Enhancing New Multiple Sclerosis Lesion Segmentation via Self-supervised Pre-training and Synthetic Lesion Integration

Peyman Tahghighi[1(✉)], Yunyan Zhang[2,3], Roberto Souza[3,4], and Amin Komeili[1]

[1] Department of Biomedical Engineering, University of Calgary, Calgary, Canada
peyman.tahghighi@ucalgary.ca
[2] Cumming School of Medicine, Department of Clinical Neurosciences, University of Calgary, Calgary, Canada
[3] Hotchkiss Brain Institute, University of Calgary, Calgary, Canada
[4] Department of Electrical and Software Engineering, University of Calgary, Calgary, Canada

Abstract. Multiple Sclerosis (MS) is a chronic and severe inflammatory disease of the central nervous system. In MS, the myelin sheath covering nerve fibres is attacked by the self-immune system, leading to communication issues between the brain and the rest of the body. Image-based biomarkers, such as lesions seen with Magnetic Resonance Imaging (MRI), are essential in MS diagnosis and monitoring. Further, detecting newly formed lesions provides crucial information for assessing disease progression and treatment outcomes. However, annotating changes between MRI scans is time-consuming and subject to inter-expert variability. Methods proposed for new lesion segmentation have utilized limited data available for training the model, failing to harness the full capacity of the models and resulting in limited generalizability. To enhance the performance of the new MS lesion segmentation model, we propose a self-supervised pre-training scheme based on image masking that is used to initialize the weights of the model, which then is trained for the new lesion segmentation task using a mix of real and synthetic data created by a synthetic lesion data augmentation method that we propose. Experiments on the MSSEG-2 challenge dataset demonstrate that utilizing self-supervised pre-training and adding synthetic lesions during training improves the model's performance. We achieved a Dice score of 56.15±7.06% and an F1 score of 56.69±9.12%, which is 2.06% points and 3.3% higher, respectively, than the previous best existing method. Code is available at: https://github.com/PeymanTahghighi/SSLMRI.

Keywords: Multiple Sclerosis · White Matter Lesion · MRI Segmentation · Self-supervised learning.

1 Introduction

Multiple Sclerosis (MS) stands as a chronic autoimmune disease targeting myelinated axons in the Central Nervous System (CNS). It impacts approximately 2.8

million individuals worldwide, with a prevalence rate of 35.9 cases per 100,000 and a female-to-male ratio of 2:1. [24]. These attacks in the CNS lead to the formation of focal lesions in the white matter. Magnetic Resonance Imaging (MRI) and, particularly, FLuid Attenuated Inversion Recovery (FLAIR) images are commonly used for understanding the quantity, distribution, and evolution of these white matter lesions, especially the emergence of new ones, which can be utilized during clinical treatment for assessing the disease status [14]. Manual delineation of newly appeared MS brain lesions from MRIs is a labor-intensive task and prone to expert annotation variability [1]. Consequently, the automated segmentation of the newly appeared lesions associated with MS is imperative for establishing a Computer-Aided Diagnosis system suitable for clinical applications.

Recent advancements in medical image segmentation using deep learning have succeeded in segmenting different structures from MRI scans [15, 23], and since delineating new lesions can be defined as a segmentation problem, attempts have been made to solve this problem using well-established segmentation methods [7, 2, 25]. For instance, Cabezas *et al.* [7] employed attention modules for feature fusion in the UNet [22], achieving third place in the MSSEG-2 [7] challenge with a 48.5% Dice score. Ashtari *et al.* [2] utilized a 3D UNet model with group normalization before convolutional layers and achieved a 45.6% Dice score on the MSSEG-2 challenge training set. Wu *et al.* [25] addressed data scarcity by proposing a heterogeneous model incorporating multi-time and single-time point MRI scans and introducing a new dataset for single-time point MRI scans. Despite achieving state-of-the-art results with a 59.91% Dice score on a single fold of the MSSEG-2 challenge training dataset, this came at the expense of gathering and labelling a new dataset. Tackling data scarcity has also been studied by Basaran *et al.* in three different studies in which they utilized CarveMix [4], Generative Adversarial Networks (GANs) [5] and proposed LesionMix [10] to add new synthetic lesions and in all cases, it improved the segmentation performance of the model.

