Subject-Specific Lesion Generation and Pseudo-Healthy Synthesis for Multiple Sclerosis Brain Images

Berke Doga Basaran[1,2(✉)][0000−0001−7235−3962], Mengyun Qiao[2][0000−0002−5157−1079], Paul M. Matthews[3,4][0000−0002−1619−8328], and Wenjia Bai[1,2,3][0000−0003−2943−7698]

[1] Department of Computing, Imperial College London, London, UK
[2] Data Science Institute, Imperial College London, London, UK
bdb19@imperial.ac.uk
[3] Department of Brain Sciences, Imperial College London, London, UK
[4] UK Dementia Research Institute, Imperial College London, London, UK

Abstract. Understanding the intensity characteristics of brain lesions is key for defining image-based biomarkers in neurological studies and for predicting disease burden and outcome. In this work, we present a novel foreground-based generative method for modelling the local lesion characteristics that can both generate synthetic lesions on healthy images and synthesize subject-specific pseudo-healthy images from pathological images. Furthermore, the proposed method can be used as a data augmentation module to generate synthetic images for training brain image segmentation networks. Experiments on multiple sclerosis (MS) brain images acquired on magnetic resonance imaging (MRI) demonstrate that the proposed method can generate highly realistic pseudo-healthy and pseudo-pathological brain images. Data augmentation using the synthetic images improves the brain image segmentation performance compared to traditional data augmentation methods as well as a recent lesion-aware data augmentation technique, CarveMix. The code will be released at `https://github.com/dogabasaran/lesion-synthesis`.

Keywords: Lesion generation · Image synthesis · Generative modelling · Data augmentation · Attention mechanisms.

1 Introduction

Multiple sclerosis (MS) is a demyelinating neurodegenerative disease of the nervous system [9]. On brain magnetic resonance (MR) images such as fluid-attenuated inversion recovery (FLAIR), MS is often manifested as hyperintense lesions distributed mostly in the white matter [1]. Identification and understanding the intensity patterns of these lesions thus plays a crucial role in diagnosing and tracking the progression of the disease. Being able to model the lesion intensity patterns also enables us to generate synthetic images for data augmentation

purposes and to perform counter-factual inference to understand disease progression. Although many efforts have been dedicated in the past to the general image synthesis problem [11,14,18], little has been investigated for subject-specific MS lesion synthesis.

In this work, we propose a novel generative method to model the MS lesion characteristics. It is able to perform subject-specific lesion synthesis given a healthy image, as well as synthesise a pseudo-healthy image given a pathological one. An attention mechanism is designed so as to only modify the intensities in the lesion region while maintaining the structure in other brain regions. We demonstrate that the proposed method achieves high realism in both lesion generation and pseudo-healthy synthesis. In addition, using the synthetic images for training image segmentation networks, we can improve the lesion segmentation performance even in a low-data setting.

1.1 Related Works

Brain Lesion Synthesis Learning-based brain lesion synthesis has received a lot of attention recently [3,22]. Salem et al. uses an encoder-decoder U-Net structure to fuse lesion masks acquired from MS patients to healthy subjects [20]. Bissoto et al., Jin et al., and Li et al. utilise generative adversarial networks (GANs) to synthesize new image samples [2,13,15]. Yet, [2] and [15] use semantic maps from pathological subjects to generate images, which lead to augmented images which have identical foreground shapes and labels to the samples that they are drawn from. [13] uses a user-specified tumour mask to generate synthetic tumours on healthy scans. These methods do not generate diverse data samples with high foreground variance (lesion/tumour masks) which differ from already present samples in the training dataset. Reinhold et al. employs a casual model using variational autoencoders to generate lesions with predefined lesion load [19]. As for subject-specific pathological to healthy image synthesis, known as pseudo-healthy synthesis, Xia et al. uses a framework that consists of a generator, a segmentor and a reconstructor, along with a mask discriminator which distinguishes segmented masks from pathology masks [24].

