Self-supervised Learning Versus Segmentation-aware Augmentation for Acute Ischemic Stroke CT Segmentation Under Annotation Scarcity

Livingstone Eli Ayivor[1,2][0009−0008−3032−204X], Prince Ebenezer Adjei[1,2,3][0000−0001−6244−2582], Belinda Esinam Akpaloo[5], Kojo Obed Otoo[5][0000−0001−8828−3188], Joseph D. A. Nti Koduah[1,2][0009−0001−5910−0693], and John Amuasi[2,3,4][0000−0002−8640−2662]

[1] Department of Computer Engineering, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana.
[2] Global Health and Infectious Disease Group, Kumasi Centre for Collaborative Research in Tropical Medicine, Kumasi, Ghana.
[3] Global One Health Group, Bernhard Notch Institute of Tropical Medicine (BNITM), Hamburg, Germany
[4] Department of Global Health, School of Medical Sciences, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
[5] Department of Radiology, Komfo Anokye Teaching Hospital, Kumasi, Ghana
peadjei.coe@knust.edu.gh

Abstract. Deep learning-based ischemic stroke lesion segmentation from non-contrast computed tomography (NCCT) is constrained by the limited availability of voxel-level annotations, especially in low-resource clinical settings. Two commonly explored approaches for improving data efficiency are self-supervised learning (SSL), which learns transferable representations from unlabeled data, and data augmentation, which increases the diversity of available labeled examples. However, their relative contribution under extreme annotation scarcity remains unclear.

In this work, we perform a controlled evaluation of SSL pretraining and segmentation-aware augmentation for acute ischemic stroke lesion segmentation using a locally acquired NCCT dataset. The dataset comprises 29 manually annotated and 178 unlabeled CT volumes collected from a tertiary hospital. We adapt a segmentation-aware augmentation framework to three-dimensional stroke CT by incorporating lesion-localized elastic deformation together with CT-compatible geometric and intensity transformations. Using Swin UNETR as the segmentation architecture, we evaluate four training configurations by independently varying the use of SSL pretraining and segmentation-aware augmentation.

Five-fold cross-validation shows that segmentation-aware augmentation provides the largest performance improvement, increasing the mean Dice score from 0.502 to 0.555 without SSL pretraining and from 0.511 to 0.551 with SSL pretraining. In comparison, SSL pretraining alone provides only marginal improvement when using the original annotated dataset and does not yield additional gains when combined with augmentation.

These findings indicate that segmentation-aware augmentation can substantially improve data efficiency in severely annotation-limited stroke CT segmentation, while the benefit of SSL may depend on the scale and diversity of the available unlabeled data.

Keywords: Stroke lesion segmentation · Non-contrast CT · Self-supervised learning · Data augmentation · Swin UNETR · Low-resource medical imaging

1 Introduction

Stroke is a major global health challenge and remains the second leading cause of death and long-term disability worldwide [7]. There is an estimated 15 million cases of stroke annually, with one-third of these cases leading to death [6]. The burden of stroke is increasingly concentrated in low- and middle-income countries (LMICs), where existing healthcare infrastructures remain unprepared to deal with the surge in cases [7,11,14,18]. Ischemic stroke represents the majority of stroke cases, and timely assessment of infarct extent is important for clinical decision-making. Non-contrast computed tomography (NCCT) remains the primary imaging modality for acute stroke evaluation because of its availability and relatively low operational requirements compared with advanced imaging techniques [2, 5, 16].

Despite the widespread availability of NCCT, accurate quantification of ischemic lesions remains challenging. Early ischemic changes on CT can be subtle, and lesion appearance varies considerably with factors such as time from symptom onset, infarct location, and acquisition characteristics [8]. Deep learning-based segmentation methods have demonstrated potential for automated stroke lesion delineation and quantitative analysis [17, 22, 24]. However, their success depends heavily on access to large, carefully annotated datasets. Generating voxel-level stroke lesion labels requires expert interpretation and substantial annotation effort, creating a significant bottleneck for model development. This challenge is severe in LMIC environments, where locally acquired datasets are often limited in size and may lack the extensive annotations required for training high-capacity models [15].

