Synthetic training for long-tail haemorrhagic lesion segmentation in data-scarce settings

Yuan Cao[1⋆][0009−0007−4323−0737], Sumeet Dash[1⋆], Antonia Zachariadis[2], Stefanie Schreiber[2,3][0000−0003−4439−4374], Katja Neumann[2], and Jose Bernal[1,4][0000−0003−3167−5134]

[1] Department Artificial Intelligence in Biomedical Engineering (AIBE), Faculty of Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Nürnberger Strasse 74, 91052 Erlangen, Germany
{yuan.cao, sumeet.dash, jose.bernal}@fau.de
[2] Otto von Guericke University Magdeburg, Medical Faculty, Department of Neurology, Leipziger Straße 44, 39120 Magdeburg, Germany
antonia.zachariadis@st.ovgu.de
{stefanie.schreiber, katja.neumann}@med.ovgu.de
[3] German Centre for Neurodegenerative Diseases (DZNE), Leipziger Straße 44, 39120 Magdeburg, Germany
[4] Institute for Neuroscience and Cardiovascular Research, Row Fogo Centre for Research into Ageing and The Brain, Department of Neuroimaging Sciences, The University of Edinburgh, 49 Little France Crescent, Edinburgh EH16 4TJ, UK

Abstract. Cerebral microbleeds (CMBs) and cortical superficial siderosis (cSS) are imaging markers of cerebral small vessel disease, but their automated segmentation is limited by the scarcity of positive cases and voxel-level annotations. We propose a synthetic training framework for long-tail haemorrhagic lesion segmentation that requires no real lesion annotations for training and leverages radiological description of the lesions. Starting from anatomical brain parcellations, the framework applies spatial augmentation and voxel resampling, procedurally inserts cSS and CMB labels using clinical priors on lesion location and morphology, and synthesises images through randomised intensity assignment, blurring, and Rician noise simulation. Models were trained on dynamically generated image–label pairs and evaluated against manual delineations in 10 cSS cases and 13 CMB cases. The proposed configurations outperformed classical filter baselines. For cSS, the hypointensity-constrained model achieved higher AUPRC and AUROC than the Frangi filter (AUPRC: 0.284 vs 0.083; AUROC: 0.907 vs 0.731). For CMBs, explicit synthesis of blood vessels as lesion mimics improved performance over the classical baseline (AUPRC: 0.538 vs 0.004; AUROC: 0.999 vs 0.968). These results support our proposal as a feasible strategy for data-scarce haemorrhagic lesion segmentation.

Keywords: Synthetic training · Cortical superficial siderosis · Cerebral microbleeds · Segmentation · Cerebral small vessel disease · Brain MRI

⋆ Yuan Cao and Sumeet Dash contributed equally to this work.

1 Introduction

Cerebral microbleeds (CMBs) and cortical superficial siderosis (cSS) are magnetic resonance imaging (MRI) markers of cerebral small vessel disease (CSVD), a major contributor to stroke, vascular cognitive impairment, dementia, and haemorrhagic adverse events during anti-amyloid therapies [8, 5, 16, 6, 14]. Accurate quantification of CMB and cSS burden is therefore important for characterising CSVD severity, stratifying haemorrhagic and dementia risk, assessing treatment eligibility, and monitoring disease progression and treatment safety.

When large, heterogeneous annotated datasets are available, standard supervised segmentation pipelines such as nnU-Net can achieve strong lesion-segmentation performance [12]. This premise, however, breaks down for long-tail haemorrhagic lesion-segmentation tasks such as CMB and cSS segmentation. Positive cases are uncommon even in clinically relevant cohorts, with reported prevalences below 3% for cSS [15] and below 20% for CMBs in memory-clinic settings [7]. Moreover, lesion voxels are extremely sparse, and lesion burden can range from a few isolated voxels to disseminated involvement across the whole brain [8]. Building representative training datasets therefore requires targeted expert case identification and voxel-level annotation across diverse cohorts, both of which are laborious and resource-intensive. For CMBs and cSS, such large, diverse, publicly available datasets remain unavailable, motivating alternative training strategies for data-scarce, long-tail lesion segmentation.