Data Augmentation With the assistance of data augmentation methods, deep learning has defined the state-of-the-art for lesion segmentation [26]. Data augmentation has improved model generalisability by increasing dataset variance, and has alleviated some of the issues caused due to scarcity of labelled data. However, traditional data augmentation methods, which include flipping, rotating, scaling, intensity alterations, and elastic deformations, do not significantly alter the original image, e.g. the appearance of lesions. Zhang et al. proposed CarveMix, derived from CutMix [25], which utilises a lesion-aware mix-based method to combine carved lesion patches with non-lesion images [27]. Both of these methods directly employ foreground masks from other subjects to compose an augmented image with already present lesion shapes and characteristics. Current data augmentation methods for brain lesions do not drastically change the

local features of lesions, and a technique which can realistically generate new lesions from a learnt distribution will increase training sample diversity and segmentation performance.

1.2 Contributions

We present a novel brain lesion synthesis method for multiple sclerosis images. Our contributions are four-fold: (1) Differing from other techniques, our method is able to preserve subject identity and generate a new lesion foreground class without directly sampling a lesion mask from another subject. (2) An attention-based foreground discriminator is introduced to generate realistic images focusing at the lesion region. (3) The method is able to generate both synthetic lesions as well as pseudo-healthy images, achieved via a cyclic structure. This enables generating a diverse set of brain images for data augmentation purposes. (4) Our method is stochastic, and is able to generate multiple augmented images with different lesion appearances from a single input image.

2 Methods

The goal is to generate pathological images with lesions, given healthy images as input, while preserving the subject identity, and vice versa, create pseudo-healthy images without the lesions in the pathological subjects. Inspired by CycleGAN [28] and AttentionGAN [23], we adopt a two-branch architecture to learn both the lesion generation and pseudo-healthy synthesis tasks. We employ the attention and content decoders from [23], which provide maps on *where* and *what* the important features of the foreground and background classes of the image are, respectively. However, we encourage the foreground decoder to learn lesion specific features. In particular, we introduce a foreground-based discriminator to focus on the generation of the lesion region. We assume that two unpaired datasets are available for learning: a healthy dataset, $X_H = \{x_H\}$, and a pathological dataset with lesions or white matter hyperintensities (WMHs), $X_P = \{x_P\}$. Lower case x_H and x_P denote single image samples from these datasets. For the pathological dataset, a corresponding foreground dataset, $X_F = \{x_F\}$, is constructed using the lesion masks. Each foreground image is a masked lesion intensity image, generated by multiplying the original pathological image from X_P with its respective lesion mask.

Figure 1 illustrates the proposed framework, which consists of two branches, one generator, G_P, for synthetic lesion generation and the other generator, G_H, for pseudo-healthy synthesis. We represent the lesion generation branch as $H \rightarrow P$, and the pseudo-healthy synthesis branch as $P \rightarrow H$, where H denotes healthy images, and P denotes pathological images. The lesion generator, G_P, takes in a healthy brain image, x_H, and generates a pathological image with lesions, $\hat{x}_P$, while also aiming to preserve subject identity. It is composed of an encoder, Enc, which maps the input image into a latent code z, a content decoder, Dec_C, that decodes z into a foreground content map, C^{fore}, and an

attention decoder, Dec_A, that decodes z into the two attention maps, namely A^{fore} for foreground and A^{back} for background. Subsequently, the foreground content, C^{fore}, in our case the lesion-related regions, and attention map, A^{fore}, are fused to generate a foreground output, O^{fore}. The input image, x_H, and background attention, A^{back}, is fused to generate a background output, O^{back}. The summation of O^{fore} and O^{back} produces the generated synthetic image, $\hat{x}_P$. The pseudo-healthy synthesis branch, G_H, takes an image with lesions, x_P, as input and produces an pseudo-healthy image, $\hat{x}_H$. This branch is constructed in a similar way as the lesion generation branch.

