Blood Pressure Assisted Cerebral Microbleed Segmentation via Meta-matching

Junmo Kwon[1][0000−0002−2242−1637], Jonghun Kim[1,2][0009−0002−2790−2090], Taehyeon Kim[3][0000−0002−0609−4867], Sang Won Seo[4], Hwan-ho Cho[5][0000−0002−1926−6891], and Hyunjin Park[1,2][0000−0001−5681−8918]

[1] Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon 16419, South Korea
hyunjinp@skku.edu

[2] Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon 16419, South Korea

[3] Department of Computer and Information Technology, Purdue University, West Lafayette, IN 47907, USA

[4] Department of Neurology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul 06351, South Korea

[5] Department of Electronics Engineering, Incheon National University, Incheon 22012, South Korea
hwanho@inu.ac.kr

Abstract. Cerebral microbleeds (CMBs) are small hemorrhagic lesions that pose significant challenges for accurate segmentation due to the high rate of false positives and false negatives. CMBs have two subtypes: lobar and deep microbleeds (MBs). Motivated by the strong association between deep MBs and hypertension, we propose a **b**lood **p**ressure-driven **nnU-Net** (**BP-nnUNet**) that integrates blood pressure (BP) prompt into the state-of-the-art nnU-Net framework through three key strategies. First, we estimate BP using the pre-trained Meta-matching model, that requires only MRI images. This allows our method to be successfully applied to public datasets with missing clinical demographics. Second, we categorize CMBs into lobar and deep MB, enriching input text prompts with multiple classes while constraining the BP effect to deep MBs. Lastly, we introduce a novel anatomically-aware joint prompt fusion module that combines lobar and deep MB prompts. Experiments on both in-house and public datasets demonstrate that our BP-nnUNet outperforms existing CMB segmentation models and universal models incorporating medical prompts. Ablation studies validate the effectiveness of integrating subtype-level and case-level prompts, as well as our fusion module. Our method paves the way for the incorporation of clinically relevant information into a segmentation framework. Our code is available at https://github.com/junmokwon/BP-nnUNet.

Keywords: Cerebral microbleeds · Blood pressure · Meta-matching · Prompt-driven medical image segmentation

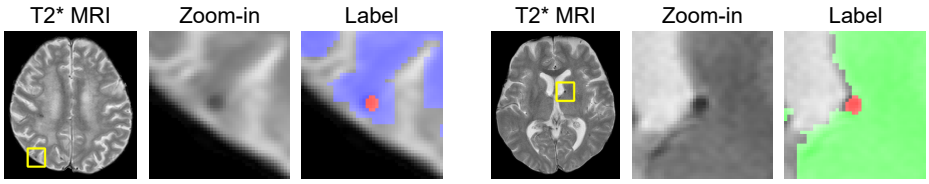

Fig. 1. Visualization of two subtypes of cerebral microbleeds: (a) Lobar microbleed and (b) deep microbleed. Red denotes microbleeds, blue denotes the lobar brain region, and green denotes the deep supratentorial region.

1 Introduction

Cerebral microbleeds (CMBs) are small brain hemorrhages that appear as hypointense lesions on T2*-weighted MRI [6,21]. The term *micro* indicates their small size, typically ranging from 2 to 10 mm in diameter [21]. Hemorrhages exceeding this size threshold are no longer classified as microbleeds (MBs) but as larger intracerebral hemorrhages [4]. Therefore, CMB segmentation is a challenging task that requires accurate localization of small lesions while dealing with diverse voxel resolutions, varying from a fine-grained spacing of 0.2 mm^3 to a coarse-grained spacing of 3.0 mm^3 [23].

Segmenting small lesions requires additional modifications to tune state-of-the-art (SOTA) segmentation backbones, ranging from modifying loss functions [19,14] to applying task-specific false positive reduction modules [15]. Existing studies on CMB segmentation have proposed various false positive reduction algorithms to overcome these challenges [15,16]. AG-nnUNet [15] extends the original binary segmentation of CMBs to multi-label segmentation across lobar, supratentorial, and infratentorial regions and CMB lesions using an nnU-Net [10] backbone. This multi-label approach enhances the identification of CMB lesions surrounded by brain parenchyma, thereby reducing false positives [5]. CMB-UNETR++ [16] adapted the masked image modeling [7] technique to pre-train both the encoder and decoder components of UNETR++ [22], demonstrating superior performance over plain nnU-Net and UNETR++ for CMB segmentation.

