SELF-SUPERVISED FEW-SHOT LEARNING FOR ISCHEMIC STROKE LESION SEGMENTATION

Luca Tomasetti[1]   *Stine Hansen*[2]   *Mahdieh Khanmohammadi*[1]   *Kjersti Engan*[1]
Liv Jorunn Høllesli[1,3]   *Kathinka Dæhli Kurz*[1,3]   *Michael Kampffmeyer*[2]

[1] Department of Electrical Engineering and Computer Science, University of Stavanger, Norway
[2] Department of Physics and Technology, UiT The Arctic University of Norway, Norway
[3] Department of Radiology, Stavanger University Hospital, Norway

ABSTRACT

Precise ischemic lesion segmentation plays an essential role in improving diagnosis and treatment planning for ischemic stroke, one of the prevalent diseases with the highest mortality rate. While numerous deep neural network approaches have recently been proposed to tackle this problem, these methods require large amounts of annotated regions during training, which can be impractical in the medical domain where annotated data is scarce. As a remedy, we present a prototypical few-shot segmentation approach for ischemic lesion segmentation using only one annotated sample during training. The proposed approach leverages a novel self-supervised training mechanism that is tailored to the task of ischemic stroke lesion segmentation by exploiting color-coded parametric maps generated from Computed Tomography Perfusion scans. We illustrate the benefits of our proposed training mechanism, leading to considerable improvements in performance in the few-shot setting. Given a single annotated patient, an average Dice score of 0.58 is achieved for the segmentation of ischemic lesions.

Index Terms— Acute Ischemic Stroke, Few-Shot Segmentation, Self-Supervision, Supervoxels

1. INTRODUCTION

Globally, neurological disorders are the leading cause of disability-adjusted life years and the second leading cause of death [1]. Cerebral stroke is a major contributor to this burden, being the number one cause of neurological disability and the third-leading cause of death and disability combined [2, 3]. In response, the World Health Organisation has identified the optimization of brain health as a fundamental step to ensure human health and well-being [1]. Towards this goal, we focus on understanding and segmenting ischemic brain tissue in patients suspected of acute ischemic stroke (AIS), which comprises the majority of strokes [4].

Computed Tomography (CT) Perfusion (CTP) is often used as the primary diagnostic modality in the initial stages of an AIS [5]. A CTP study is a 4D examination of an area of the brain where an ischemic stroke lesion (ISL) is suspected. The study collects a series of 3D scans over the same portion of the brain during injection of an iodinated contrast agent over time. Based on the brain tissue changes over time, color-coded parametric maps (PMs) are calculated. These PMs highlight the possible variations in the brain tissue, allowing neuroradiologists to diagnose and plan treatments rapidly [6, 7]. In particular, we consider the cerebral blood flow (CBF), cerebral blood volume (CBV), time-to-peak (TTP), and time-to-maximum (TMax) PMs and maximum intensity projection (MIP), which are all highly
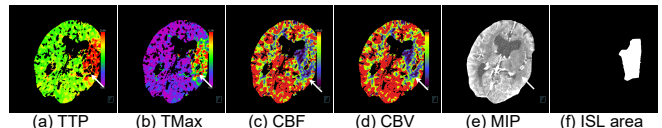

Fig. 1. Parametric maps (a-e) of a single brain slice plus the labeled ISL area (f). In this patient, there is an ischemic area in the vascular territory on the left middle cerebral artery (pointed by a white arrow).

relevant indicators for ISL segmentation [6]. Fig. 1 shows a set of PMs and MIP for an image slice and the corresponding labeled ISL.

Numerous applications have been explored to tackle this neurological disorder, traditionally using thresholding of different PMs [8, 9], but with lacking consensus on which PMs and threshold combination. Newer methods are relying on machine learning (ML) approaches [7, 10] or deep neural network (DNN) architectures [11, 12] and have generated promising results for segmenting ISLs. Nevertheless, all these architectures are supervised learning methods and require large amounts of training examples manually annotated by expert neuroradiologists. This can be problematic in the medical domain, where annotated training data is often scarce due to the time-consuming nature of the labeling process. Moreover, manual annotations might only give a coarse representation of the ISL, excluding all the minor details in the area, which can lead to imperfect predictions if used during model training.