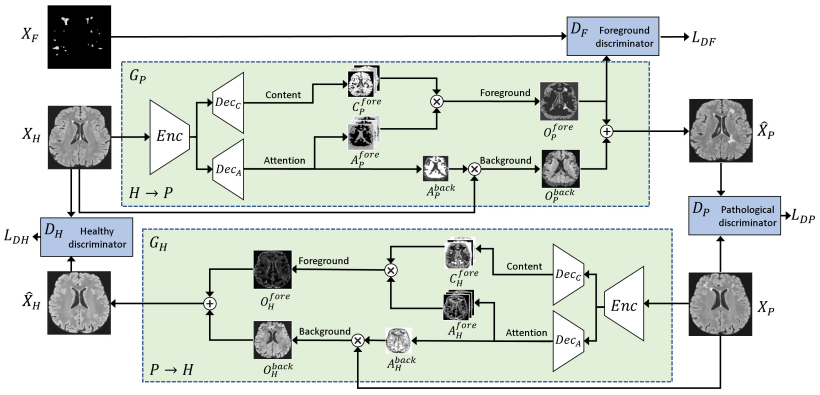

Fig. 1. Proposed framework for subject-specific lesion generation and pseudo-healthy synthesis of brain images, consisting of two generators, G_P and G_H. See text for detail.

2.1 Generators

The two generators, G_P and G_H, are both constructed using the ResNet architecture with nine blocks. Each generator features an encoder with three downsampling paths, and two decoders, one for generating content masks, and one for attention masks. For a given input image, our model generates n attention masks, 1 for background attention, A^{back}, and $n-1$ for foreground attention, A^{fore}, and also generates $n-1$ foreground content masks, C^{fore}. Here foreground broadly refers to regions relevant to lesions and background refers to the rest of the brain image. Producing multiple foreground attention masks allows lesions of varying shape and location characteristics to be decoded, thus producing synthetic images with different lesion distributions. Similar ideas have been explored in [5], where an input image is disentangled into multiple channels representing different anatomical structures. The synthetic image $G(x)$ is generated as,

$$G(x) = C^{fore} * A^{fore} + x * A^{back} \tag{1}$$

which combines the foreground content, C^{fore}, multiplied with their respective attention masks, A^{fore}, and the original input image, x, multiplied with the

background attention mask, A^{back}. The separation of foreground and background allows us to explicitly focus on the lesions in the generating process by developing a lesion-aware foreground discriminator.

2.2 Discriminators

To generate realistic brain images, three discriminators are introduced. The pathological discriminator, D_P, distinguishes synthetic images with lesions from real pathological images. The healthy discriminator, D_H, distinguishes pseudo-healthy images from real healthy subjects. The third discriminator, D_F, which is lesion-aware, encourages the foreground (synthetic lesions) to look realistic compared to real lesion masks in the pathological dataset.

2.3 Losses

The total generator loss, $\mathcal{L}_{G_{total}}$, consists of three loss terms: generation loss, $\mathcal{L}_G$, cycle-consistency loss, $\mathcal{L}_{CC}$, and identity loss, $\mathcal{L}_{idt}$. We employ the least squares GAN loss [17] for the generation loss, also known as $\mathcal{L}_2$ loss. For pseudo-healthy synthesis ($P \rightarrow H$), the generation loss is formulated as,

$$\mathcal{L}_{G_H} = \mathbb{E}_{x \sim X_P} \left[(D_H(G_H(x)) - 1)^2 \right] \tag{2}$$

where D_H denotes the healthy discriminator for the pseudo-healthy image $G_H(x)$. For lesion generation ($H \rightarrow P$), the generation loss is formulated as,

$$\mathcal{L}_{G_P} = \mathbb{E}_{x \sim X_H} \left[\left(\tfrac{1}{2}(D_P(G_P(x)) + D_F(O_P^{fore})) - 1 \right)^2 \right] \tag{3}$$

where D_P denotes the pathological discriminator for the lesion image $G_P(x)$ and D_F denotes the foreground discriminator for the foreground output O_P^{fore}, which penalises non-realistic lesion regions.