Synthetic training addresses this bottleneck by generating image–label pairs rather than requiring manual annotation for every training example. This strategy has proven effective when the target structures are already available as label maps, such as anatomical regions, tissue classes, or existing lesion masks [10, 11, 13, 1, 2, 4]. For CMBs and cSS, however, voxel-level lesion labels are precisely missing, limiting the direct applicability of standard synthetic training frameworks.

In this paper, we propose to extend synthetic training to data-scarce haemorrhagic lesion segmentation by procedurally generating the missing CMB and cSS labels themselves using clinical and imaging priors on lesion morphology, anatomical distribution, MRI appearance, and acquisition variability. Evaluation against manual delineations across two datasets supports this approach for long-tail haemorrhagic lesion segmentation in data-scarce settings.

2 Methods

The following sections describe the proposed synthetic data generation framework (Fig. 1). In brief, anatomical parcellations are first spatially augmented to increase anatomical and voxel-size variability. cSS and CMB labels are then inserted using lesion-specific anatomical and morphological priors. The resulting lesion-augmented maps are converted into synthetic images through randomised intensity assignment, blurring, and noise simulation, and finally paired with binary lesion masks for network training.

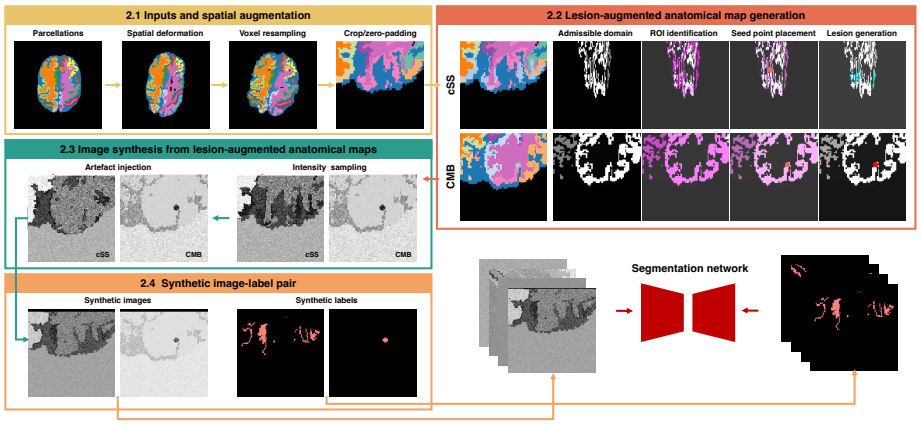

Fig. 1. Proposed synthetic image generation pipeline for long-tail haemorrhagic lesion segmentation.

2.1 Inputs and spatial augmentation

The data generation workflow starts from a pool of brain parcellations $\{S_n\}_{n=1}^{N}$, where each parcellation is a label map $S_n : \Omega_n \to \{1, \ldots, K\}$. Here, $\Omega_n \subset \mathbb{Z}^3$ denotes the discrete voxel lattice of subject n, and K is the number of anatomical classes before adding cSS/CMB labels. Although these maps contain no haemorrhagic lesion annotations, they provide the anatomical context required to place cSS and CMBs in clinically plausible locations.

At each training iteration, a parcellation S_i is sampled at random and spatially augmented to increase anatomical variability and simulate native-space acquisition differences. The augmentation comprises a randomly parameterised affine transformation $\phi_{\text{aff}}(\cdot;\theta_{\text{aff}})$, a nonlinear elastic deformation $\phi_{\text{nonlin}}(\cdot;\theta_{\text{nonlin}})$, and voxel-size resampling $\mathcal{R}_{\mathbf{r}_{\text{res}}}$, where θ_{aff} are the affine parameters, θ_{nonlin} the deformation parameters, and $\mathbf{r}_{\text{res}}$ the target voxel size.

