Our proxy label generation process is fully automated by utilizing FreeSurfer [6] and ANTs [1]. Our proxy task enables us to localize probable CMBs while effectively reducing false positives. Furthermore, our model could assist clinicians by providing an automated assessment of the visual grading of CMBs, namely the Microbleed Anatomical Rating Scale (MARS) [9], based on the predicted lobar, deep supratentorial, and infratentorial regions.

The contribution of this paper can be summarized as follows. First, we propose the integration of a proxy task in the segmentation of CMBs by introducing the segmentation of lobar, deep supratentorial, and infratentorial regions without the need for extra manual annotation. Second, we show that predicted proxy labels can effectively reduce false positives, particularly on unseen datasets. Lastly, we demonstrate the reproducibility of our model in both in-house and external public datasets.

2 Methods

The overall architecture of our proxy task integrated into the CMB segmentation network is shown in Fig. 2. Specifically, our framework consists of two parts: a segmentation network and a false positive reduction module.

2.1 Segmentation of CMB Lesions with Proxy Labels

For a given T1-weighted and T2*-weighted MRI scan, the goal of the task extends to not only segmenting CMB lesions but also segmenting the related brain regions into proxy labels, such as lobar, deep supratentorial, and infratentorial

Algorithm 1 Pseudo-code of proxy label generation.

Input: T1-weighted MRI M and JHU-DTI atlas J
Output: Lobar L, deep supratentorial D and infratentorial region I
Register J from MNI152 space to MNI305 space ▷ rigid-body transformation
for all subjects **do**
    Perform FreeSurfer parcellation on M
    Compute Talairach transformation T via FreeSurfer
    for all segmented regions P_i **do**
        if $P_i \in$ Cerebral lobes **then** $L \leftarrow L \cup P_i$
        if $P_i \in$ Deep supratentorial region **then** $D \leftarrow D \cup P_i$
        if $P_i \in$ Infratentorial region **then** $I \leftarrow I \cup P_i$
    end for
    Register J from MNI305 space to native space of M using T
    Select internal capsule and external capsule from J
    for all selected white matter regions J_i **do**
        $D \leftarrow D \cup J_i$
    end for
end for

regions. Selecting a segmentation network is a design choice. Among many possible architectures, we adopted nnUNet [13] for our segmentation module.

Configuring loss functions. In [18], choosing an appropriate loss function plays a crucial role in segmentation tasks, especially when dealing with an extreme class imbalance between the foreground and background voxels. Typically, the diameter of CMB lesions does not exceed 10 mm [9,21], which makes the segmentation task challenging. In this scenario, we adopted DiceTopK loss [18,3], which is a linear combination of soft Dice loss and TopK loss to penalize easily classified voxels and help the network focus on challenging voxels.

Generating proxy labels. We leveraged whole brain parcellation results from FreeSurfer [6] to generate proxy labels for lobar, deep supratentorial, and infratentorial regions as described in Algorithm 1 and supplementary Fig. S1. Cerebral lobes including hippocampus and amydala were categorized as lobar regions. Deep supratentorial region includes thalamus, caudate, pallidum, corpus callosum, nucleus accumbens, ventral diencephalon, internal capsule (IC), external capsule (EC), and deep white matter voxels surrounding lateral ventricle, caudate, thalamus, and pallidum. To localize IC and EC, we registered IC and EC from the JHU-DTI atlas [19,20] to the subject's T1 space. Finally, the brainstem and cerebellum voxels were classified as infratentorial regions.