The cycle-consistency loss, $\mathcal{L}_{CC}$, evaluates a full cycle of an image through the network, $P \rightarrow H \rightarrow P$ or $H \rightarrow P \rightarrow H$. We encourage the output of these cycles to be equivalent to the input as we wish to preserve subject identity. This also ensures that the generated synthetic lesions do not appear at unrealistic locations. The cycle-consistency loss, $\mathcal{L}_{CC}$, is formulated as,

$$\mathcal{L}_{CC} = \mathbb{E}_{x \sim X_H} \left[\|G_H(G_P(x)) - x\|_1 \lambda_H \right] + \mathbb{E}_{x \sim X_P} \left[\|G_P(G_H(x)) - x\|_1 \lambda_P \right] \tag{4}$$

Finally, we incorporate the identity loss, $\mathcal{L}_{idt}$, with the objective of getting an identical image out of a generator if its healthiness is not changed, such that $G_H(x_H)$ would return an image identical to x_H. We use the $\mathcal{L}_1$ loss for this task.

$$\mathcal{L}_{idt} = \mathbb{E}_{x \sim X_H} \left[\|G_H(x) - x\|_1 \lambda_H \lambda_{idt} \right] + \mathbb{E}_{x \sim X_P} \left[\|G_P(x) - x\|_1 \lambda_P \lambda_{idt} \right] \tag{5}$$

We utilise hyperparameters λ_H, λ_P, and λ_{idt} for weighting these losses. The total loss for the generator is summarized as,

$$\mathcal{L}_{G_{total}} = \mathcal{L}_{G_H} + \mathcal{L}_{G_P} + \mathcal{L}_{CC} + \mathcal{L}_{idt} \tag{6}$$

The discriminators aim to minimize the sum of the squared difference between predicted and expected values for real and synthetic images. We use the least squares loss for the three discriminators, D_H, D_P, D_F, formulated as,

$$\mathcal{L}_{D_H} = \mathbb{E}\left[(D_H(x_H) - 1)^2 + (D_H(G_H(x_P)))^2\right] \tag{7}$$

$$\mathcal{L}_{D_P} = \mathbb{E}\left[(D_P(x_P) - 1)^2 + (D_P(G_P(x_H)))^2\right] \tag{8}$$

$$\mathcal{L}_{D_F} = \mathbb{E}\left[(D_F(x_F) - 1)^2 + (D_F(O_P^{fore}))^2\right] \tag{9}$$

3 Experiments

3.1 Evaluation

We evaluate the method performance in data augmentation using the synthetic data for downstream segmentation task as well as evaluate the image quality of synthetic data. To assess performance we conduct two experiments, employing our method as a data augmentation tool and comparing segmentation performance when trained on only synthetic data. To evaluate quality we ask three raters to assess healthiness and realism of generated images.

The proposed method is compared with traditional data augmentation (TDA) techniques which come default with nnU-Net [12], including rotation, scaling, mirroring, elastic deformation, intensity perturbation, and simulation of low resolution. We also compare to a state-of-the-art lesion-aware augmentation method, CarveMix [27]. We utilise offline versions of our method and CarveMix to double the training dataset for those experiments. When performing data augmentation using the proposed method, lesion masks are required for the foreground discriminator. After lesion generation, we perform free-form deformation registration of the healthy subject image, x_H, to the corresponding synthesized pathological image, $\hat{x}_P$, threshold the difference of the two images and generate the lesion masks. We report the mean and variance of the Dice scores of the segmentation.

3.2 Implementation

GAN Network We implement the generative model with a single channel input and single channel output. We set the generated mask number n to 10, as our experiments show more variety of realistic lesion characteristics and intensities are synthesised with a higher n. Our generative framework method also features data augmentation in the form of sagittal mirroring and elastic deformation.

Training The proposed method is developed on PyTorch and trained on one NVIDIA GeForce RTX 3080. We use the Adam optimizer with an initial learning rate of 0.001 and 0.5 dropout. The network is trained for 400 epochs and with a linear decay of the learning rate after 200 epochs with a batch size of 1, due to GPU memory limit. We set the generator loss trade-off hyperparameters λ_H, λ_P, and λ_{idt} to 10, 10, and 0.5, respectively. For evaluating segmentation performance, nnU-Net is trained for 1,000 epochs using stochastic gradient descent.

3.3 Data

A private dataset is used for training the proposed method, which consists of FLAIR MRI scans of 120 subjects, including 60 healthy scans and 60 scans with MS lesions. For evaluation, we train the segmentation networks using the 2016 Multiple Sclerosis Lesion Segmentation dataset (MS2016) [6], and test on the 2008 MICCAI MS Lesion Segmentation dataset (MS2008) [21] and 2015 ISBI Longitudinal MS Lesion Segmentation dataset (ISBI2015) [4]. We resample all images into $1 \times 1 \times 1$ mm^3 voxel spacing and extract the centre axial slice of each image. These images are reshaped into 256×256 dimensions to pass through the network. The private, MS2008, ISBI2015, and MS2016 datasets provide 120, 20, 21, and 15 2D FLAIR images, respectively.