Because transformation and resampling may change the lattice dimensions, a cropping or zero-padding operator $\mathcal{M}_\Omega$ is applied to enforce a consistent output lattice $\Omega \subset \mathbb{Z}^3$. The augmented parcellation is therefore defined as

$$S_i^{\text{aug}} = \mathcal{M}_\Omega \mathcal{R}_{\mathbf{r}_{\text{res}}} \left[S_i \circ (\phi_{\text{aff}}(\cdot;\theta_{\text{aff}}) \circ \phi_{\text{nonlin}}(\cdot;\theta_{\text{nonlin}})) \right], \; S_i^{\text{aug}} : \Omega \to \{1, \ldots, K\}. \tag{1}$$

2.2 Lesion-augmented anatomical map generation

The STRIVE criteria [8] define CMBs as small, quasi-spherical blood-product deposits within the brain parenchyma, whereas cSS reflects superficial blood-product deposition along the cortical surface and adjacent subarachnoid space. Lesion labels are therefore generated by combining lesion-specific priors on morphology and anatomical location with the augmented anatomical map S_i^{aug}.

Let $\mathcal{L}_{\mathrm{GM}} \subset \{1,\dots,K\}$ and $\mathcal{L}_{\mathrm{WM}} \subset \{1,\dots,K\}$ denote the grey matter and white matter label sets, respectively, and let $l_{\mathrm{CSF}} \in \{1,\dots,K\}$ denote the CSF label. For each lesion type $t \in \{\mathrm{cSS}, \mathrm{CMB}\}$, an admissible domain $\mathcal{D}_i^t$ is first defined, seed points are sampled within this domain, and the resulting lesion mask Ω_i^t is inserted into the anatomical label map.

For cSS, the admissible domain is the CSF–cortical grey matter interface:

$$\mathcal{D}_i^{\mathrm{cSS}} = \Gamma_i^{\mathrm{sur}} = (\{x \in \Omega \mid S_i^{\mathrm{aug}}(x) = l_{\mathrm{CSF}}\} \oplus B) \cap \{x \in \Omega \mid S_i^{\mathrm{aug}}(x) \in \mathcal{L}_{\mathrm{GM}}\}, \quad (2)$$

where $\oplus$ denotes morphological dilation with structuring element B. For CMBs, the admissible domain is restricted to parenchymal tissue:

$$\mathcal{D}_i^{\mathrm{CMB}} = \Omega_i^{\mathrm{par}} = \{x \in \Omega \mid S_i^{\mathrm{aug}}(x) \in \mathcal{L}_{\mathrm{par}}\}, \qquad \mathcal{L}_{\mathrm{par}} = \mathcal{L}_{\mathrm{GM}} \cup \mathcal{L}_{\mathrm{WM}}. \quad (3)$$

Seed points $\mathcal{S}_i^t = \{s_1^t, \dots, s_{M^t}^t\} \subset \mathcal{D}_i^t$ are sampled with $M^t \sim \mathcal{U}(M_{\min}^t, M_{\max}^t)$, and each seed is assigned a radius $R_j^t \sim \mathcal{U}(R_{\min}^t, R_{\max}^t)$. Lesion geometry is then generated according to the lesion type:

$$\Omega_i^{\mathrm{cSS}} = \bigcup_{j=1}^{M^{\mathrm{cSS}}} \{x \in \Gamma_i^{\mathrm{sur}} \mid d_{\Gamma,\boldsymbol{\rho}}(x, s_j^{\mathrm{cSS}}) \le R_j^{\mathrm{cSS}}\}, \quad (4)$$

for surface-constrained cSS, where $d_{\Gamma,\boldsymbol{\rho}}$ denotes distance along the cortical surface, and

$$\Omega_i^{\mathrm{CMB}} = \bigcup_{j=1}^{M^{\mathrm{CMB}}} \left\{x \in \Omega_i^{\mathrm{par}} \mid \left\|\boldsymbol{\rho} \odot (x - s_j^{\mathrm{CMB}})\right\|_2 \le R_j^{\mathrm{CMB}}\right\}, \quad (5)$$

for approximately spherical CMBs in physical space, where $\boldsymbol{\rho}$ denotes voxel size.