A promising alternative, that has been gaining momentum in the last years, is self-supervised few-shot segmentation (FSS) [14, 15] where DNN algorithms only require a few labeled training images. This can be crucial in the medical domain for reducing the need for annotated data and enabling the models to also learn rare cases for which annotated data can be scarce. These models are trained and tested in episodes where a few labeled support images are used to guide the segmentation of the unlabeled query images. To bypass the need of manually annotating the training data, the training episodes are constructed in a self-supervised manner by leveraging automatically generated superpixel/supervoxel-based pseudolabels. Then, during inference, only a few manually labeled support images are required to segment the unlabeled query images.

Ouyang et al. [14] introduced a novel FSS framework for detecting organs in medical images using superpixel-based pseudolabels rather than manual annotations during training. Hansen et al. [15] further extended the self-supervision task to leverage 3D information via supervoxels and proposed a novel anomaly detection-inspired approach for prototypical FSS to increase the model's robustness to the heterogeneous background, resulting in the current state-of-the-art self-supervised FSS model. During each training episode, two random slices containing the same random supervoxel are sampled to
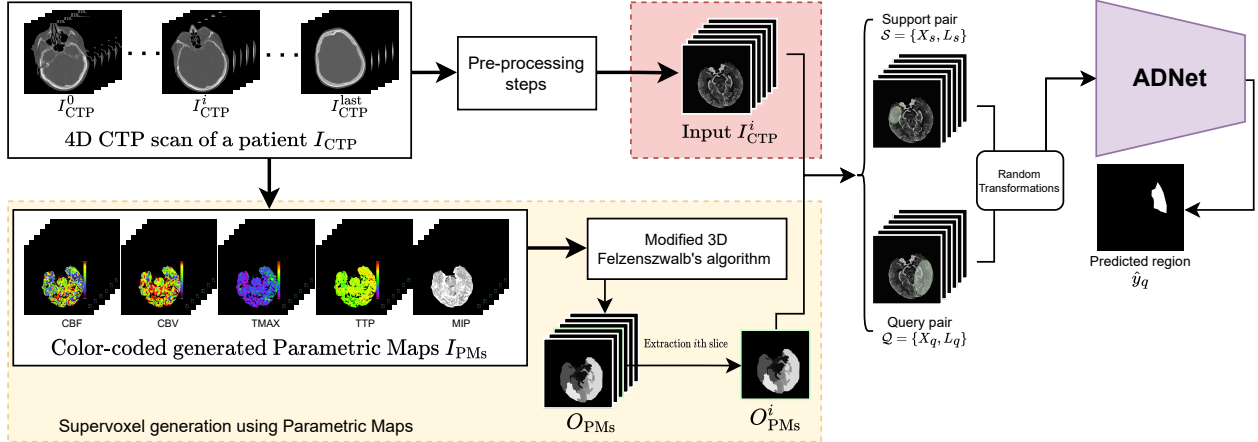
Fig. 2. General overview of the various steps involved in the method. Supervoxel regions are generated from the color-coded PMs with a modified 3D version of Felzenszwalb's image segmentation algorithm [13]. A series of pre-processing steps (details in Sec. 2.1) are implemented to extract brain tissue. Two random slices containing the same supervoxel are then sampled to act as the support and query slice, with the supervoxel region acting as the pseudolabel for the self-supervised task. Random transformations are then applied to either the support or query volume to provide the support-query pair that is used as input to train the ADNet model.

act as the support and query slice, with the supervoxel region acting as the pseudolabel for the self-supervised task. Random transformations are then applied to the support or query image-pseudolabel pair and the objective is to segment the query's pseudolabel using as reference the support pair. This gives the model a procedure to train unsupervised, i.e., without using labeled data.