Several approaches have been proposed to improve learning efficiency when annotated data are scarce. Self-supervised learning (SSL) has emerged as a promising strategy by enabling models to learn useful representations from large collections of unlabeled images before supervised fine-tuning. Recent medical SSL approaches have demonstrated improvements across segmentation and classification tasks by exploiting anatomical and imaging patterns present in unlabeled datasets [4,23]. However, the benefit of SSL is influenced by the availability and diversity of pretraining data, the similarity between pretraining and downstream domains, and the amount of labeled data available for adaptation [1, 12].

An alternative approach is to improve the utility of existing annotations through data augmentation. Conventional augmentation strategies, including geometric and intensity transformations, are widely used to increase robustness in medical image segmentation [10, 13, 21]. However, generic transformations may not capture clinically meaningful variations in lesion morphology. Segmentation-aware augmentation approaches address this limitation by introducing controlled, localized perturbations guided by anatomical structures or lesion masks while maintaining image-label consistency [13, 19, 20].

Although SSL and augmentation represent two complementary strategies for addressing annotation scarcity, their relative contribution under extreme low-data conditions remains unclear. Most previous studies evaluate these approaches independently or using relatively large datasets, leaving limited understanding of their practical value when only a small number of local annotations are available. This question is especially important for LMIC-based clinical AI development, where collecting additional expert annotations may be substantially more difficult than acquiring unlabeled scans.

In this work, we investigate the contribution of SSL pretraining and segmentation-aware augmentation for ischemic stroke lesion segmentation using a locally acquired NCCT dataset. We adapt an augmentation framework to three-dimensional stroke CT and perform a controlled 2×2 evaluation using Swin UNETR, independently varying the use of SSL pretraining and augmentation.

2 Methods

2.1 Dataset

This retrospective study used a non-contrast CT (NCCT) ischemic stroke data acquired at the Komfo Anokye Teaching Hospital (KATH) in Ghana. After quality control, the dataset comprised 207 volumetric CT examinations, including 29 cases with voxel-wise infarct annotations and 178 unlabeled cases. The annotated subset was used for supervised segmentation, while the unlabeled scans were used exclusively for self-supervised pretraining.

All CT volumes were resampled to isotropic voxel spacing of $1 \times 1 \times 1$ mm, intensity clipped to the brain CT window, normalized, and resized to 256×256 in-plane resolution before training.

The study was conducted under institutional ethical approval from the KATH IRB, and all imaging data were de-identified prior to analysis.

2.2 Self-supervised Pretraining

Self-supervised learning (SSL) was used to exploit the 178 unlabeled NCCT volumes. A Swin UNETR [9] model was pretrained using masked volume inpainting, image rotation prediction and contrastive coding with Root Mean Squared Error (RMSE) loss on the unlabeled dataset. The resulting pretrained weights were subsequently used to initialize the downstream stroke lesion segmentation model.

For comparison, an identical Swin UNETR architecture was trained from random initialization using the same downstream training protocol and labeled data. The unlabeled cases were not used during supervised segmentation training.

2.3 Data Augmentation

To mitigate the limited number of annotated cases, we adapted the segmentation-aware augmentation strategy proposed in [3] for three-dimensional NCCT stroke segmentation. Rather than applying global geometric transformations alone, the adapted approach introduces localized anatomical variations around the lesion while maintaining spatial correspondence between the CT image and segmentation mask.

For each annotated training volume, the infarct mask was first morphologically dilated by 12 voxels to define a local deformation region surrounding the lesion. Within this region, elastic deformation was applied jointly to the CT image and segmentation mask using 16 control points and a maximum displacement of 8 voxels. The deformation region was itself transformed jointly with the image and label to preserve consistent lesion boundaries.

Additional CT-compatible transformations were applied probabilistically, including random left-right flipping ($p = 0.5$), random affine transformations involving scaling between 0.9 and 1.1, rotations up to 8° and translations of up to 10 voxels ($p = 0.4$), and low-amplitude Gaussian noise with a standard deviation of up to 0.02 ($p = 0.3$). The localized elastic deformation was applied with probability $p = 0.5$.