Finally, the simulated lesion label is inserted into the augmented parcellation:

$$S_{i,t}^{\mathrm{final}}(x) = \begin{cases} l_t, & \text{if } x \in \Omega_i^t, \\ S_i^{\mathrm{aug}}(x), & \text{otherwise,} \end{cases} \qquad S_{i,t}^{\mathrm{final}} : \Omega \to \{1,\dots,K\} \cup \{l_t\}. \quad (6)$$

2.3 Image synthesis from lesion-augmented anatomical maps

The intensity model assigns each label l an intensity parameter μ_l. For anatomical labels $k \in \mathcal{L}_{\mathrm{anat}} = \{1,\dots,K\}$, intensities are sampled as $\mu_k \sim \mathcal{U}(0,1)$. For haemorrhagic lesion labels, intensities are constrained by an imaging prior. On T2*-weighted imaging and SWI, blood-breakdown products cause local signal loss, making CMBs and cSS hypointense relative to neighbouring tissues [5, 8]. Therefore, for lesion labels $h \in \mathcal{L}_{\mathrm{hem}} = \{l_{\mathrm{CMB}}, l_{\mathrm{cSS}}\}$, intensities are sampled under the constraint

$$\max_{h \in \mathcal{L}_{\mathrm{hem}}} \mu_h < \min_{k \in \mathcal{L}_{\mathrm{anat}}} \mu_k. \quad (7)$$

The initial synthetic image is then obtained by replacing each voxel label in the lesion-augmented map with its sampled intensity $I_0(x) = \mu_{S_i^{\mathrm{final}}(x)},\ x \in \Omega$.

To mimic partial-volume effects and acquisition noise, the initial synthetic image I_0 is first smoothed with a Gaussian kernel $G_{\sigma_{\text{blur}}}$ and then corrupted with Rician noise. Noise is simulated by adding independent complex Gaussian components to the k-space representation, with variance determined by a randomly sampled $\text{SNR}_{\text{dB}} \sim \mathcal{U}(\text{SNR}_{\min}, \text{SNR}_{\max})$. The resulting image is obtained by inverse Fourier transformation and magnitude reconstruction:

$$I_{\text{final}}(x) = \left| \mathcal{F}^{-1} \left(\mathcal{F}(I_0 * G_{\sigma_{\text{blur}}}) + \eta_{\text{real}} + i\eta_{\text{imag}} \right)(x) \right|, \tag{8}$$

where $\eta_{\text{real}}, \eta_{\text{imag}} \sim \mathcal{N}(0, \sigma^2_{\text{noise}})$ are independent Gaussian noise components, with $\sigma_{\text{noise}} = \mu_{\text{signal}} / 10^{\text{SNR}_{\text{dB}}/20}$.

2.4 Final synthetic image-label pair

Training synthetic image–mask pairs (I_{final}, Y) are constructed by using I_{final} as the input image and binarising S_i^{final} to isolate the target cSS/CMB label.

3 Training and evaluation

3.1 Training

Brain parcellations. We used a pool of 840 brain parcellations derived from Alzheimer's Disease Neuroimaging Initiative (ADNI) scans. Parcellations were generated with SynthSeg [1] and SIAM [18], which provides labels to guide cSS and CMB placement, respectively.

Generation parameters. Brain parcellations were spatially augmented with affine transformations (scaling ±20%, rotation ±15°, shearing ±0.012) and nonlinear elastic deformations (displacement-field std 3.0, control-point scale 0.0625). To mimic acquisition variability, volumes were randomly resampled to isotropic resolutions of 0.5–1.2 mm or anisotropic resolutions with 0.5–1.2 mm in-plane and 1.2–4.0 mm through-plane spacing. For cSS, 3–20 lesions were generated from sampled seeds using geodesic propagation, with a base cost of 1.0, noise weight of 0.25, and $\boldsymbol{\rho} \sim \mathcal{U}(10, 30)$ voxels. For CMB, radii were sampled uniformly from 1–5 mm, with 1–4 lesions per crop. In both cases, SNR_{dB} was drawn from a uniform distribution $\mathcal{U}(10\ dB, 40\ dB)$.