2.2 Clinically-derived False Positive Reduction

Upon successful segmentation of proxy labels, our framework can identify candidate CMB regions by calculating the total ratio of surrounding brain parenchyma

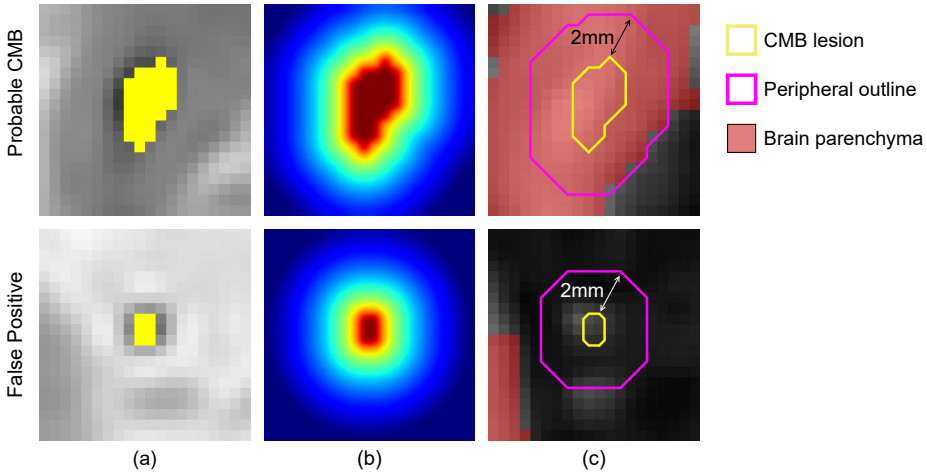

Fig. 3. Our clinically derived false positive reduction module. (a) Illustration of predicted CMB candidates. (b) Calculation of the Euclidean distance map, assuming the voxel resolution meets the in-plane isotropy. (c) Segmentation of peripheral voxels within a 2 mm range, followed by the quantification of peripheral voxels associated with brain parenchyma. Candidate regions are marked as false positive if a majority of the surrounding voxels are identified as background class.

as illustrated in Fig. 3. Specifically, brain parenchyma is derived from the union of predicted proxy labels. Euclidean distance map is then calculated to select voxels adjacent to CMB areas while ensuring in-plane isotropy, particularly in an axial view. If the voxel resolution is isotropic, adjacent out-of-plane slices are also considered during the selection of peripheral CMB voxels. When dealing with anisotropic voxel resolutions, especially when the out-of-plane slice thickness is larger than 2 mm, the selection criterion is limited to the in-plane axes. Finally, false positives are discarded if more than half of the voxels surrounding a CMB region are classified as non-brain regions, thereby leveraging anatomical insights to improve the specificity of the CMB detection.

3 Experiments

3.1 Datasets

In-house dataset. Our study used an in-house dataset obtained from Samsung Medical Center comprising 335 participants with 582 cases including longitudinal scans. We applied stratified sampling based on the number of CMB lesions and the total volume of CMBs for each case to split the training set and internal validation set with a ratio of 0.8 and 0.2, respectively. In sum, there are 483 training cases and 99 internal validation cases, ensuring no overlap of subjects between the different sets. The images were acquired using a 3T Philips scanner

with a spatial resolution of $0.5 \times 0.5 \times 0.5$ mm^3 for T1-weighted MRI scans and $0.4286 \times 0.4286 \times 6.5$ mm^3 for T2*-weighted MRI scans.

External public dataset. To further validate our framework, we incorporated an external public dataset collected from the MICCAI VALDO 2021 challenge [24]. Specifically, we used 72 publicly available cases from three different cohorts (SABRE, RSS, and ALFA). While the SABRE and ALFA studies utilized a 3T MRI scanner for MRI acquisition, the RSS study was performed with a 1.5T GE MRI scanner.

3.2 Data Preprocessing

T1-weighted MRI preprocessing. Given a T1-weighted MRI scan, we performed FreeSurfer cortical and subcortical parcellation [6,7]. The parcellation pipeline includes motion correction, skull-stripping, intensity inhomogeneity correction, and Talairach MNI space transformation. The parcellation results will then be used for proxy label generation.

T2*-weighted MRI preprocessing. Since T2*-weighted MRI scans are not ideally suited for processing with the FreeSurfer pipeline, we adopted SynthStrip [12] to perform skull-stripping including the surrounding cerebrospinal fluid (CSF) regions. Following skull-stripping, we applied the N4ITK algorithm [25] for bias field correction.