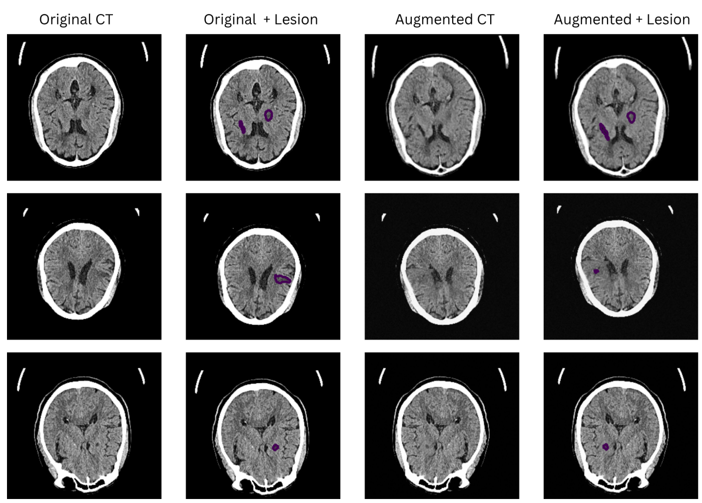

Fig. 1: Qualitative examples of the adapted segmentation-aware augmentation strategy for ischemic stroke NCCT. Each row shows an original CT slice with its corresponding infarct mask and an augmented counterpart generated using the transformation pipeline.

Table 1: **Effect of self-supervised pretraining and segmentation-aware augmentation on five-fold cross-validation DSC. Values represent mean DSC $\pm$ standard deviation.**

	Augmentation	
SSL pretraining	No	Yes
No	0.502 ± 0.003	$\mathbf{0.555 \pm 0.024}$
Yes	0.511 ± 0.023	0.551 ± 0.018

Augmented samples were generated offline, with three augmented variants generated from each annotated training case. Importantly, augmentation was performed independently within each training fold after patient-level fold assignment.

2.4 Training and Evaluation

A single Swin UNETR [9] architecture was used for all experiments. To evaluate the individual and combined contributions of self-supervised pretraining and segmentation-aware augmentation, a 2×2 factorial study was performed comprising: (i) SSL with augmentation, (ii) SSL without augmentation, (iii) augmentation without SSL, and (iv) neither SSL nor augmentation. Models were optimized using Dice loss with Adam optimization for 50 epochs with a batch size of 2. The same training protocol was applied across all configurations. Five-fold cross-validation was performed, and segmentation performance was measured using the Dice similarity coefficient (DSC).

3 Results

3.1 Segmentation Performance

Table 1 summarizes the five-fold cross-validation Dice similarity coefficient (DSC) for the four experimental conditions. The model trained using only the original annotated dataset with random initialization achieved a mean DSC of 0.502 ± 0.003. Introducing segmentation-aware augmentation increased performance to $\mathbf{0.555 \pm 0.024}$ without SSL pretraining and to 0.551 ± 0.018 when combined with SSL pretraining.

In contrast, the effect of SSL pretraining was smaller. Without augmentation, SSL increased the mean DSC from 0.502 ± 0.003 to 0.511 ± 0.023. However, when segmentation-aware augmentation was applied, SSL pretraining resulted in a comparable DSC (0.551 ± 0.018) to augmentation alone (0.555 ± 0.024).

Across the evaluated configurations, segmentation-aware augmentation provided the largest improvement, with an average increase of approximately 0.05 DSC compared with the corresponding non-augmented settings. SSL pretraining resulted in only a marginal improvement when used without augmentation.