Segmentation network. We used a five-level 3D U-Net backbone. Each block comprised two convolutional layers with batch normalisation and ELU activations. The encoder used max-pooling for downsampling, the decoder used upsampling, and skip connections preserved local spatial detail. Voxel-wise predictions were produced with a final softmax layer. The model was implemented in PyTorch/MONAI and trained for 70 epochs, with 1,000 on-the-fly synthesised image–label pairs per epoch and a batch size of 2. Optimisation used Adam with a learning rate of 10^{-5}. To address severe class imbalance, training minimised a soft Dice loss with a squared denominator and dynamic inverse-volume weighting [1, 3]. Training took approximately 24 hours on an NVIDIA A100 GPU.

3.2 Evaluation

Test cases and ground truth. For cSS, evaluation used an in-house test set of 10 patients with confirmed cSS (6 females; median age 74.5 [IQR 67–78] years) scanned with susceptibility- or T2*-weighted MRI. Lesions were manually delineated by a trained investigator and validated by an experienced neuroradiologist. For CMBs, evaluation used 13 T2*-weighted scans from the ALFA cohort made available through the MICCAI 2021 VALDO Challenge [17], under the corresponding ethics approval and data use agreements.

Ablation experiments. To quantify the effect of key generation choices, we performed ablation experiments. First, we compared fully randomised label intensities with the proposed intensity prior that enforces cSS and CMB hypointensity, testing whether preserving haemorrhagic signal dropout improves segmentation beyond anatomical and morphological randomisation alone. Second, we compared image-only input with input augmented by a Laplacian filter response, testing whether an explicit cue for dark lesion-like structures improves learning. Under the adopted sign convention, hypointense structures yield positive Laplacian responses, highlighting candidate CMB and cSS patterns while preserving the original synthetic image information. Third, we also evaluated lesion-specific refinements motivated by observed failure modes. For CMBs, false positives frequently occurred along blood vessels; we therefore tested whether including blood vessel labels from SIAM during synthesis improved robustness to hypointense vessel-like structures. For cSS, errors tended to occur at the CSF–cortical boundary. We therefore tested whether a two-class target formulation improved separation between cSS and CSF.

Comparison against other methods. For cSS, no established method provides fully automated segmentation. The closest existing approach is a semi-automated Frangi-filter-based method [9], which we therefore used as comparator. For CMBs, we also used the Frangi as the corresponding classical baseline.

Evaluation metrics. Segmentation performance was evaluated using threshold-free metrics: area under the precision–recall curve (AUPRC) and area under the receiver operating characteristic curve (AUROC). These metrics allow comparison with classical non-probabilistic baselines, such as the Frangi filter.

4 Results and discussion

Results are shown in Figs. 2–3 and discussed in the following sections.

Random intensities vs intensity prior. Fully randomised intensity sampling yielded substantially lower performance than the proposed hypointensity prior for both cSS and CMB segmentation, especially for AUPRC (cSS: 0.086 [IQR 0.028–0.152] vs 0.284 [IQR 0.124–0.363]; CMB: 0.024 [IQR 0.002–0.446] vs 0.261 [IQR 0.056–0.431]), while AUROC remained largely unchanged. This suggests that enforcing haemorrhagic-lesion hypointensity improves precision in the sparse lesion class by ranking true lesion voxels higher among lesion-like candidates and reducing false positives. In contrast, fully randomised intensities make the synthetic task less aligned with blood-sensitive MRI.