These recent methods rely on the extraction of superpixels/supervoxels from the raw input images, which works well when the task is to segment classes that are relatively homogeneous and distinctly different from the background (e.g., liver or kidneys in abdominal CT scans) and where regions of interest thus are likely to be grouped in supervoxels. However, their performance decreases considerably when the class to segment is not distinctly different from the background, as is the case in ISL segmentation. For this reason, we propose an extension of the approach developed by Hansen et al. [15] that is particularly suited for the task of ISL segmentation. Unlike Hansen et al. [15], we do not generate supervoxels directly from the raw 4D CTP study, but instead we leverage domain knowledge generating supervoxels as targets for the self-supervised learning task from the PMs. We argue that while supervoxels generated from raw 4D CTP studies cannot correctly describe the boundaries inside the brain tissue, supervoxels extracted from the PMs can capture informative sub-regions, thus resulting in improved training targets. To the best of our knowledge, this is the first study exploiting self-supervised mechanisms for ISL segmentation.

The main contributions of this work can be summarized as follows:

1. We propose a novel self-supervised mechanism specifically tailored for ISL segmentation that leverages supervoxels extracted from a set of stacked PMs.
2. Exploiting the self-supervised mechanism, we propose a prototypical FSS framework for ISL segmentation and demonstrate the benefit of leveraging PMs for self-supervision.

2. DATA MATERIAL

We have analyzed 152 CTP scans from distinct patients acquired with a Siemens CT scanner between January 2014 and August 2020. Patients were divided into two groups based on the level of vessel occlusion: 77 patients had a large vessel occlusion (LVO) and 60 patients had a non-large vessel occlusion (Non-LVO). In addition, we included 15 patients admitted with stroke-like symptoms, but after diagnostic work-up were determined not to have suffered from an ischemic stroke, defined as without ischemic stroke (WIS). A split of 89/30/33 patients was conducted for training, validation, and testing. The CTP scans and the relative PMs were obtained using the "syngo.via" software from Siemens Healthineers (Siemens Medical Solutions, USA). Two expert neuroradiologists with 17 and 4.5 years of clinical experience manually annotated the ischemic areas using an in-house software. The annotations were delineated by examining the entire set of PMs (CBV, CBF, TTP, TMax) and MIP generated from a CTP study. Moreover, they used as assistance follow-up examination, including Magnetic Resonance Imaging (MRI) for most patients, performed within 1 to 3 days after the initial CT examination to assess the final infarction areas.

2.1. Pre-processing steps

Each CTP study underwent a series of pre-processing steps to extrapolate brain tissue from the raw scans and to enhance the contrast. The pre-processing steps can be summarized as follow:

1. Each study is converted into Hounsfield unit (HU) values to have a known quantitative scale for describing radio-density.
2. Co-registration of each slice is executed using the first slice as a frame of reference.
3. Brain tissue is extracted using an algorithm by Najm et al. [16]. This algorithm was selected due to its public availability and proven efficiency for CT scans.
4. Gamma correction and histogram equalization are performed after the brain's extraction for contrast enhancement.

3. METHOD

In this work, we propose a PM-based self-supervised few-shot segmentation model for ISLs in patients suspected of AIS[1]. The task is formulated as a binary segmentation, where the foreground corresponds to the area of interest. Instead of extracting pseudolabels from the raw input as in previous approaches [14, 15], which for ISLs will lead to sub-optimal performance due to the homogeneous nature of the brain region, we leverage domain-knowledge to improve the self-supervision task. In particular, we extract pseudolabels from the color-coded PMs that capture more informative sub-

---

[1]Code is publicly available at https://github.com/Biomedical-Data-Analysis-Laboratory/ADNet-for-AIS-segmentation
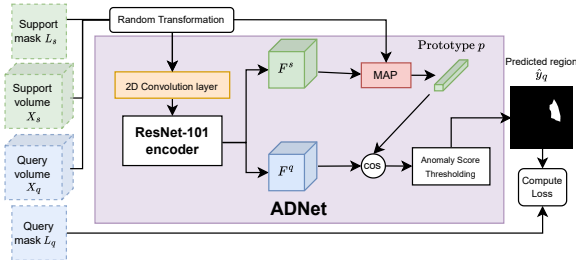
Fig. 3. Visual overview of the ADNet architecture [15].