Acquiring a large, labelled dataset of longitudinal MRI scans for MS lesions is challenging, and it requires time-consuming expert annotations of multi-time-point MRI data. Additionally, the frequency of appearance of new lesions in the longitudinal data is often low, causing data imbalance and challenges for training deep neural networks. Motivated by this, a self-supervised pre-training based on image masking is proposed. We also propose a synthetic lesion data augmentation approach, which can be used during the model's downstream task training in combination with real data to address new lesion examples scarcity and the high cost of gathering annotated longitudinal MRI scans. Detailed experiments on the MSSEG-2 challenge dataset showed a Dice score of 56.15±7.06% and F1 of 56.69±9.12%, which is higher by 2.06% points and 3.3%, respectively, than the current sate-of-the-art method [25]. Overall, the contributions of this work are: (1) Proposal of a self-supervised pre-training paradigm to enhance the performance of the lesion segmentation model. (2) Development of a new augmentation strategy for creating synthetic white matter lesions to alleviate the

problem of data scarcity. (3) Achieving state-of-the-art results on the training set of the MSSEG-2 challenge dataset using 5-fold cross-validation.

2 Methods

2.1 Self-supervised pre-training

To enhance the performance of the segmentation model, we pre-trained the model with a masking pretext task similar to [23, 3], in which we masked certain parts of an MR image to remove brain structures and tasked the model to regress the differences between two given MR images, i.e., the original and the masked images. We defined a model $f_\theta\left(x_1, x_2,\ d\right) \rightarrow \mathbb{R}$ for self-supervised pre-training, where x_1 and x_2 are the first and second input MR images, respectively, and similar to [25], d is the difference between the first and second inputs, *i.e.*, $d = x_1 - x_2$.

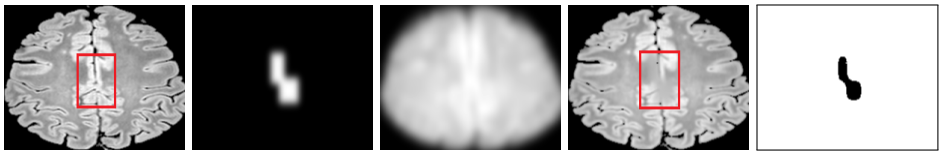
Let m^{base} be an MR image, and we want to mask certain parts from it to create m^{masked}. To find candidate locations for masking on the brain surface, Otsu threshold [21] was done on m^{base} to ignore black regions such as Cerebrospinal Fluid (CSF). Then, 1 to 5 voxels were sampled as the center location to generate cuboids. The cuboids' side sizes were sampled from a uniform distribution between 3 and 10 mm to mask regions of varying sizes. These cuboids were stored in a tensor, c, filled with ones inside the sampled cuboid regions and zeros everywhere else. A Gaussian filter with σ randomly sampled from a uniform distribution between 2 and 4 was applied to c next to generate c^{blur} (Figure 1b) for smoother masking, and we chose a small σ to maintain the overall shape of cuboid regions. A Gaussian filter with $\sigma = 7$ for a strong blurring effect was applied to m^{base} to obtain m^{blur} (Figure 1c). Then m^{base} and m^{blur} were combined to produce the masked MR image:

$$m^{masked} = (1 - c^{blur}) \odot m^{base} + c^{blur} \odot m^{blur} \tag{1}$$

Where $\odot$ represents the point-wise multiplication.

To generate a binary mask of the masked region to guide the training, c^{blur} was then lower thresholded with $T = 0.8$ to only include strongly masked areas, since a voxel value close to 1 in c^{blur} cause that voxel to take most of its value from m^{blur} leading to a stronger removal. We call this binary mask h (Figure 1e). Regions represented with ones in h correspond to regions that did not change, and regions filled with zeros correspond to regions that have changed during the masking process. To make the trained model insensitive to changes such as noise, Gaussian smoothing, Gaussian Noise and Gibbs Noise were applied to m^{masked} randomly after masking.