Recently, leveraging text prompts has shown remarkable improvements in medical image segmentation [17,18,26]. The CLIP-driven Universal Model [17,18] uses both image and CLIP-driven text embeddings as convolutional weights and biases in the final convolutional layers. UniSeg [26] partitions a prompt into universal and task-specific parts, achieving superior performance in both single and multiple tasks. These studies emphasize that fusing imaging features and text embeddings is particularly beneficial for medical image segmentation.

In this study, we propose BP-nnUNet, a blood pressure-driven nnU-Net framework for CMB segmentation. CMB lesions are pathologically categorized into lobar and deep MBs (Fig. 1). High blood pressure (BP) is associated with

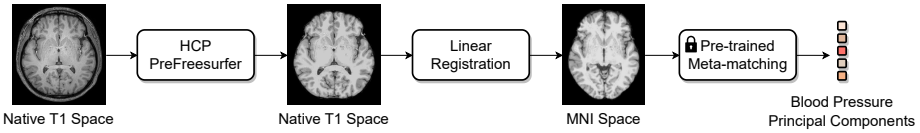

Fig. 2. Overview of the blood pressure (BP) estimation pipeline. T1-weighted MRI scans are processed using (1) HCP PreFreesurfer pipeline, (2) FMRIB's Linear Registration Tool to register T1 images to a 1 mm^3 MNI space, and (3) a pre-trained Meta-matching model, which outputs five phenotypes representing the principal components of BP measurements.

deep MB [6]. Our framework leverages medical text prompts from the pre-trained text encoder of BiomedCLIP [27]. In addition, we use the estimated BP from a pre-trained phenotype prediction model called Meta-matching [8,2,25], trained on large datasets. This BP estimation procedure only requires MRI data and thus can be applied not only to in-house datasets, but also to public datasets where clinical demographics are missing. Furthermore, we introduce a novel anatomically-aware joint prompt fusion module to effectively integrate our lobar and deep MB prompts into the segmentation network. We collected 621 cases of T1- and T2*-weighted MRI as an in-house dataset and evaluated our framework using the public MICCAI VALDO 2021 challenge dataset [23]. Our BP-nnUNet was evaluated against comparative methods, including task-specific models as well as SOTA universal models.

Our main contributions are three-fold:

1. We are the first to incorporate BP data, reflecting the pathology of CMB lesions, into the CMB segmentation framework. Unlike the task-level prompts used in universal models, BP reflects the underlying subject-level features.
2. The BP estimation procedure can be applied to anonymized public datasets without requiring access to clinical demographics.
3. We introduce a novel anatomically-aware joint prompt fusion module that achieves superior performance by leveraging a joint prompt of medical text and BP data.

2 Methods

2.1 Blood Pressure Estimation

Fig. 2 summarizes the overall procedure for BP estimation using a pre-trained Meta-matching framework [8,2,25]. Meta-matching was trained using 265 phenotypes from the UK Biobank dataset [1], including diastolic BP, systolic BP, and intraocular pressure [8]. While the original Meta-matching [8] was proposed using resting-state fMRI inputs, the subsequent study [25] extended its application to T1 MRI inputs, leveraging the simple fully convolutional network as proposed in [20]. Following the pre-processing pipeline for anatomical Meta-matching [3], each T1 MRI scan yields five distinct blood pressure components, referred to as *BP eye* C2 through *BP eye* C6.