regions. After generating the PMs from the 4D CTP input data and extracting pseudolabels, we take inspiration from the ADNet model proposed in [15] to learn a binary classifier that can segment the CTP input based on one annotated image slice. Fig. 2 gives a general overview of all the steps involved in the proposed method.

3.1. Parametric Map-Based Self-Supervision network

The proposed PM-based self-supervised network is an extension of the ADNet [15] and is trained using supervoxel-based pseudolabels extracted from the unlabeled training images. A supervoxel can be considered a group of neighboring voxels in an image with a similar nature. To obtain the supervoxels, we follow prior approaches in FSS [14, 15] and use the Felzenszwalb's algorithm [13], as it provides more diverse supervoxels compared to alternative approaches such as Simple Linear Iterative Clustering (SLIC) [17]. We leverage the 3D version of Felzenszwalb's algorithm to produce supervoxels from the 3D image volume[2] and extend the Euclidean distance calculation within the algorithm to support that each voxel in the image is represented by a vector, thus accepting 4D input tensors $I \in \mathbb{R}^{(W \times H \times Z \times M)}$ and returning a 3D volume $O \in \mathbb{R}^{(W \times H \times Z)}$ with the corresponding segmented regions. The four dimensions of the Felzenszwalb's input I are defined as width (W), height (H), depth (Z), and an additional dimension (M), corresponding to the number of modalities [3], hence $M = 5$; we named the Felzenszwalb's input I_{PMs} and the corresponding output O_{PMs}. A CTP scan after the pre-processing steps, which corresponds to the input of the proposed method, is called $I_{\text{CTP}} \in \mathbb{R}^{(W \times H \times Z \times M)}$, where the M corresponds to the time dimension ($M = T = 30$). Let X_s and X_q represent the support and query 2D+time slices extracted from the same brain volume I_{CTP}, respectively. The algorithm relies on one parameter, ρ, which controls the supervoxel size. Experiments were performed to tune ρ to improve the segmentation accuracy of the ischemic areas.

An illustration of the ADNet architecture is shown in Fig. 3. During training, the proposed model accepts support $\mathcal{S} = \{X_s, L_s\}$ and query $\mathcal{Q} = \{X_q, L_q\}$ pairs as input, where L_s and L_q are the 2D binary label images for X_s and X_q. In particular, each training episode is constructed from one unlabeled CTP scan by first choosing a random supervoxel extracted from the PMs to act as the foreground class, and then samples two 2D image slices (+time) containing this supervoxel to act as support and query image (X_s and X_q). Training labels L_s and L_q are constructed from the supervoxel segmentations of the corresponding PMs: the selected supervoxel is foreground while the union of the remaining supervoxels acts as background class. Random transformations are applied to one of the support or query volumes, based on a 50% probability, to encourage invariance to shape, and intensity differences [14]. X_s and X_q are fed to a shared feature encoder to extract high-level features F_s and F_q, respectively. The support foreground mask L_s is used to per-

form masked average pooling (MAP) of the support features F^s to compute a foreground prototype $p \in \mathbb{R}^d$, where d is the dimension of the embedding space. To segment the query volume, the negative cosine similarity (CS) is computed between the prototype p and the feature map F_q. The predicted foreground mask $\hat{y}_q$ is found by thresholding the computed CS with a learned parameter Th and the loss proposed in [15] is optimized given the query label L_q. During inference phase, we follow the evaluation protocol of [15] and sample one single 2D+time slice X_s from a support patient. The selected slice is always the middle slice of the support volume to avoid limited or non-existing ischemic regions at the edges of the volumes. L_s is the corresponding 2D slice with labeled ISL. The support pair $\{X_s, L_s\}$ is used to segment ISLs in the entire query patient one slice at a time. For a detailed explanation of the steps involved in the training and inference phases, we refer the reader to [15].