For training, the output was a heatmap highlighting the difference from the first input (x_1) to the second (x_2), in which, if added to the x_1, the result would match x_2. Motivated by recent methods in depth estimation [12], the model was run twice to generate two outputs op_1 and op_2 by permuting the order the inputs



(a) m^{base} (b) c^{blur} (c) m^{blur} (d) m^{masked} (e) h

Fig. 1. Overall masking steps. (a) MR image to be masked (m^{base}). (b) Randomly generated smoothed regions that will be used for masking (c^{blur}). (c) Smoothed MR image (m^{blur}). (d) Masked image (m^{masked}), and (e) binary map indicating the voxels that changed during masking in black. The red rectangles in (a) and (d) highlight the changes inserted due to the masking process.

m^{base} and m^{masked} were presented to the model (Figure 2a):

$$\begin{aligned} op_1 &= f_\theta\left(x_1 = m^{base},\ x_2 = m^{masked},\ d = m^{base} - m^{masked}\right) \\ op_2 &= f_\theta\left(x_1 = m^{masked},\ x_2 = m^{base},\ d = m^{masked} - m^{base}\right) \end{aligned} \tag{2}$$

Five different pre-training loss terms were utilized. Absolute reconstruction error between the reconstructed masked MR image, r^{masked}, and the actual masked MR image, m^{masked}:

$$Lp_1 = \left|r^{masked} - m^{masked}\right| \quad s.t \quad r^{masked} = op_1 + m^{base} \tag{3}$$

Absolute reconstruction error between reconstructed base MR image r^{base} and the actual base MR image, m^{base}:

$$Lp_2 = \left|r^{base} - m^{base}\right| \quad s.t \quad r^{base} = op_2 + m^{masked} \tag{4}$$

As both outputs should represent the same changes with different signs ($op_1 = -op_2$) and following recent methods in depth estimation [12], a consistency loss, Lp_3, was defined to enforce this:

$$Lp_3 = |op_1 + op_2| \tag{5}$$

To force the model to neglect changes due to noise added through augmentations to m^{masked}, Lp_4 and Lp_5 were defined as below:

$$Lp_4 = |op_1 \odot h| \qquad Lp_5 = |op_2 \odot h| \tag{6}$$

By minimizing losses in Equation 6, we enforced the model only to generate an output for the target masked region, as shown in black in Figure 1e and ignore changes in other areas, shown in white, which came from augmentations we applied to m^{masked}. The final loss function was the sum of all loss functions defined here with equal contribution. We utilized the VNet [20] as the deep learning model for both segmentation and self-supervised pre-training. For the self-supervised pre-training, following [23], a convolutional head was added on top of the last encoder output, replacing the VNet decoder, which consists of *upsample*→*conv3d*→*instance norm*→*ReLU*. Features from this level of the encoder were upsampled first by this head before the final convolutional layer.

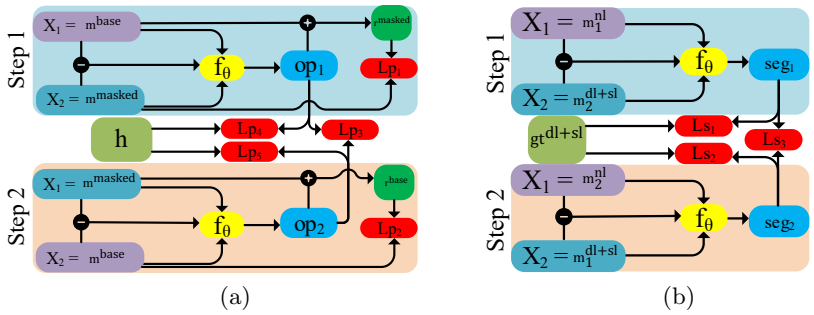


Fig. 2. (a) For self-supervised pre-training, our model generates two outputs, op_1 and op_2 which represent changes from m^{base} to m^{masked} and m^{masked} to m^{base} respectively. To only focus on masked regions and ignore changes in other areas due to noise, both op_1 and op_2 were compared against h, which is a binary mask only highlighting masked regions. (b) For new lesion segmentation, our model outputted new appeared lesions from m_1^{nl} to m_2^{dl+sl} in step 1 and m_2^{nl} to m_1^{dl+sl} in step 2. Since both seg_1 and seg_2 predicted the same set of newly appeared lesions, we defined Ls_3 to enforce this.