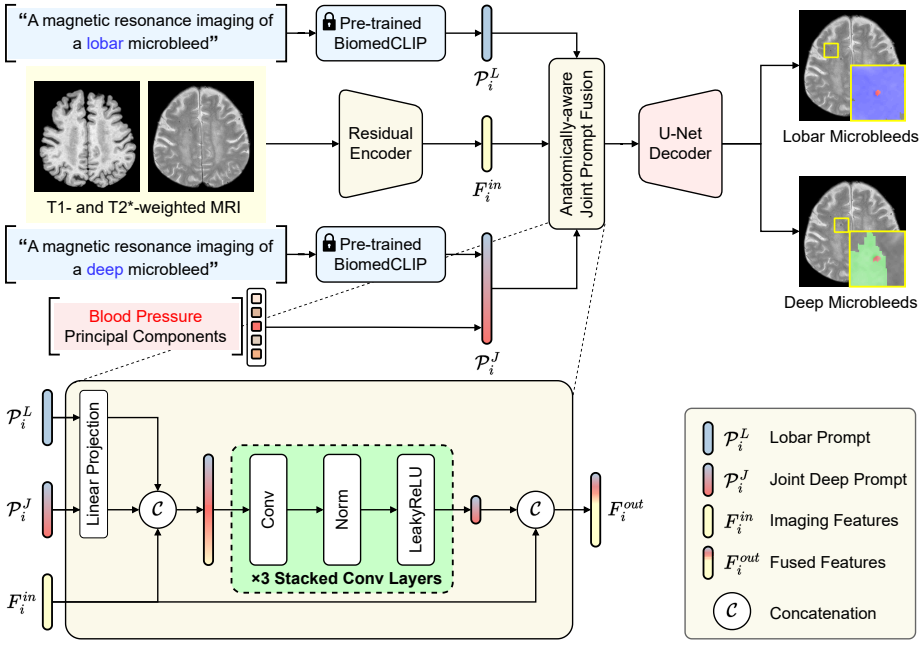

Fig. 3. Overview of our **BP-nnUNet**. Built upon nnU-Net ResEnc, we integrate lobar and deep microbleed prompts along with imaging features at the final layer of the residual encoder to predict binary segmentation masks for lobar and deep microbleeds.

2.2 Cerebral Microbleed Segmentation

We aim to develop a fully automated CMB segmentation framework by leveraging a joint prompt of medical text and BP data. Our proposed network, **BP-nnUNet**, is illustrated in Fig. 3. To seamlessly integrate imaging features and joint prompts, we propose an anatomically-aware joint prompt fusion module.

Subtyping approach. We categorize CMB lesions based on their pathology as lobar and deep MBs, as illustrated in Fig. 1, applying different strategies for each subtype. The process is automated and followed by proxy label generation as proposed in AG-nnUNet [15]. MBs in lobar brain regions are classified as lobar MBs, while those in infratentorial and deep supratentorial regions are categorized as deep MBs. We follow the standardized medical prompt of "A magnetic resonance imaging of [CLS]" as used in the CLIP-driven Universal Model [17,18], obtaining 512-dimensional sentence embeddings using BiomedCLIP [27]. For lobar MBs, we use only a medical prompt as an input prompt. By contrast, for deep MBs, we use both a medical prompt and the estimated BP as a joint prompt, motivated by the fact that hypertension is the primary cause of deep MBs [6]. Compared with a plain task-level prompt, our approach can enrich medical prompts based on their subtypes.

Anatomically-aware joint prompt fusion. We introduce a novel fusion approach, *anatomically-aware joint prompt fusion*, for CMB segmentation. As illustrated in Fig. 3, our module integrates all input imaging features and joint prompts, and processes them through stacked convolutional layers:

$$F_i^{out} = \text{cat}(F_i^{in}, \text{StackedConv}(\text{MLP}_L(\mathcal{P}_i^L), \text{MLP}_D(\mathcal{P}_i^J), F_i^{in})) \quad (1)$$

where F_i^{in} denotes the imaging features from the two MRI modalities (i.e., T1 and T2*) of the i-th subject, obtained from the last convolutional layer of the residual encoder. $\text{cat}(\cdot)$ denotes concatenation along the channel dimension, assuming given inputs have the same spatial dimensions. $\text{StackedConv}(\cdot)$ first concatenates all inputs and then processes them through a stacked convolutional layer, followed by instance normalization and LeakyReLU layers. There are two separate linear projection layers: MLP_L and MLP_D, each projecting lobar prompt $\mathcal{P}_i^L \in \mathbb{R}^{512}$ and joint deep prompt $\mathcal{P}_i^J = \text{cat}(\mathcal{P}_i^D, \mathcal{P}_i^B)$ into spatial vectors to ensure that both prompts have the same spatial dimensions as the imaging feature F_i^{in}. $\mathcal{P}_i^D \in \mathbb{R}^{512}$ denotes the medical prompt for deep MBs and $\mathcal{P}_i^B \in \mathbb{R}^5$ represents the estimated BP principal components.