4. EXPERIMENTS

The proposed method is implemented to segment ISLs in patients suspected of AIS. In the following, we illustrate the benefit of leveraging the domain knowledge in the form of color-coded PMs for supervoxel extraction, rather than the raw CTP. For this, we compare the proposed method to a baseline approach that directly applies ADNet to the raw CTP scans, called *CTP-Baseline*. To further demonstrate the need to jointly leverage both the raw CTP scans and the PMs, we compare to another baseline that utilizes PMs both as input and for the pseudolabel generation, which we term *PMs-Baseline*.

Baselines Our *CTP-Baseline* model uses supervoxels generated from raw 4D CTP studies. For the Felzenszwalb's algorithm the 4D input is I_{CTP}. The same input is used in the model. The *PMs-Baseline* method utilizes supervoxels generated directly from a set of 3D PMs stacked over the fourth dimension (Sec. 3.1). The 4D input, for producing supervoxels and for the *PMs-Baseline*, is I_{PMs}.

Evaluation To demonstrate the applicability of our proposed method, we performed evaluations on the following datasets:

- **CTP-ALL** comprises the full dataset consisting of all three groups involved in this study (152 subjects).
- **CTP-LVO** contains only patients within the LVO group (77 subjects). Patients from this group are a clinically significant share of patients with AIS, considering the extension of the ischemic area and the grim natural course of the condition. This dataset is implemented to illustrate the variation effects during training.

Following standard practice in the literature [14, 15], we evaluate the performance by measuring the Dice score (DS); furthermore, we measure the predictions with the Matthews Correlation Coefficient (MCC) [18], which has been proven to provide a better measurement for imbalanced datasets [19]. Moreover, we employ the absolute difference in volumes $\Delta V(V_y, V_{\hat{y}}) = |V_y - V_{\hat{y}}|$ between the predicted volume $V_{\hat{y}}$ and the manually annotated volume V_y. This metric is essential for the WIS group because of the lack of labeled areas in this group, which makes it impossible to rely on the DS or the MCC.

We evaluate the effect of leveraging supervoxels obtained from the PMs via a comparison of the performance of our proposed method and the baseline approaches. In order to select the best hyperparameter ρ for the three models, we further performed a sensitivity study where we analyzed the parameter ρ in relation to the aggregate DS and the ΔV from the set of validation patients. These two-step experiments help create a fair comparison among the analyzed models.

5. RESULTS & DISCUSSION

Table 1 compares our proposed method with the baseline approaches for the test datasets split into the three groups (LVO, Non-LVO, WIS). Moreover, a comparison with the current state-of-the-art supervised method [12] is presented. Results are given as the mean

---

[2]Code is publicly available at https://github.com/sha168/Felzenszwalb-supervoxel-segmentation