2.2 Synthetic lesions

Let m be the given MR image; in the first step, we applied a lower threshold to find a rough white matter mask. The threshold value was set to be the peak of the histogram computed on the voxels inside the brain mask since white matter voxels have the highest occurrence [13], and we call this tr_1. To filter out black regions in FLAIR scans, such as CSF, higher Otsu thresholding was performed, and we call it tr_2. Gradients of m were calculated using the Sobel filter [11] to put lesions only on smooth regions without abrupt changes in intensity and removing edges, and lower thresholded using Otsu, to have gr. The set of possible locations for a synthetic lesion, s, was the intersection of tr_1, tr_2 and gr. To reduce outliers, a morphological opening [11] was performed on s using a $3 \times 3 \times 3$ structuring element. A random number of points between 1 and 5 were sampled from s to determine synthetic lesion locations. An initial cuboid was used for each lesion, with side sizes randomly sampled from a uniform distribution between 1.5 and 4 mm. This range was determined based on the sizes of lesions in the dataset, with the aim of enhancing the model's generalizability. These lesions were then stored in a binary tensor l, where lesions were filled with ones and zeros elsewhere. To have a smoother blending of the lesions to m and to randomize the shape of the initial cuboid, Gaussian smoothing was applied with σ randomly sampled from a uniform distribution between 2 and 8, resulting in l^{blur}, as shown in Figure 3c. Note that although we started with a cuboid, Gaussian smoothing can transform it into arbitrary shapes, such as ellipsoid. As the final step, synthetic lesions were blended to m to have m^{sl} through the following equation (Figure 3d):

$$m^{sl} = \left(\left(1 - l^{blur}\right) \odot m\right) + l^{blur} \tag{7}$$

(a) m (b) s (c) l^{blur} (d) m^{sl}

Fig. 3. Steps of adding synthetic lesion. (a) Input MRI. (b) Candidate locations on white matter. (c) Random lesion shape generated. (d) Added lesion.

2.3 New lesion segmentation

The segmentation model objective was to generate a binary output that marks newly appeared lesions. Let m_1^{nl} be the baseline MR image with no new lesions and m_2^{dl} be the follow-up MR image with dataset-provided newly appeared lesions marked by binary mask gt^{dl}. We first added dataset-provided new lesions in m_2^{dl} to m_1^{nl} using CarveMix [26] and similar to [4] to have m_1^{dl} and then removed these lesions from m_2^{dl} using masking method explained in section 2.1 to have m_2^{nl} with no newly appeared lesions. Then, we added the same synthetic lesions to both m_1^{dl} and m_2^{dl} to have m_1^{dl+sl} and m_2^{dl+sl} with gt^{dl+sl} marking both dataset-provided and synthetic lesions. To this point, we have two different MR images with the same new lesions m_1^{dl+sl} and m_2^{dl+sl}, and new lesions mask gt^{dl+sl} (including synthetic and dataset-provided new lesions) and two MR images m_1^{nl} and m_2^{nl} with no newly appeared lesions. Augmentations such as Gibss Noise and Gaussian noise were applied next to both m_1^{dl+sl} and m_2^{dl+sl}. Similar to self-supervised pre-training, we defined the model as $f_\theta\left(x_1, x_2,\ d\right) \rightarrow \{0, 1\}$. The model outputs sigmoid probabilities, seg_1 and seg_2 as below (shown in Figure 2b):

$$\begin{aligned} seg_1 &= f_\theta\left(x_1 = m_1^{nl},\ x_2 = m_2^{dl+sl},\ d = m_1^{nl} - m_2^{dl+sl}\right) \\ seg_2 &= f_\theta\left(x_1 = m_2^{nl},\ x_2 = m_1^{dl+sl},\ d = m_2^{nl} - m_1^{dl+sl}\right) \end{aligned} \tag{8}$$