Multi-modal image registration. In order to feed both T1-weighted and T2*-weighted MRI scans onto the networks, T1-weighted MRIs underwent a registration process onto the T2* space by employing a rigid-body transformation via ANTs [1]. Moreover, to unify voxel spacings between our in-house dataset and the external public dataset, all of the VALDO 2021 cases were subsampled to fit a voxel resolution of $0.4286 \times 0.4286 \times 6.5$ mm^3, ensuring both datasets to have the same spatial dimensions.

3.3 Implementation Details

Due to the compatibility limitations of nnDetection [2] with PyTorch versions beyond 2.0, all networks in this study were implemented in Python 3.8 using PyTorch 1.11.0 for a fair comparison. Instead of using the default trainer in 3D full-resolution nnUNet [13], we applied a modified trainer to employ a linear combination of soft Dice loss and TopK loss with $k = 10$, $\lambda_{\text{dice}} = 0.5$, and $\lambda_{\text{topk}} = 0.5$.

Table 1. Qualitative results of CMB segmentation and detection task on internal validation set and external dataset. Bold is the best-performing case.

Method	Internal (n=99)		External (n=72)	
	Dice	F1-score	Dice	F1-score
nnDetection [2]	0.2420	0.2484	0.3146	0.3167
nnUNet [13]	0.4627	0.4714	0.2972	0.2860
Ours	**0.5138**	**0.5199**	**0.4688**	**0.4600**

Table 2. Ablation analysis on loss configurations and module designs. Bold is the best-performing case. DiceCE: Linear combination of Dice loss and cross-entropy loss. DiceTopK: Linear combination of Dice loss and TopK loss. Proxy: Proxy task integration. CFPR: Clinically-derived false positive reduction.

Method	Loss		Module		Internal (n=99)		External (n=72)	
	DiceCE	DiceTopK	Proxy	CFPR	Dice	F1-score	Dice	F1-score
nnUNet [13]	✓		✓		0.4485	0.5104	0.3628	0.3696
	✓		✓	✓	0.4488	0.5118	0.4281	0.4116
		✓	✓		0.5137	0.5182	0.4549	0.4461
		✓	✓	✓	**0.5138**	**0.5199**	**0.4688**	**0.4600**

3.4 Evaluation Metrics

For the CMB segmentation task, we use the Dice coefficient to quantify overlaps between predictions and ground truth labels. When identifying CMB lesions, the F1-score is utilized by calculating the shortest Euclidean distance between the centroids of predicted CMB regions and the centroids of ground truth regions, followed by applying a distance threshold of not exceeding 5 mm [24].

4 Results

To verify the effectiveness of our proxy task integration and clinically-derived false positive reduction (CFPR), we compared our results with state-of-the-art methods such as nnUNet [13] and nnDetection [2]. We reported quantitative metrics in Table 1 and qualitative results in Fig. 4 and supplementary Fig. S2. Our method consistently shows promising results compared to baseline methods on an unseen dataset with a different field strength and repetition time. We visualized challenging cases acquired in the 1.5T MRI scanner in supplementary Fig. S2, which did not show signal hyperintensity in CSF regions. This emphasizes that our model is robust to imaging quality as well as annotation quality, which is affected by clinicians' bias.

Ablation studies. We performed an ablation analysis on loss configurations and our proposed modules. The baseline result, excluding proxy tasks and CFPR, is reported as the nnUNet baseline in Table 1. Table 2 shows little performance improvement on the internal dataset while a significant boost in Dice coefficient