[3]CBV, CBF, TTP, TMax, MIP in our proposed approach

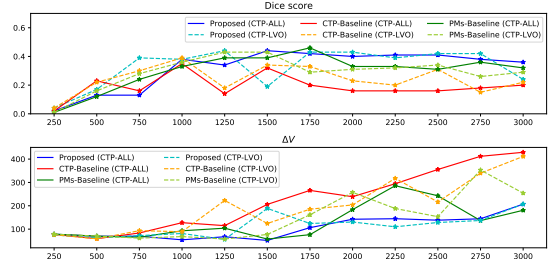| Method	Best ρ	Dataset	DS ⇑			MCC ⇑			ΔV (ml) ⇓			
			LVO	Non-LVO	All	LVO	Non-LVO	All	LVO	Non-LVO	WIS	All
Self-Supervised Methods												
CTP-Baseline	1000	CTP-LVO	0.50±0.14	0.07±0.03	0.43±0.12	0.48±0.14	0.10±0.05	0.43±0.12	72.6±15.8	85.4±26.5	212.1±63.0	89.5±15.4
PMs-Baseline	1250		0.50±0.15	0.09±0.03	0.45±0.14	0.48±0.15	0.11±0.03	0.44±0.14	74.8±15.7	50.3±42.8	75.5±41.3	66.7±19.5
Proposed	1250		**0.60±0.16**	**0.13±0.05**	**0.58±0.16**	**0.58±0.17**	**0.14±0.05**	**0.57±0.16**	**56.4±14.7**	**23.0±40.0**	**16.0±50.0**	**41.6±25.4**
CTP-Baseline	1000	CTP-ALL	0.47±0.11	0.09±0.02	0.39±0.09	0.45±0.11	**0.14±0.04**	0.39±0.10	84.2±13.4	128.9±23.3	176.3±31.5	107.5±17.4
PMs-Baseline	1750		0.50±0.13	**0.10±0.05**	0.45±0.12	0.48±0.14	0.14±0.07	0.43±0.13	86.3±15.1	70.5±34.1	219.6±57.9	93.2±15.6
Proposed	1500		**0.55±0.10**	0.06±0.03	**0.51±0.10**	**0.54±0.10**	0.07±0.03	**0.50±0.10**	**60.4±10.9**	**32.3±6.8**	**13.8±12.0**	**46.8±9.8**
Supervised Method (used only for comparison)												
Tomasetti et al. [12]	N.A.	CTP-ALL	0.85±0.06	0.49±0.49	0.66±0.32	0.83±0.06	0.49±0.29	0.64±0.31	19.6±15.2	4.9±5.4	0.3±0.4	12.9±14.2

Table 1. Mean (+ standard deviation) test results on the two datasets. All the methods (with the exception of Tomasetti et al. [12]) utilize a single labeled 2D+time slice during inference phase. Tomasetti et al. [12] uses a supervised approach and relies on the full labeled dataset for training and inference. Note that higher values are preferable for the DS and MCC (⇑), while for ΔV, lower values are better (⇓). Best performance are highlighted in **bold** (last row is only used for comparing supervised method and our self-supervised approaches).
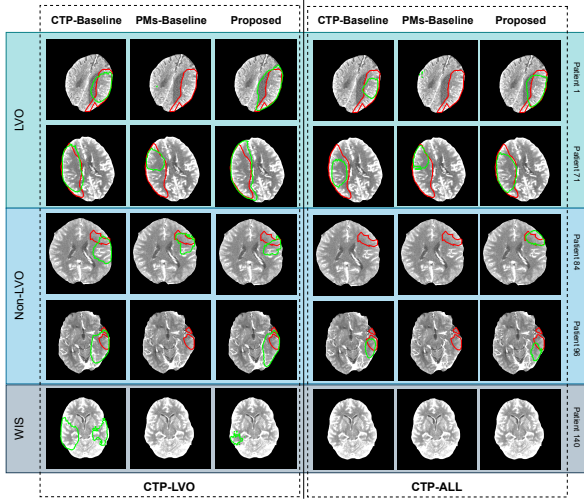
Fig. 4. Qualitative results of predicted ISL from random brain slices of different patients from the test set. Predictions are in green, while the ISL manually annotated ground truth are in red.

Fig. 5. Analysis of various ρ parameters (x-axis) in relation to the DS results (top graph) and the ΔV (bottom graph) on the y-axis for the validation set for the two datasets used.

(+ standard deviation) of 22 inference runs with a single slice from different patients as support. The proposed method and baseline results are shown with the best-selected ρ parameter chosen over a set of experiments on the validation set. In Table 1, we can see that our proposed model outperforms the baseline approaches for both datasets. The overall best performance is obtained by our proposed method using the **CTP-LVO** dataset for training, which indicates that the use of a single group of patients during the training phase gives a better understanding of the ischemic region.