3 Experiments and Results

3.1 Datasets

We collected 621 MRI scans as an in-house dataset and internally divided them into 430 training, 87 validation, and 104 test cases. External validation was conducted using the MICCAI VALDO 2021 challenge dataset [23], utilizing 38 of the 72 cases. We excluded 34 cases obtained using 1.5T MRI because the Meta-matching framework requires 3T MRI inputs [25].

3.2 Implementation Details

All models except the CLIP-driven-UNet [17,18] were trained using the nnU-Net framework [10,11]. The input shape of $40 \times 192 \times 160$ and the voxel resolution of $1.0 \times 1.0 \times 3.0$ mm^3 were automatically determined during the nnU-Net preprocessing steps. For the CMB-UNETR++ [16], an input shape of 192×160 and a pixel resolution of 1 mm^2 were used. The CLIP-driven-UNet was trained for 2000 epochs, employing cosine annealing with a warm restart for the initial 100 epochs, and an initial learning rate of 1e-4.

All T1-weighted MRI scans were pre-processed using HCP Prefreesurfer pipeline [3] and registered to a 1 mm^3 MNI space [12,13] to estimate blood pressure as described in Fig. 2. In addition, bias-corrected T1 MRI images were registered to the native T2* space. For the T2*-weighted MRI scans, N4 bias field correction [24] was performed, followed by skull-stripping using SynthStrip [9]. Finally, the co-registered T1- and T2*-weighted MRI images were used as multi-modal inputs.

Table 1. Quantitative results of cerebral microbleed segmentation. The best performance is highlighted in **bold**, and the second-best is underlined.

		In-house Dataset				VALDO2021			
Method	Venue	DSC↑	F_1↑	SEN↑	FP_{avg}↓	DSC↑	F_1↑	SEN↑	FP_{avg}↓
UNETR++ [22]	IEEE TMI '24	44.02	61.17	60.47	0.990	37.48	42.63	42.05	0.711
nnU-Net [10]	Nat. Methods '21	51.68	67.54	66.56	0.423	36.56	44.21	43.09	0.395
nnU-Net ResEncL [11]	MICCAI '24	47.69	66.71	64.93	0.385	41.08	47.40	49.12	0.447
CMB-UNETR++ [16]	BIBM '24	47.28	65.34	62.23	0.750	41.40	47.54	46.33	0.684
AG-nnUNet [15]	MICCAI '24	47.02	65.97	66.14	0.433	39.82	44.08	45.34	0.395
AG-nnUNet ResEncL [15]	MICCAI '24	45.85	**68.43**	67.09	0.385	44.66	48.16	46.88	**0.184**
CLIP-driven-UNet [17,18]	MedIA '24	44.05	50.36	44.46	**0.221**	34.15	40.79	39.04	0.368
UniSeg [26]	MICCAI '23	48.85	63.40	69.82	1.163	40.09	46.02	56.52	1.421
BP-nnUNet (Ours)	MICCAI '25	**57.48**	67.95	**70.97**	0.269	**49.01**	**56.98**	**57.18**	0.316

3.3 Evaluation Metrics

In addition to the Dice score (DSC), we used F_1 score, sensitivity (SEN), and the average number of false positives per subject (FP_{avg}) as the evaluation metrics. To assess whether the predicted CMB lesions were correct, we measured the distance between the centroid of each predicted lesion and its nearest ground truth by applying a distance threshold of 5 mm, following the CMB detection criteria of the MICCAI VALDO 2021 challenge [23].

3.4 Quantitative and Qualitative Results

Table 1 summarizes the performance of our proposed framework and competing methods. We compared two task-specific models, CMB-UNETR++ [16] and AG-nnUNet [15], with their respective backbones, UNETR++ [22] and nnU-Net [10]. In addition, we included nnU-Net with a residual encoder (i.e., nnU-Net ResEncL [11]) and two prompt-driven universal models, CLIP-driven-UNet [17,18] and UniSeg [26] as the SOTA baselines. All baseline methods were trained for binary CMB segmentation, as they were not originally designed for multi-label segmentation. By contrast, our BP-nnUNet was trained to predict two classes: lobar and deep MBs. For a fair comparison with the baselines, the predictions of BP-nnUNet for lobar and deep MBs were converted into binary format.