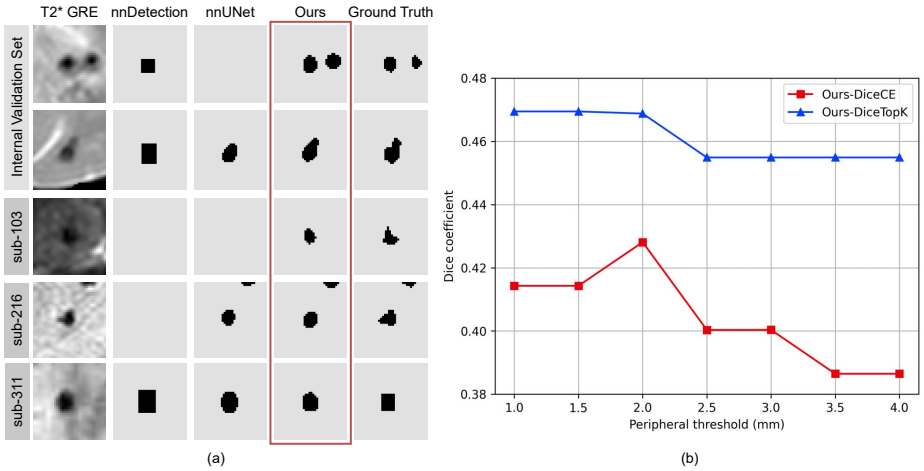

Fig. 4. (a) Qualitative results of our proposed method with nnDetection [2] and nnUNet [13] on internal validation set and external dataset. (b) Ablation analysis of clinically-derived false positive reduction module with varying peripheral thresholds on the external dataset. We chose 2 mm as an optimal peripheral threshold based on our model with DiceCE loss and DiceTopK loss.

and F1-score when applying proxy tasks and CFPR on an unseen dataset. The limited impact of the CFPR module on the internal dataset may be due to the training and validation sets originating from the same MRI scanner. In contrast, our proxy task is robust to unseen datasets due to its large size compared to CMB lesions, which contributes to enhancing the overall performance. The successful segmentation of proxy labels could further benefit our CFPR module both trained on DiceCE loss and DiceTopK loss. We also evaluated our model with different peripheral thresholds in the CFPR module as described in Fig. 4 (b), showing that a peripheral threshold of 2 mm is optimal in terms of the Dice coefficient.

5 Conclusion

In this paper, we propose proxy task integration and clinically-derived false positive reduction to improve the performance of CMB segmentation and detection. Specifically, we automatically derived proxy labels via FreeSurfer and ANTs to improve the overall performance of our model especially on unseen datasets. Furthermore, the prediction of proxy labels can be utilized in reducing false positives of CMBs. Without extra manual annotation done by clinicians, our method shows a remarkable improvement in terms of the Dice coefficient and F1-score in the external dataset.

Acknowledgments. We greatly thank the organizers and participants of the MICCAI VALDO 2021 challenge for the use of the public dataset. This study was supported

by National Research Foundation (NRF-2020M3E5D2A01084892), Institute for Basic Science (IBS-R015-D1), AI Graduate School Support Program (2019-0-00421), ICT Creative Consilience program (IITP-2024-2020-0-01821), and the Artificial Intelligence Innovation Hub program (RS-2021-II212068).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C.: A reproducible evaluation of ants similarity metric performance in brain image registration. Neuroimage **54**(3), 2033–2044 (2011)
2. Baumgartner, M., Jäger, P.F., Isensee, F., Maier-Hein, K.H.: nndetection: a self-configuring method for medical object detection. In: 24th International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 530–539. Springer (2021)
3. Brugnara, G., Isensee, F., Neuberger, U., Bonekamp, D., Petersen, J., Diem, R., Wildemann, B., Heiland, S., Wick, W., Bendszus, M., et al.: Automated volumetric assessment with artificial neural networks might enable a more accurate assessment of disease burden in patients with multiple sclerosis. European radiology **30**, 2356–2364 (2020)
4. Charidimou, A., Werring, D.J.: Cerebral microbleeds: detection, mechanisms and clinical challenges. Future Neurology **6**(5), 587–611 (2011)
5. Chen, H., Yu, L., Dou, Q., Shi, L., Mok, V.C., Heng, P.A.: Automatic detection of cerebral microbleeds via deep learning based 3d feature representation. In: IEEE 12th international symposium on biomedical imaging (ISBI). pp. 764–767. IEEE (2015)
6. Fischl, B.: Freesurfer. Neuroimage **62**(2), 774–781 (2012)
7. Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., et al.: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron **33**(3), 341–355 (2002)
8. Greenberg, S.M., Vernooij, M.W., Cordonnier, C., Viswanathan, A., Salman, R.A.S., Warach, S., Launer, L.J., Van Buchem, M.A., Breteler, M.M.: Cerebral microbleeds: a guide to detection and interpretation. The Lancet Neurology **8**(2), 165–174 (2009)
9. Gregoire, S., Chaudhary, U., Brown, M., Yousry, T., Kallis, C., Jager, H., Werring, D.: The microbleed anatomical rating scale (mars) reliability of a tool to map brain microbleeds. Neurology **73**(21), 1759–1766 (2009)
10. Haller, S., Vernooij, M.W., Kuijer, J.P., Larsson, E.M., Jäger, H.R., Barkhof, F.: Cerebral microbleeds: imaging and clinical significance. Radiology **287**(1), 11–28 (2018)
11. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask r-cnn. In: 2017 IEEE International Conference on Computer Vision (ICCV). pp. 2980–2988 (2017)
12. Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B., Hoffmann, M.: Synthstrip: skull-stripping for any brain image. NeuroImage **260**, 119474 (2022)
13. Isensee, F., Jaeger, P.F., Kohl, S.A., Petersen, J., Maier-Hein, K.H.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. Nature methods **18**(2), 203–211 (2021)