The qualitative results in Fig. 4 support the quantitative results in Table 1. Training the models with images from the **CTP-LVO** dataset results in satisfactory segmentations for the ISLs, especially for the LVO group. With the adoption of the **CTP-ALL** as the training dataset, the results follow the correct structure of the ISL but with a tendency of under-segmentation in comparison with the method trained with the **CTP-LVO** dataset. Outcomes from our proposed approaches (third and last columns in Fig. 4) are not perfect, but they strongly reflect the area in the brain with an ISL, especially for patients in the LVO group. The *CTP-Baseline* model brings unnecessary under-segmentation (fourth column in Fig. 4) or excessive over-segmentation (first column in Fig. 4), depending on the dataset used for training. The *PMs-Baseline* approach yields more stable segmentations across different support images. Employing I_{PMs} as input in creating the supervoxels and using them in combination with I_{CTP} as input for the model provides a superior balance in the predicted segmentation regions; however, none of the methods are suitable for predicting ISL in Non-LVO patients. The meager nature of these areas makes it harder for the model to understand and adequately segment the ISL. The proposed approach, trained with the **CTP-LVO** dataset, achieved the best metrics among all the analyzed groups. This implies that the proposed PMs-Based Self-Supervision method is more suitable to segment ischemic regions and that there is a benefit to using PMs for supervoxel generation rather than relying on supervoxels extracted from raw CTP studies. Although the performances are not as high as the supervised method [12], which utilizes the entire labeled dataset during training, the achieved segmentation accuracy for the LVO group is still encouraging. This accuracy was attained using only a single labeled 2D+time slice, emphasizing the significance of employing domain knowledge in a limited labeled data setting

Sensitivity Study   The performance for a range of ρ values is illustrated in Fig. 5. As it is possible to evince from the graphs, the performance of the proposed method are relatively robust in the size of the supervoxels inside this range ($1000 < \rho < 2500$), while for the baseline, the range was $1000 < \rho < 1750$. Outside this range, the performances are inferior for the DS, while the ΔV increase with the increment of the parameter ρ.

6. CONCLUSION

We proposed a novel self-supervised mechanism to perform few-shot segmentation of ISLs and integrated it into the current state-of-the-art self-supervised FSS model ADNet. Instead of leveraging supervoxels directly generated from the raw CTP data for self-supervised training, which leads to sub-optimal performance, the proposed model leveraged domain knowledge by extracting supervoxels from parametric maps using a modified 3D version of Felzenszwalb's algorithm. The present study demonstrated the benefit of leveraging domain knowledge, especially for patients affected by a large vessel occlusion. This study is a first step of exploring few-shot learning approaches for ISL segmentations. Further research is still required to validate the results.

7. COMPLIANCE WITH ETHICAL STANDARDS

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] World Health Organization et al., "Optimizing brain health across the life course: Who position paper," 2022.

[2] Valery L Feigin, Emma Nichols, Tahiya Alam, Marlena S Bannick, Ettore Beghi, Natacha Blake, William J Culpepper, E Ray Dorsey, Alexis Elbaz, Richard G Ellenbogen, et al., "Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the global burden of disease study 2016," *The Lancet Neurology*, vol. 18, no. 5, pp. 459–480, 2019.

[3] Valery L Feigin, Benjamin A Stark, Catherine Owens Johnson, Gregory A Roth, Catherine Bisignano, Gdiom Gebreheat Abady, Mitra Abbasifard, Mohsen Abbasi-Kangevari, Foad Abd-Allah, Vida Abedi, et al., "Global, regional, and national burden of stroke and its risk factors, 1990–2019: A systematic analysis for the global burden of disease study 2019," *The Lancet Neurology*, vol. 20, no. 10, pp. 795–820, 2021.

[4] Mira Katan and Andreas Luft, "Global burden of stroke," in *Seminars in neurology*. Thieme Medical Publishers, 2018, vol. 38, pp. 208–211.

[5] European Stroke Organisation (ESO) Executive Committee, ESO Writing Committee, et al., "Guidelines for management of ischaemic stroke and transient ischaemic attack 2008," *Cerebrovascular diseases*, vol. 25, no. 5, pp. 457–507, 2008.