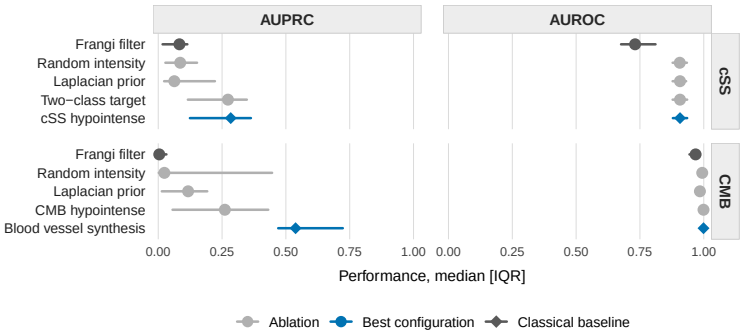

Fig. 2. Quantitative evaluation of framework components for cSS and CMB segmentation. Forest plots show median AUPRC and AUROC with interquartile ranges for each approach. The proposed hypointensity prior improved performance for both lesion types, while explicit blood-vessel synthesis led to the best CMB performance.

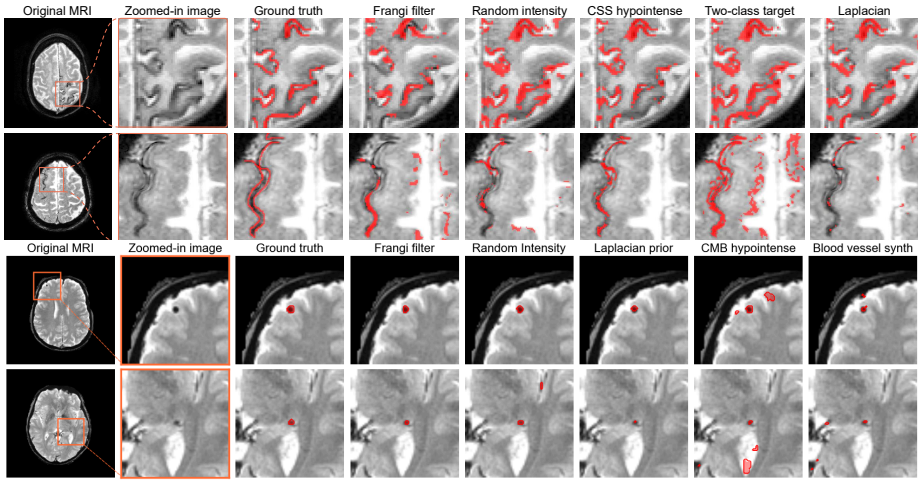

Fig. 3. Visual comparison of segmentation performance. Columns show, from left to right, the original MRI scan, zoomed-in view, manual ground-truth annotation, Frangi filter output, and outputs from the four ablation configurations: random intensity sampling, Laplacian prior, cSS/CMB hypointensity constraint, and lesion-specific refinement. The lesion-specific refinement corresponds to a two-class target for cSS and blood-vessel synthesis for CMBs.

Image prior only vs image and Laplacian prior. Although the Laplacian channel was intended to provide an explicit cue for local hypointense structures, it also enhanced non-lesion features, including vessel boundaries and sulcal interfaces. Consequently, adding the Laplacian prior degraded performance relative to the image prior alone for both lesion types (cSS AUPRC: 0.063 [IQR 0.023–0.222] vs 0.284 [IQR 0.124–0.363]; CMB AUPRC: 0.117 [IQR 0.014–0.192] vs 0.261 [IQR 0.056–0.431]). These findings suggest that the constrained synthetic images already encoded sufficient intensity and morphological information, whereas the Laplacian channel introduced a less specific representation that impaired transfer to real data.

cSS-specific refinements. For cSS, visual inspection showed that errors frequently occurred at the CSF–cortical boundary, particularly when cSS was adjacent to sulcal CSF. To address this failure mode, we tested an alternative target formulation in which CSF was modelled explicitly as a separate class, alongside background and cSS, rather than training only a binary cSS-versus-background model. However, this modification produced little numerical change. AUPRC was similar for the binary cSS target and the explicit-CSF formulation (0.284 [IQR 0.124–0.363] vs 0.273 [IQR 0.116–0.347]). This suggests that most boundary errors were not primarily caused by the absence of an explicit CSF class, but may instead reflect the intrinsic ambiguity of thin hypointense structures at the cortical surface, partial-volume effects, and the limited spatial separation between cSS and adjacent CSF.