Our BP-nnUNet achieved either the best or second-best performance across all metrics. On the in-house dataset, BP-nnUNet achieved the best Dice score and sensitivity, while achieving a close second in F_1 score and FP_{avg}. On the VALDO 2021 dataset, BP-nnUNet outperformed the second-best baselines by a significant margin in the Dice and F_1 scores. This emphasizes the robustness and generalizability of our framework to unseen datasets. Fig. 4 shows the superiority of BP-nnUNet in minimizing both false positives and false negatives, compared with the most competitive baselines (AG-nnUNet ResEncL [15] and UniSeg [26]). These findings highlight the importance of designing input prompts that range from task-level to subject-level prompts.

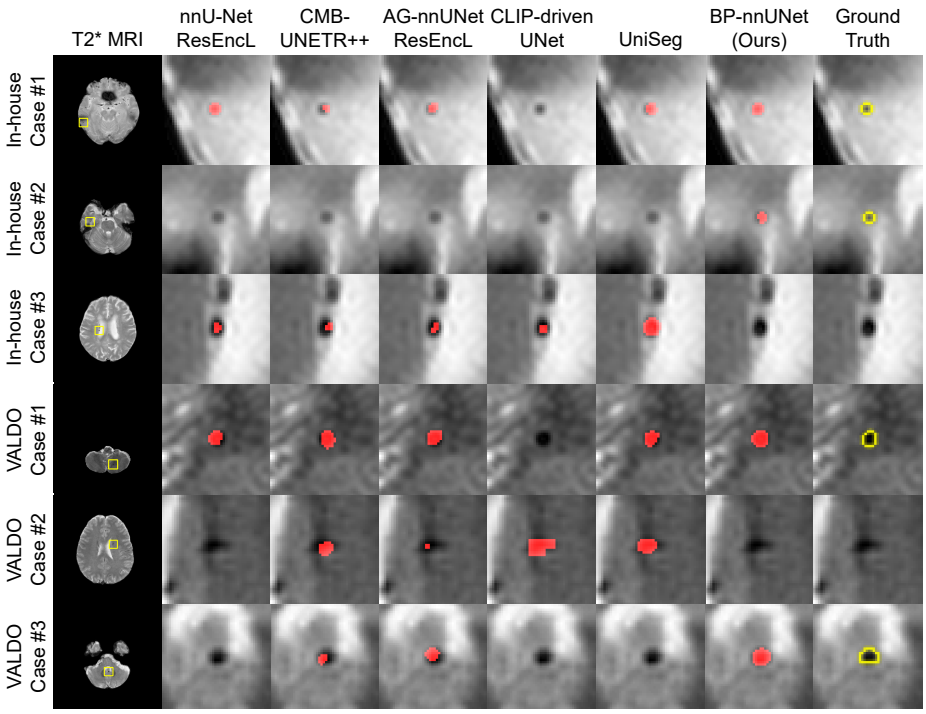

Fig. 4. Qualitative results of cerebral microbleed segmentation. Predictions are given in red while ground truth is given in yellow.

3.5 Ablation Analysis

We conducted ablation studies to evaluate the effectiveness of the BP data on nnU-Net ResEncL [11]. We considered three scenarios: *no prompt* (nnU-Net ResEncL baseline), *medical text prompt* for both lobar and deep MBs, and our *joint prompt* of medical text and BP data. Table 2 shows that our *joint prompt* consistently outperformed other scenarios in all metrics except FP_{avg}. Even the *medical text prompt* scenario outperformed all baselines in the Dice and F_1 scores, particularly on the unseen datasets, demonstrating the superiority of our CMB subtyping approach in designing input prompts. Furthermore, incorporating case-level features, such as BP, significantly enhances the overall performance of our BP-nnUNet, at the minor cost of only a slight increase in false positives on unseen datasets. Hence, BP plays a crucial role in boosting the CMB segmentation performance.

We also compared our anatomically-aware joint prompt fusion module with other fusion methods, as shown in Table 3. Specifically, *cross-attention* refers to the application of cross-attention between imaging features and joint prompt embeddings. *Single prompt fusion* refers to the use of a medical text prompt for a single class of CMB, whereas *joint prompt fusion* incorporates subtype-level

Table 2. Ablation results on input prompts. BP denotes using blood pressure prompt. The best performance is highlighted in **bold**.