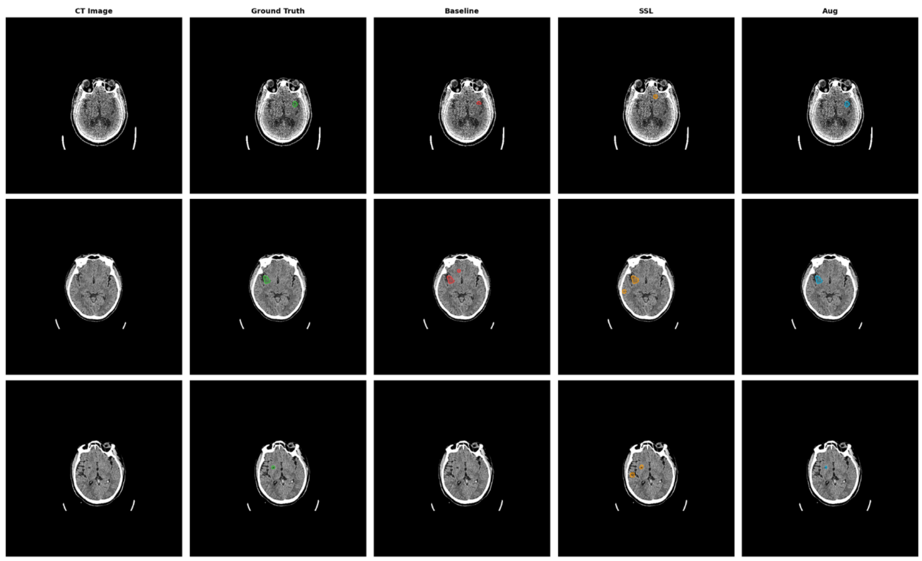

Fig. 2: Representative segmentation results from three test cases. Compared with the baseline model trained on the original annotated dataset, augmentation produced more accurate lesion delineation, particularly for small and poorly defined infarcts. SSL pretraining alone resulted in predictions similar to the baseline.

4 Discussion

This study investigated the relative contributions of self-supervised pretraining and segmentation-aware augmentation for ischemic stroke lesion segmentation under severe annotation scarcity. Across the evaluated configurations, segmentation-aware augmentation produced the largest observed improvement in Dice similarity coefficient (DSC), increasing performance from 0.502 ± 0.003 to 0.555 ± 0.024 without SSL and from 0.511 ± 0.023 to 0.551 ± 0.018 with SSL. In contrast, SSL pretraining provided only a modest improvement when the original annotated dataset was used, increasing DSC from 0.502 ± 0.003 to 0.511 ± 0.023.

The observed benefit of segmentation-aware augmentation may be attributed to its direct contribution to the diversity of supervised training examples. Rather than introducing only global geometric or intensity variations, the adapted strategy generates localized changes around the lesion while preserving image-label correspondence. This may be particularly useful when only 29 annotated cases are available, as the resulting variations expose the segmentation model to a broader range of lesion appearances and morphologies.

However, the effect of augmentation should be interpreted alongside an important optimization-related confound. The augmented data were generated offline, resulting in approximately three times as many training samples as in the corresponding non-augmented configurations under the same 50-epoch training protocol. Consequently, the observed improvement may reflect both increased sample diversity and greater exposure to training examples during optimization. The current experimental design does not allow these two effects to be separated.

The contribution of SSL was comparatively modest in the evaluated setting. SSL improved performance when the original 29 annotated cases were used, but the difference between augmentation alone (0.555 ± 0.024) and SSL with augmentation (0.551 ± 0.018) was small. Given the limited number of validation folds and annotated cases, this difference should not be interpreted as evidence that SSL provides no benefit when combined with augmentation. Rather, the results indicate that no additional improvement from SSL was observed under the conditions evaluated in this study. Larger and more diverse unlabeled datasets, together with alternative pretraining objectives, may produce different outcomes.

These findings should be interpreted in the context of the evaluated dataset and experimental design. The study was conducted using a single-center cohort with a small number of annotated cases and focused on a single segmentation architecture and SSL strategy. Consequently, the observed balance between augmentation and SSL may change with larger datasets, different imaging environments, or alternative pretraining approaches. Nevertheless, the results give insight into selecting data-efficient strategies for stroke segmentation when annotation resources are severely constrained.

Future work should investigate the sensitivity of the proposed strategy to the number of generated variants, compare segmentation-aware augmentation with conventional augmentation schemes, and evaluate additional segmentation architectures and self-supervised objectives. External validation on larger multi-center datasets would also be important for assessing the generalizability of these findings.