CMB-specific refinements. For CMBs, visual inspection showed that false positives often occurred along cerebral blood vessels, which can also appear hypointense on blood-sensitive MRI. To address this failure mode, we tested whether explicit vessel modelling during synthesis improved robustness. Because SynthSeg does not provide vascular labels, SIAM parcellations were used to introduce vessels as a separate anatomical class in the synthetic label maps. This CMB-specific refinement substantially improved performance, increasing AUPRC from 0.261 [IQR 0.056–0.431] to 0.538 [IQR 0.470–0.723]. This suggests that explicitly modelling relevant lesion mimics during synthetic training is important for reducing false positives.

Comparison against other methods Compared with the classical filter baseline, the best synthetic-training configurations showed higher median AUPRC and AUROC for both lesion types. For cSS, the hypointensity-constrained model achieved higher AUPRC than the Frangi filter (0.284 [IQR 0.124–0.363] vs 0.083 [IQR 0.017–0.114]) and higher AUROC (0.907 [IQR 0.879–0.935] vs 0.731 [IQR 0.676–0.811]). For CMBs, the model trained with blood-vessel synthesis achieved higher AUPRC than the classical filter baseline (0.538 [IQR 0.470–0.723] vs 0.004 [IQR 0.002–0.032]) and higher AUROC (0.999 [IQR 0.997–1.000] vs 0.968 [IQR 0.944–0.978]). Together, these results indicate better discrimination of the sparse positive class, with a more favourable balance between true-positive lesion voxels and false-positive detections, as reflected by AUPRC. The higher AUROC further suggests improved lesion-versus-background rank-

ing across thresholds, although this metric is less sensitive to false positives in highly imbalanced voxel-wise segmentation.

5 Conclusion

We introduced a synthetic training strategy for long-tail haemorrhagic lesion segmentation and demonstrated its feasibility for CMB and cSS quantification. By procedurally generating lesion labels from anatomical and imaging priors, the proposed framework enabled model training without real voxel-level lesion annotations. Experimental results showed agreement with manual delineations across two datasets. Future work will evaluate the approach in larger, more heterogeneous cohorts and extend it from segmentation to lesion detection.

Acknowledgments. None.

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

References

1. Billot, B., Greve, D.N., Puonti, O., Thielscher, A., Van Leemput, K., Fischl, B., Dalca, A.V., Iglesias, J.E., et al.: Synthseg: Segmentation of brain mri scans of any contrast and resolution without retraining. Medical image analysis **86**, 102789 (2023)
2. Billot, B., Magdamo, C., Cheng, Y., Arnold, S.E., Das, S., Iglesias, J.E.: Robust machine learning segmentation for large-scale analysis of heterogeneous clinical brain mri datasets. Proceedings of the National Academy of Sciences **120**(9), e2216399120 (2023)
3. Bitar, L., Díaz, M., Duarte Coello, R., Valdés-Hernández, M.d., Mattern, H., Neumann, K., Pfister, M., Beck, C., Mai, H.T., Fuchs, E., Tang, S., Tosun, D., Besteher, B., Rocktäschel, T., Reuken, P.A., Stallmach, A., Opel, N., Gaser, C., Walter, M., Dörner, M., Arndt, P., Behme, D., Piechowiak, C., Lading, Y., Müller, P., Braun-Dullaeus, R., Meuth, S.G., Alzheimer's Disease Neuroimaging Initiative, Wardlaw, J.M., Schreiber, S., Trujillo, M., Düzel, E., Ziegler, G., Bernal, J.: DRIPS: Domain randomisation for image-based perivascular spaces segmentation. medRxiv preprint 2025.10.22.25337423 (2025)
4. Chalcroft, L., Pappas, I., Price, C.J., Ashburner, J.: Synthetic data for robust stroke segmentation. Machine Learning for Biomedical Imaging **3**(August 2025), 317–346 (Aug 2025). https://doi.org/10.59275/j.melba.2025-f3g6, http://dx.doi.org/10.59275/j.melba.2025-f3g6
5. Charidimou, A., Boulouis, G., Frosch, M.P., Baron, J.C., Pasi, M., Albucher, J.F., Banerjee, G., Barbato, C., Bonneville, F., Brandner, S., et al.: The boston criteria version 2.0 for cerebral amyloid angiopathy: a multicentre, retrospective, mri–neuropathology diagnostic accuracy study. The Lancet Neurology **21**(8), 714–725 (2022)