[6] KD Kurz, G Ringstad, A Odland, R Advani, E Farbu, and MW Kurz, "Radiological imaging in acute ischaemic stroke," *European journal of neurology*, vol. 23, pp. 8–17, 2016.

[7] Luca Tomasetti, Liv Jorunn Høllesli, Kjersti Engan, Kathinka Dæhli Kurz, Martin Wilhelm Kurz, and Mahdieh Khanmohammadi, "Machine learning algorithms versus thresholding to segment ischemic regions in patients with acute ischemic stroke," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 2, pp. 660–672, 2021.

[8] Girish Bathla, Kaustubh Limaye, Bruno Policeni, Ernst Klotz, Markus Juergens, and Colin Derdeyn, "Achieving comparable perfusion results across vendors. the next step in standardizing stroke care: a technical report," *Journal of neurointerventional surgery*, vol. 11, no. 12, pp. 1257–1260, 2019.

[9] A Bivard, C Levi, V Krishnamurthy, J Hislop-Jambrich, P Salazar, B Jackson, S Davis, and M Parsons, "Defining acute ischemic stroke tissue pathophysiology with whole brain ct perfusion," *Journal of neuroradiology*, vol. 41, no. 5, pp. 307–315, 2014.

[10] Wu Qiu, Hulin Kuang, Johanna M Ospel, Michael D Hill, Andrew M Demchuk, Mayank Goyal, and Bijoy K Menon, "Automated prediction of ischemic brain tissue fate from multiphase ct-angiography in patients with acute ischemic stroke using machine learning," *medRxiv*, 2020.

[11] Kimberly Amador, Matthias Wilms, Anthony Winder, Jens Fiehler, and Nils D Forkert, "Predicting treatment-specific lesion outcomes in acute ischemic stroke from 4d ct perfusion imaging using spatio-temporal convolutional neural networks," *Medical Image Analysis*, p. 102610, 2022.

[12] Luca Tomasetti, Mahdieh Khanmohammadi, Kjersti Engan, Liv Jorunn Høllesli, and Kathinka Dæhli Kurz, "Multi-input segmentation of damaged brain in acute ischemic stroke patients using slow fusion with skip connection," *arXiv preprint arXiv:2203.10039*, 2022.

[13] Pedro F Felzenszwalb and Daniel P Huttenlocher, "Efficient graph-based image segmentation," *International journal of computer vision*, vol. 59, no. 2, pp. 167–181, 2004.

[14] Cheng Ouyang, Carlo Biffi, Chen Chen, Turkay Kart, Huaqi Qiu, and Daniel Rueckert, "Self-supervision with superpixels: Training few-shot medical image segmentation without annotation," in *European Conference on Computer Vision*. Springer, 2020, pp. 762–780.

[15] Stine Hansen, Srishti Gautam, Robert Jenssen, and Michael Kampffmeyer, "Anomaly detection-inspired few-shot medical image segmentation through self-supervision with supervoxels," *Medical Image Analysis*, p. 102385, 2022.

[16] Mohamed Najm, Hulin Kuang, Alyssa Federico, Uzair Jogiat, Mayank Goyal, Michael D Hill, Andrew Demchuk, Bijoy K Menon, and Wu Qiu, "Automated brain extraction from head ct and cta images using convex optimization with shape propagation," *Computer Methods and Programs in Biomedicine*, vol. 176, pp. 1–8, 2019.

[17] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk, "Slic superpixels," Tech. Rep., 2010.

[18] Brian W Matthews, "Comparison of the predicted and observed secondary structure of t4 phage lysozyme," *Biochimica et Biophysica Acta (BBA)-Protein Structure*, vol. 405, no. 2, pp. 442–451, 1975.

[19] Davide Chicco, "Ten quick tips for machine learning in computational biology," *BioData mining*, vol. 10, no. 1, pp. 1–17, 2017.