Where seg_1 specifies new lesions from m_1^{nl} to m_2^{dl+sl}, and seg_2 specifies the new lesions from m_2^{nl} to m_1^{dl+sl}. By doing so, we increased the number of examples of two timepoints MR images with new lesions to which our model was exposed. We defined two loss function terms, Ls_1 and Ls_2 as below:

$$\begin{aligned} Ls_1 &= Dice\left(seg_1, gt^{dl+sl}\right) + \alpha\ BL\left(seg_1, gt^{dl+sl}\right) \\ Ls_2 &= Dice\left(seg_2, gt^{dl+sl}\right) + \alpha\ BL\left(seg_2, gt^{dl+sl}\right) \end{aligned} \tag{9}$$

Here, Dice was combined with boundary loss, which was introduced to mitigate challenges in highly imbalanced segmentation tasks [16, 18], and α controls the contribution of boundary loss. Since both seg_1 and seg_2 are predicting the same

set of newly appeared lesions and motivated by recent research in depth estimation [12], a consistency loss between two output segmentations was also defined below, as shown in Figure 2b:

$$Ls_3 = Dice\left(seg_1, seg_2\right) \tag{10}$$

The final loss function was the summation of all the defined loss functions with equal contributions.

3 Experiments and Results

3.1 Dataset

We evaluated the proposed method on the MICCAI-21 MS new lesion segmentation challenge dataset (MSSEG-2) [7], which contains FLAIR images of MS patients at two timepoints. The follow-up data were obtained within 1–3 years after the first examination from 15 different scanners. The annotations were limited to new lesions at the second time point, omitting the delineation of growing or shrinking lesions. Four expert neuroradiologists initially segmented new lesions. A senior expert resolved disputes, followed by consensus fusion via majority voting. Since the test set from the challenge was not provided, we only had access to 40 MRI scans of the training set to train and evaluate our model using 5-fold cross-validation. We utilized the same folding split to train and evaluate our model and all the compared methods here. The MICCAI-16 (MSSEG) [9] dataset was utilized for self-supervised pre-training, including 53 single timepoint FLAIR MRI scans from four scanners. Note that no overlap exists between MSSEG and MSSEG-2 datasets.

3.2 Implementation details

The MSSEG and MSSEG-2 datasets were pre-processed using rigid registration, bias correction, denoising, and skull stripping using the official code provided by the event organizers. For training the segmentation model and self-supervised pre-training, patches of size $96 \times 96 \times 96$ were used, and voxel intensity values were normalized by subtracting the mean and dividing by the standard deviation of each MRI. A batch size of four was used with an initial learning rate of $1e^{-4}$ with AdamW [17] as the optimizer and models were trained for 500 epochs. Python 3.9.7, Pytorch 2.1.2 and MONAI [6] were utilized to implement models. We chose $\alpha = 10$ for the boundary loss.

3.3 MS lesion segmentation

To report segmentation results, besides common segmentation metrics, including the Dice and Hausdorff Distance (HD) [19], the F1-score was also calculated, indicating how many new lesions have been (in)correctly detected independent of the precision of the contour [8]. Here, small lesions, smaller than 3 mm^3, were

Table 1. Comparison between our proposed method and alternative approaches. All results are expressed as mean ± std. Ablation study results are also provided.

Methods		**Dice**(%) ↑	**HD**(mm) ↓	**F1**(%) ↑
Coact [25]		54.09±8.71	44.20±8.47	53.39±6.01
SNAC [7]		52.05±13.16	50.12±15.50	45.94±13.71
Pre-activation UNet [2]		45.6±9.5	40.1±13.2	51.9±11.3
Our method		**56.15±7.06**	**37.13±13.29**	**56.69±9.12**
Ablation	without pre-training	53.98±7.6	42.15±15.29	49.53±8.91
	without consistency loss	51.25±8.6	45.12±8.50	50.36±11.71
	without boundary loss	55.56±8.2	38.25±15.36	56.01±7.55
	without dataset-provided lesions	52.36±6.55	54.37±10.37	51.95±6.34

excluded from F1 computations [8]. Comparison results in Table 1 indicate that our proposed method outperformed previously proposed methods.