Prompt		In-house Dataset				VALDO2021			
Text	BP	DSC↑	F_1↑	SEN↑	FP_{avg}↓	DSC↑	F_1↑	SEN↑	FP_{avg}↓
		47.69	66.71	64.93	0.385	41.08	47.40	49.12	0.447
✓		46.55	67.55	67.05	0.442	45.41	55.13	53.29	**0.237**
✓	✓	**57.48**	**67.95**	**70.97**	**0.269**	**49.01**	**56.98**	**57.18**	0.316

Table 3. Ablation results on fusion methods. Single prompt denotes a common task-level prompt for cerebral microbleeds. Joint prompt denotes subtype-level prompts for both lobar and deep microbleeds. The best performance is highlighted in **bold**.

	In-house Dataset				VALDO2021			
Fusion Method	DSC↑	F_1↑	SEN↑	FP_{avg}↓	DSC↑	F_1↑	SEN↑	FP_{avg}↓
Cross-attention	44.52	66.90	65.06	0.337	45.80	51.07	49.84	**0.211**
Single Prompt Fusion	49.40	65.87	63.83	0.327	41.51	42.22	42.65	0.368
Joint Prompt Fusion	**57.48**	**67.95**	**70.97**	**0.269**	**49.01**	**56.98**	**57.18**	0.316

prompts for both lobar and deep MBs. Table 3 shows the superiority of *joint prompt fusion* on both the in-house and external datasets. Our ablation study highlights the importance of CMB subtyping, which enables prompt fusion in a joint fashion. By contrast, adopting a single task-level prompt was particularly unsuccessful on unseen datasets. This aligns with the poor performance of the CLIP-driven-UNet [17,18], which was also trained with a task-level prompt derived from a pre-trained CLIP model. Therefore, applying subtype-level prompts is essential for enhancing the CMB segmentation performance.

4 Conclusion

In this paper, we propose a joint prompt of medical text and BP data to fully leverage the clinical characteristics of CMB subtypes. By enriching medical text prompts from a task level to a subtype level and leveraging pathological details through BP prompt, our framework achieved promising performance compared with SOTA CMB segmentation methods and universal prompt-driven models. Ablation studies revealed the effectiveness of our three design components: CMB subtyping, incorporation of BP, and joint prompt fusion, all of which significantly enhanced the performance on unseen datasets. Specifically, employing medical text prompts based on CMB subtypes outperformed the baseline methods on unseen datasets. Leveraging BP data effectively exploited subject-level features, significantly boosting the CMB segmentation performance. Lastly, our joint prompt fusion effectively utilized both BP and subtype-level prompts, delivering the best performance via BP-nnUNet.

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 the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2025-00518533), K-Brain Project of the National Research Foundation (NRF), funded by the Korean government (MSIT) (RS-2023-00265393), National Research Foundation (RS-2024-00408040), AI Graduate School Support Program (Sungkyunkwan University) (RS-2019-II190421), ICT Creative Consilience program (IITP-2025-RS-2020-II201821), 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. Alfaro-Almagro, F., Jenkinson, M., Bangerter, N.K., Andersson, J.L., Griffanti, L., Douaud, G., Sotiropoulos, S.N., Jbabdi, S., Hernandez-Fernandez, M., Vallee, E., et al.: Image processing and quality control for the first 10,000 brain imaging datasets from uk biobank. Neuroimage **166**, 400–424 (2018)
2. Chen, P., An, L., Wulan, N., Zhang, C., Zhang, S., Ooi, L.Q.R., Kong, R., Chen, J., Wu, J., Chopra, S., et al.: Multilayer meta-matching: Translating phenotypic prediction models from multiple datasets to small data. Imaging Neuroscience **2**, 1–22 (2024)
3. Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., et al.: The minimal preprocessing pipelines for the human connectome project. Neuroimage **80**, 105–124 (2013)
4. Greenberg, S.M., Nandigam, R.K., Delgado, P., Betensky, R.A., Rosand, J., Viswanathan, A., Frosch, M.P., Smith, E.E.: Microbleeds versus macrobleeds: evidence for distinct entities. Stroke **40**(7), 2382–2386 (2009)
5. 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)
6. 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)
7. He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R.: Masked autoencoders are scalable vision learners. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 15979–15988 (2022)
8. He, T., An, L., Chen, P., Chen, J., Feng, J., Bzdok, D., Holmes, A.J., Eickhoff, S.B., Yeo, B.T.: Meta-matching as a simple framework to translate phenotypic predictive models from big to small data. Nature neuroscience **25**(6), 795–804 (2022)
9. Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B., Hoffmann, M.: Synthstrip: skull-stripping for any brain image. NeuroImage **260**, 119474 (2022)
10. 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)