5 Conclusion

This study investigated self-supervised pretraining and segmentation-aware augmentation for acute ischemic stroke lesion segmentation under severe annotation scarcity. Using a controlled evaluation with Swin UNETR, segmentation-aware augmentation produced the largest observed improvement in segmentation performance, while SSL provided only modest gains under the evaluated data conditions. These findings highlight the potential of task-aware augmentation for improving data efficiency when expert annotations are limited, while motivating further evaluation of SSL with larger and more diverse unlabeled datasets.

Acknowledgments. The authors thank the Department of Radiology at Komfo Anokye Teaching Hospital for providing access to the clinical data used in this study. This work was supported by the Kwame Nkrumah University of Science and Technology Research Fund (KReF).

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

References

1. Al Kader Hammoud, H.A., Das, T., Pizzati, F., Torr, P.H., Bibi, A., Ghanem, B.: On pretraining data diversity for self-supervised learning. In: European Conference on Computer Vision. pp. 54–71. Springer (2024)
2. Al-Salahat, A., Pirahanchi, Y., Dhasakeerthi, T., Nayar, D., Almasri, S., Verma, K., Thatikala, A., Alkhaldi, M., Ebrahimzadeh, S.A., Shahsavari, N., Godasi, R., Kapoor, N., Sharma, R.: Current state and advancements of imaging in acute ischemic stroke: a practical review. Neurological Sciences **47** (12 2025). https://doi.org/10.1007/s10072-025-08685-8
3. Ankomah, C.T., Ayivor, L.E., Nyame, I., Wambo, L., Bonsu, P.Y., Iorumbur, A.M., Confidence, R., Musah, T.: How we won brats-ssa 2025: Brain tumor segmentation in the sub-saharan african population using segmentation-aware data augmentation and model ensembling. In: Bakas, S., Dennis, E., Astaraki, M., Baid, U., Conte, G.M., Foltyn-Dumitru, M., Jiang, Z., Labella, D., Metz, M.C., Anazodo, U., de Verdier, M.C., Kofler, F., Li, H.B., Linguraru, M.G., Maleki, N. (eds.) Segmentation, Classification, and Synthesis for Brain Tumors and Traumatic Brain Injuries. pp. 296–306. Springer Nature Switzerland, Cham (2026)
4. Bartlett, S., Rajan, S., Rawat, A., Christie, A., Tripathi, A.: Self-supervised pretraining of deep learning models for unlabeled medical image datasets. World Journal of Advanced Engineering Technology and Sciences **2**, 100–103 (05 2021). https://doi.org/10.30574/wjaets.2021.2.2.0031
5. Dhamija, R.K., Donnan, G.A.: The role of neuroimaging in acute stroke. Ann Indian Acad Neurol **11**(1), S12–S23 (Jan 2008)
6. Feigin, V.L., Abajobir, A.A., Abate, K.H., Abd-Allah, F., Abdulle, A.M., Abera, S.F., Abyu, G.Y., Ahmed, M.B., Aichour, A.N., Aichour, I., et al.: Global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the global burden of disease study 2015. The Lancet Neurology **16**(11), 877–897 (2017)
7. Feigin, V.L., Brainin, M., Norrving, B., Martins, S.O., Pandian, J., Lindsay, P., F Grupper, M., Rautalin, I.: World stroke organization: Global stroke fact sheet 2025. Int. J. Stroke **20**(2), 132–144 (Feb 2025)
8. Gao, J., Parsons, M.W., Kawano, H., Levi, C.R., Evans, T.J., Lin, L., Bivard, A.: Visibility of CT early ischemic change is significantly associated with time from stroke onset to baseline scan beyond the first 3 hours of stroke onset. J. Stroke **19**(3), 340–346 (Sep 2017). https://doi.org/10.5853/jos.2016.01424
9. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images (2022), https://arxiv.org/abs/2201.01266
10. Kamsari, M., Sadeghi, S., de Oliveira, G.G., Alves, A.M., Sarshar, N.T., Anari, S., Ranjbarzadeh, R.: The role of data augmentation and attention mechanisms in UNet and ConvNeXt architectures for optimizing breast tumor segmentation. Sci. Rep. **15**(1), 45268 (Nov 2025). https://doi.org/10.1038/s41598-025-29841-2
11. Kengne, A.P., Anderson, C.S.: The neglected burden of stroke in sub-saharan africa. International Journal of Stroke **1**(4), 180–190 (2006). https://doi.org/https://doi.org/10.1111/j.1747-4949.2006.00064.x, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1747-4949.2006.00064.x
12. Konstantakos, S., Cani, J., Mademlis, I., Chalkiadaki, D.I., Asano, Y.M., Gavves, E., Papadopoulos, G.T.: Self-supervised visual learning in the low-data regime: A comparative evaluation. Neurocomputing **620**, 129199 (2025).