14. Kim, J.H., Al-masni, M.A., Lee, S., Lee, H., Kim, D.H.: Cerebral microbleeds detection using a 3d feature fused region proposal network with hard sample prototype learning. In: 25th International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 452–460. Springer (2022)
15. Kim, J.H., Noh, Y., Lee, H., Lee, S., Kim, W.R., Kang, K.M., Kim, E.Y., Al-masni, M.A., Kim, D.H.: Toward automated detection of microbleeds with anatomical scale localization: A complete clinical diagnosis support using deep learning. arXiv preprint arXiv:2306.13020 (2023)
16. Lee, H., Kim, J.H., Lee, S., Jung, K.J., Kim, W.R., Noh, Y., Kim, E.Y., Kang, K.M., Sohn, C.H., Lee, D.Y., et al.: Detection of cerebral microbleeds in mr images using a single-stage triplanar ensemble detection network (tpe-det). Journal of Magnetic Resonance Imaging **58**(1), 272–283 (2023)
17. Liu, S., Utriainen, D., Chai, C., Chen, Y., Wang, L., Sethi, S.K., Xia, S., Haacke, E.M.: Cerebral microbleed detection using susceptibility weighted imaging and deep learning. Neuroimage **198**, 271–282 (2019)
18. Ma, J., Chen, J., Ng, M., Huang, R., Li, Y., Li, C., Yang, X., Martel, A.L.: Loss odyssey in medical image segmentation. Medical Image Analysis **71**, 102035 (2021)
19. Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., Faria, A.V., Mahmood, A., Woods, R., et al.: Stereotaxic white matter atlas based on diffusion tensor imaging in an icbm template. Neuroimage **40**(2), 570–582 (2008)
20. Oishi, K., Zilles, K., Amunts, K., Faria, A., Jiang, H., Li, X., Akhter, K., Hua, K., Woods, R., Toga, A.W., et al.: Human brain white matter atlas: identification and assignment of common anatomical structures in superficial white matter. Neuroimage **43**(3), 447–457 (2008)
21. Potter, G.M., Sarah Keir, M.: Improving inter-rater agreement about brain microbleeds: development of the brain observer microbleed. Stroke **35**, 1831–1835 (2004)
22. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234–241. Springer (2015)
23. Sanguesa, M.G., Kutnar, D., van der Velden, B.H., Kuijf, H.J.: Mixmicrobleed: Multi-stage detection and segmentation of cerebral microbleeds. arXiv preprint arXiv:2108.02482 (2021)
24. Sudre, C.H., Van Wijnen, K., Dubost, F., Adams, H., Atkinson, D., Barkhof, F., Birhanu, M.A., Bron, E.E., Camarasa, R., Chaturvedi, N., et al.: Where is valdo? vascular lesions detection and segmentation challenge at miccai 2021. Medical Image Analysis **91**, 103029 (2024)
25. 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)