11. Isensee, F., Wald, T., Ulrich, C., Baumgartner, M., Roy, S., Maier-Hein, K., Jaeger, P.F.: nnu-net revisited: A call for rigorous validation in 3d medical image segmentation. In: 27th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). pp. 488–498 (2024). https://doi.org/10.1007/978-3-031-72114-4_47
12. Jenkinson, M., Bannister, P., Brady, M., Smith, S.: Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage **17**(2), 825–841 (2002)
13. Jenkinson, M., Smith, S.: A global optimisation method for robust affine registration of brain images. Medical image analysis **5**(2), 143–156 (2001)
14. Kwon, J., Kim, J., Park, H.: Leveraging segmentation-guided spatial feature embedding for overall survival prediction in glioblastoma with multimodal magnetic resonance imaging. Computer Methods and Programs in Biomedicine **255**, 108338 (2024)
15. Kwon, J., Seo, S.W., Park, H.: Anatomically-guided segmentation of cerebral microbleeds in t1-weighted and t2*-weighted mri. In: 27th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). pp. 24–33 (2024). https://doi.org/10.1007/978-3-031-72069-7_3
16. Kwon, J., Seo, S.W., Park, H.: Enhancing cerebral microbleed segmentation with pretrained unetr++. In: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 3372–3377. IEEE (2024). https://doi.org/10.1109/BIBM62325.2024.10822393
17. Liu, J., Zhang, Y., Chen, J.N., Xiao, J., Lu, Y., Landman, B.A., Yuan, Y., Yuille, A., Tang, Y., Zhou, Z.: Clip-driven universal model for organ segmentation and tumor detection. In: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 21095–21107 (2023)
18. Liu, J., Zhang, Y., Wang, K., Yavuz, M.C., Chen, X., Yuan, Y., Li, H., Yang, Y., Yuille, A., Tang, Y., et al.: Universal and extensible language-vision models for organ segmentation and tumor detection from abdominal computed tomography. Medical image analysis **97**, 103226 (2024)
19. 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)
20. Peng, H., Gong, W., Beckmann, C.F., Vedaldi, A., Smith, S.M.: Accurate brain age prediction with lightweight deep neural networks. Medical image analysis **68**, 101871 (2021)
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. Shaker, A.M., Maaz, M., Rasheed, H., Khan, S., Yang, M.H., Khan, F.S.: Unetr++: Delving into efficient and accurate 3d medical image segmentation. IEEE Transactions on Medical Imaging **43**(9), 3377–3390 (2024)
23. 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)
24. 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)
25. Wulan, N., An, L., Zhang, C., Kong, R., Chen, P., Bzdok, D., Eickhoff, S.B., Holmes, A.J., Yeo, B.T.: Translating phenotypic prediction models from big to

small anatomical mri data using meta-matching. Imaging Neuroscience **2**, 1–21 (2024)
26. Ye, Y., Xie, Y., Zhang, J., Chen, Z., Xia, Y.: Uniseg: A prompt-driven universal segmentation model as well as a strong representation learner. In: 26th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). pp. 508–518 (2023). https://doi.org/10.1007/978-3-031-43898-1_49
27. Zhang, S., Xu, Y., Usuyama, N., Xu, H., Bagga, J., Tinn, R., Preston, S., Rao, R., Wei, M., Valluri, N., Wong, C., Tupini, A., Wang, Y., Mazzola, M., Shukla, S., Liden, L., Gao, J., Crabtree, A., Piening, B., Bifulco, C., Lungren, M.P., Naumann, T., Wang, S., Poon, H.: A multimodal biomedical foundation model trained from fifteen million image–text pairs. NEJM AI **2**(1) (2024)