https://doi.org/https://doi.org/10.1016/j.neucom.2024.129199, https://www.sciencedirect.com/science/article/pii/S0925231224019702

13. Kozah, N., Dornaika, F., Charafeddine, J., El Jaam, J.: Data augmentation techniques for medical image segmentation – a review. In: 2024 International Conference on Computer and Applications (ICCA). pp. 1–8 (2024). https://doi.org/10.1109/ICCA62237.2024.10927851
14. Krishnamurthi, R.V., Ikeda, T., Feigin, V.L.: 2020) global, regional and Country-Specific burden of ischaemic stroke, intracerebral haemorrhage and subarachnoid haemorrhage: A systematic analysis of the global burden of disease study. Neuroepidemiology **54**, 171–179 (2017)
15. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A., van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. Medical Image Analysis **42**, 60–88 (2017). https://doi.org/https://doi.org/10.1016/j.media.2017.07.005, https://www.sciencedirect.com/science/article/pii/S1361841517301135
16. Menon, B.K., Campbell, B.C., Levi, C., Goyal, M.: Role of imaging in current acute ischemic stroke workflow for endovascular therapy. Stroke **46**(6), 1453–1461 (2015). https://doi.org/10.1161/STROKEAHA.115.009160, https://www.ahajournals.org/doi/abs/10.1161/STROKEAHA.115.009160
17. Musah, T., Adjei, P.E., Otoo, K.O.: Automated Segmentation of Ischemic Stroke Lesions in Non-contrast Computed Tomography Images for Enhanced Treatment and Prognosis, p. 73–80. Springer Nature Switzerland (2025). https://doi.org/10.1007/978-3-031-79103-1_8, http://dx.doi.org/10.1007/978-3-031-79103-1_8
18. Prust, M.L., Forman, R., Ovbiagele, B.: Addressing disparities in the global epidemiology of stroke. Nat. Rev. Neurol. **20**(4), 207–221 (Apr 2024)
19. Sharma, V.K., Rout, R.K., Tanwar, R., Singh, P.K.: An efficient approach to synthesize medical images for data augmentation with constrained training data. Discov. Comput. **29**(1) (Feb 2026). https://doi.org/10.1007/s10791-026-09980-4
20. Tian, S., Yang, L., Zhao, Z., Xie, B., Lv, Y.: Lesion-centric data augmentation: A novel strategy to enhance robustness in medical image segmentation. In: Proceedings of the 2025 6th International Symposium on Artificial Intelligence for Medical Sciences. p. 481–485. ISAIMS '25, Association for Computing Machinery, New York, NY, USA (2026). https://doi.org/10.1145/3777577.3777656, https://doi.org/10.1145/3777577.3777656
21. Uniyal, Y., Uniyal, V., Bachheti, Y., Lakhera, N., Rawat, R.: Data augmentation techniques applied to medical images **5**, 483–501 (07 2024)
22. Verma, K., Kumar, S., Paydarfar, D.: Automatic segmentation and quantitative assessment of stroke lesions on MR images. Diagnostics (Basel) **12**(9), 2055 (Aug 2022)
23. Wald, T., Ulrich, C., Suprijadi, J., Ziegler, S., Nohel, M., Peretzke, R., Köhler, G., Maier-Hein, K.H.: An openmind for 3d medical vision self-supervised learning (2025), https://arxiv.org/abs/2412.17041
24. Zare Lahijan, L., Meshgini, S., Afrouzian, R.: Deep learning-based automatic segmentation of ischemic stroke lesions in ct perfusion imaging. Biomimetics **11**(5) (2026). https://doi.org/10.3390/biomimetics11050334, https://www.mdpi.com/2313-7673/11/5/334