Figure 4 qualitatively compares the output segmentation of our method, and Coact [25]. Having been trained on lesions of diverse shapes, sizes, and locations, our model exhibited a more robust performance, particularly with tiny, low-contrast lesions. To further explore the effectiveness of the self-supervised pre-

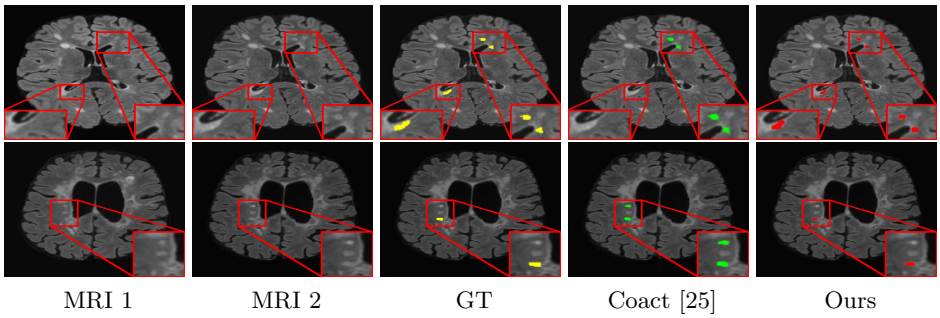


Fig. 4. Comparative visualization of predictions by the proposed model and Coact [25].

training model, consistency loss, boundary loss, and synthetic lesions, ablation studies were performed, and the results are provided in Table 1. As shown, 1) removing self-supervised pre-training led to the largest decrease, with a 7.16% drop in F1 score, and a 2.17% reduction in Dice, 2) adding consistency loss had the highest impact on Dice, with an increase of 4.9%, 3) adding boundary loss slightly improved Dice, HD and F1 by 0.59%, 1.12 mm, and 0.68%, respectively, 4) utilizing synthetic lesions, without using the dataset-provided new lesions, our model beat two of the three compared methods for Dice and F1.

4 Discussion and Conclusion

In this work, a self-supervised pre-training scheme was proposed to enhance the performance of segmenting new lesions between the baseline and follow-up MRI scans. To combat data scarcity challenges, a simple data augmentation strategy was proposed, to imitate the shape, size, and location of lesions. Experiments on the MSSEG-2 dataset showed that both proposed methods significantly improve segmentation accuracy. A limitation of this work in the context of clinical applications is that only newly appeared lesions between two MRIs were segmented because only new lesions were labelled in the dataset. However, growing and shrinking lesions must also be addressed as they have a clinical impact on monitoring MS progression, suggesting a future research direction.

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

References

1. Altay, E.E., Fisher, E., Jones, S.E., Hara-Cleaver, C., Lee, J.C., Rudick, R.A.: Reliability of classifying multiple sclerosis disease activity using magnetic resonance imaging in a multiple sclerosis clinic. JAMA neurology **70**(3), 338–344 (2013)
2. Ashtari, P., Barile, B., Van Huffel, S., Sappey-Marinier, D.: New multiple sclerosis lesion segmentation and detection using pre-activation u-net. Frontiers in Neuroscience **16**, 975862 (2022)
3. Atito, S., Awais, M., Kittler, J.: Sit: Self-supervised vision transformer. arXiv preprint arXiv:2104.03602 (2021)
4. Basaran, B.D., Matthews, P.M., Bai, W.: New lesion segmentation for multiple sclerosis brain images with imaging and lesion-aware augmentation. Frontiers in Neuroscience **16**, 1007453 (2022)
5. Basaran, B.D., Qiao, M., Matthews, P.M., Bai, W.: Subject-specific lesion generation and pseudo-healthy synthesis for multiple sclerosis brain images. In: International Workshop on Simulation and Synthesis in Medical Imaging. pp. 1–11. Springer (2022)
6. Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., et al.: Monai: An open-source framework for deep learning in healthcare. arXiv preprint arXiv:2211.02701 (2022)
7. Commowick, O., Cervenansky, F., Cotton, F., Dojat, M.: Multiple sclerosis new lesions segmentation challenge using a data management and processing infrastructure. MICCAI 2021 MSSEG-2 Challenge Proceedings (2021)
8. Commowick, O., Istace, A., Kain, M., Laurent, B., Leray, F., Simon, M., Pop, S.C., Girard, P., Ameli, R., Ferré, J.C., et al.: Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure. Scientific reports **8**(1), 13650 (2018)
9. Commowick, O., Kain, M., Casey, R., Ameli, R., Ferré, J.C., Kerbrat, A., Tourdias, T., Cervenansky, F., Camarasu-Pop, S., Glatard, T., et al.: Multiple sclerosis lesions segmentation from multiple experts: The miccai 2016 challenge dataset. Neuroimage **244**, 118589 (2021)