3.4 Results

Data Augmentation Performance We train twelve separate segmentation models using different dataset sizes and data augmentation methods. We use nnU-Net in 3D full resolution, utilizing 100%, 53.3%, 26.6% and 12.3% of the training dataset, corresponding to 15, 8, 4, and 2 subjects.

Table 1. Mean and standard deviations of the Dice scores (%), at different sizes of training data. Best results are in bold. Asterisks indicate statistical significance (*: p$\leq$ 0.05, **: p $\leq$ 0.001) when using a paired Student's t-test compared to baselines.

Size	TDA		TDA+CarveMix [27]		TDA+Proposed	
	MS2008	ISBI2015	MS2008	ISBI2015	MS2008	ISBI2015
100%	$33.06^{**}_{17.37}$	$63.20_{16.70}$	$32.32^{**}_{15.23}$	$57.82_{16.56}$	$\mathbf{44.77_{15.87}}$	$\mathbf{63.90_{16.45}}$
53.3%	$34.03_{17.83}$	$58.84_{24.54}$	$35.39_{17.36}$	$61.25_{19.06}$	$\mathbf{39.70_{18.44}}$	$\mathbf{62.35_{13.21}}$
26.6%	$31.55_{18.62}$	$\mathbf{64.43_{12.64}}$	$\mathbf{39.21_{18.86}}$	$60.28_{21.22}$	$34.58_{15.47}$	$60.83_{10.03}$
13.3%	$28.96_{23.55}$	$37.34^{**}_{23.81}$	$23.57*_{23.18}$	$38.94^{**}_{21.81}$	$\mathbf{32.29_{16.16}}$	$\mathbf{56.64_{13.23}}$

Table 1 indicates that the average Dice score is increased using the proposed method in three of the four dataset sizes when compared to TDA and CarveMix, while also decreasing the standard deviation in most cases. Figure 2 demonstrates different stages of the proposed method and augmentation examples from CarveMix. [27]. The generator of the method is stochastic, which employs dropout to generate multiple diverse augmented images for each given sample (Supple. Fig. 2). This increases the diversity of training image samples. We report the mean and variance of the Hausdorff distances in Supple. Table 1.

Synthetic data segmentation performance We compare segmentation performance when training a model using a fully synthetic dataset against a model trained on a real dataset. We generate 15 synthetic images from the MS2016 dataset and test on the MS2008 and ISBI2015 datasets. We compare our results

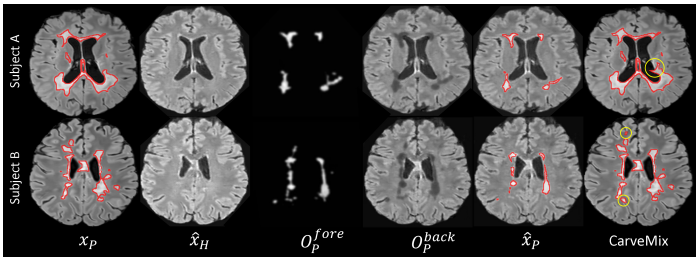

Fig. 2. Different stages of the proposed method. Lesions are contoured in red. **Left to right**: Real pathological image, x_P; pseudo-healthy image $\hat{x}_H$; foreground output for lesion synthesis on pseudo-healthy image O_P^{fore}; background output for lesion synthesis O_P^{back}; the synthetic pathological image generated from the pseudo-healthy image $\hat{x}_P$; augmented image using CarveMix, unrealistic lesions circled in yellow.

to nnU-Net with TDA at 100% dataset size. Table 2 suggests a model trained on our generated data performs comparable to a model trained on a real dataset.

Table 2. Segmentation performance using fully synthetic data for model training, compared to using real data. Mean and standard deviations of the Dice scores (%) are reported with best results in bold.