6. Claus, J.J., Vom Hofe, I., van Ijlzinga Veenstra, A., Licher, S., Seelaar, H., de Jong, F.J., Neitzel, J., Vernooij, M.W., Ikram, M.A., Wolters, F.J.: Generalizability of trial criteria on amyloid-lowering therapy against alzheimer's disease to individuals with mild cognitive impairment or early alzheimer's disease in the general population. European Journal of Epidemiology **40**(3), 327–337 (2025)
7. Cordonnier, C., Van der Flier, W., Sluimer, J., Leys, D., Barkhof, F., Scheltens, P.: Prevalence and severity of microbleeds in a memory clinic setting. Neurology **66**(9), 1356–1360 (2006)
8. Duering, M., Biessels, G.J., Brodtmann, A., Chen, C., Cordonnier, C., de Leeuw, F.E., Debette, S., Frayne, R., Jouvent, E., Rost, N.S., et al.: Neuroimaging standards for research into small vessel disease—advances since 2013. The Lancet Neurology **22**(7), 602–618 (2023)
9. van Harten, T., Koemans, E., Voigt, S., Rasing, I., van Osch, M., van Walderveen, M., Wermer, M.: Quantitative measurement of cortical superficial siderosis in cerebral amyloid angiopathy. NeuroImage: Clinical **38**, 103447 (2023)
10. Hoffmann, M., Hoopes, A., Greve, D.N., Fischl, B., Dalca, A.V.: Anatomy-aware and acquisition-agnostic joint registration with synthmorph. Imaging Neuroscience **2**, imag–2 (2024)
11. Hoopes, A., Mora, J.S., Dalca, A.V., Fischl, B., Hoffmann, M.: Synthstrip: skull-stripping for any brain image. NeuroImage **260**, 119474 (2022)
12. 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)
13. Laso, P., Cerri, S., Sorby-Adams, A., Guo, J., Mateen, F., Goebl, P., Wu, J., Liu, P., Li, H.B., Young, S.I., et al.: Quantifying white matter hyperintensity and brain volumes in heterogeneous clinical and low-field portable mri. In: 2024 IEEE International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2024)
14. Lohner, V., Badhwar, A., Detcheverry, F.E., García, C.L., Gellersen, H.M., Khodakarami, Z., Lattmann, R., Li, R., Low, A., Mazo, C., et al.: Machine learning applications in vascular neuroimaging for the diagnosis and prognosis of cognitive impairment and dementia: a systematic review and meta-analysis. Alzheimer's research & therapy **17**(1), 183 (2025)
15. Shams, S., Martola, J., Charidimou, A., Cavallin, L., Granberg, T., Shams, M., Forslin, Y., Aspelin, P., Kristoffersen-Wiberg, M., Wahlund, L.O.: Cortical superficial siderosis: prevalence and biomarker profile in a memory clinic population. Neurology **87**(11), 1110–1117 (2016)
16. Sin, M.K., Zamrini, E., Ahmed, A., Nho, K., Hajjar, I.: Anti-amyloid therapy, ad, and aria: untangling the role of caa. Journal of clinical medicine **12**(21), 6792 (2023)
17. 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)
18. Valabregue, R., Khemir, I., Badinet, E., Rousseau, F., Auzias, G., Dorent, R.: SIAM: Head and brain MRI segmentation from few high-quality templates via synthetic training. arXiv preprint arXiv:2605.02737 (2026)