10. Doga Basaran, B., Zhang, W., Qiao, M., Kainz, B., Matthews, P.M., Bai, W.: Lesionmix: A lesion-level data augmentation method for medical image segmentation. arXiv e-prints pp. arXiv–2308 (2023)
11. E Woods, R., C Gonzalez, R.: Digital image processing (2008)
12. Fang, I.S., Wen, H.C., Hsu, C.L., Jen, P.C., Chen, P.Y., Chen, Y.S.: Es3net: Accurate and efficient edge-based self-supervised stereo matching network. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. pp. 4472–4481 (June 2023)
13. Fischl, B.: FreeSurfer. Neuroimage **62**(2), 774–781 (Jan 2012)
14. Ghasemi, N., Razavi, S., Nikzad, E.: Multiple sclerosis: pathogenesis, symptoms, diagnoses and cell-based therapy. Cell Journal (Yakhteh) **19**(1), 1 (2017)
15. Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., Roth, H.R., Xu, D.: Unetr: Transformers for 3d medical image segmentation. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision. pp. 574–584 (2022)
16. Kervadec, H., Bouchtiba, J., Desrosiers, C., Granger, E., Dolz, J., Ayed, I.B.: Boundary loss for highly unbalanced segmentation. In: International conference on medical imaging with deep learning. pp. 285–296. PMLR (2019)
17. Loshchilov, I., Hutter, F.: Decoupled weight decay regularization. arXiv preprint arXiv:1711.05101 (2017)
18. Ma, J., Wei, Z., Zhang, Y., Wang, Y., Lv, R., Zhu, C., Gaoxiang, C., Liu, J., Peng, C., Wang, L., et al.: How distance transform maps boost segmentation cnns: an empirical study. In: Medical Imaging with Deep Learning. pp. 479–492. PMLR (2020)
19. Maier-Hein, L., Menze, B., et al.: Metrics reloaded: Pitfalls and recommendations for image analysis validation. arXiv. org (2206.01653) (2022)
20. Milletari, F., Navab, N., Ahmadi, S.A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 fourth international conference on 3D vision (3DV). pp. 565–571. Ieee (2016)
21. Otsu, N.: A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics **9**(1), 62–66 (1979)
22. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. pp. 234–241. Springer (2015)
23. Tang, Y., Yang, D., Li, W., Roth, H.R., Landman, B., Xu, D., Nath, V., Hatamizadeh, A.: Self-supervised pre-training of swin transformers for 3d medical image analysis. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 20730–20740 (2022)
24. Walton, C., King, R., Rechtman, L., Kaye, W., Leray, E., Marrie, R.A., Robertson, N., La Rocca, N., Uitdehaag, B., van Der Mei, I., et al.: Rising prevalence of multiple sclerosis worldwide: Insights from the atlas of ms. Multiple Sclerosis Journal **26**(14), 1816–1821 (2020)
25. Wu, Y., Wu, Z., Shi, H., Picker, B., Chong, W., Cai, J.: Coactseg: Learning from heterogeneous data for new multiple sclerosis lesion segmentation. In: International conference on medical image computing and computer-assisted intervention. pp. 3–13. Springer (2023)
26. Zhang, X., Liu, C., Ou, N., Zeng, X., Zhuo, Z., Duan, Y., Xiong, X., Yu, Y., Liu, Z., Liu, Y., et al.: Carvemix: a simple data augmentation method for brain lesion segmentation. NeuroImage **271**, 120041 (2023)