Test data	Real training data	Synthetic training data
MS2008	$33.06_{17.37}$	$\mathbf{35.74_{19.85}}$
ISBI2015	$\mathbf{63.20_{16.70}}$	$52.72_{18.61}$

Expert Evaluation In order to assess quality of generated images, we ask three human raters to evaluate images in healthiness and realism. We define "healthiness" as a brain MR image with no visible pathological features, and "realism" as an image which possesses the same quality and authentic look as one obtained by a scanner. We collate 50 real healthy, 50 real pathological, 50 pseudo-healthy, and 50 pseudo-pathological images. Raters are asked to classify each image as healthy (1) or pathological (0), and real (1) or fake (0). We produce the results for the real images as a benchmark to compare our generated images. Table 3 shows that the generated pseudo-healthy and pseudo-pathological images closely follow real images for the assessed metrics.

Correlation with Radiological Findings The spatial heatmap of the synthetic lesions closely resemble those of the real lesions from clinical data, including the private dataset and the MS2016 dataset (Supple. Fig. 3). Furthermore, generated pseudo-pathological images show an increase in the ventricular volume, while pseudo-healthy images tend to have a decreased ventricular volume

Table 3. Rater classification of real and generated images on realism and healthiness. Asterisk (*) indicates that a lower score is better.

Realism		
	Real	Proposed
Healthy images	85%	82%
Pathological images	84%	82%

Healthiness		
	Real	Proposed
Healthy images	85%	83%
Pathological images*	5%	1%

(Supple. Fig. 4). The change in the ventricular volume is consistent with literature [7, 8, 10, 16].

4 Conclusion

We have proposed a novel method for generating lesions and pseudo-healthy images while preserving subject-specific features, providing a useful tool for data augmentation in brain image analysis to complement real patient datasets. It is lesion-aware in the generating process. In quantitative evaluation, the method improves lesion segmentation performance in downstream tasks. In our human rater assessment, the generated images achieved a high realism score close to the real images, and generated lesions follow a spatial distribution consistent with real clinical datasets.

Acknowledgements This work is supported by the UKRI CDT in AI for Healthcare http://ai4health.io (Grant No. EP/S023283/1). For the purpose of open access, the author has applied a 'Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

References

1. Bakshi, R., Ariyaratana, S., Benedict, R.H.B., Jacobs, L.: Fluid-Attenuated Inversion Recovery Magnetic Resonance Imaging Detects Cortical and Juxtacortical Multiple Sclerosis Lesions. Archives of Neurology **58**(5), 742–748 (2001). https://doi.org/10.1001/archneur.58.5.742
2. Bissoto, A., Perez, F., Valle, E., Avila, S.: Skin Lesion Synthesis with Generative Adversarial Networks. In: OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis. pp. 294–302. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-030-01201-4_32
3. Bowles, C., Qin, C., Guerrero, R., et al.: Brain lesion segmentation through image synthesis and outlier detection. NeuroImage: Clinical **16**, 643–658 (2017). https://doi.org/10.1016/j.nicl.2017.09.003
4. Carass, A., Roy, S., Jog, A., et al.: Longitudinal multiple sclerosis lesion segmentation: Resource and challenge. NeuroImage **148**, 77–102 (2017). https://doi.org/10.1016/j.neuroimage.2016.12.064

5. Chartsias, A., Joyce, T., Papanastasiou, G., et al.: Disentangled representation learning in cardiac image analysis. Medical Image Analysis **58**, 101535 (2019). https://doi.org/10.1016/j.media.2019.101535
6. Commowick, O., Istace, A., Kain, M., et al.: Objective evaluation of multiple sclerosis lesion segmentation using a data management and processing infrastructure. Scientific Reports **8**(1), 13650 (Sep 2018). https://doi.org/10.1038/s41598-018-31911-7
7. Dalton, C.M., Brex, P.A., Jenkins, R., et al.: Progressive ventricular enlargement in patients with clinically isolated syndromes is associated with the early development of multiple sclerosis. Journal of Neurology, Neurosurgery & Psychiatry **73**(2), 141–147 (2002). https://doi.org/10.1136/jnnp.73.2.141
8. Dalton, C.M., Miszkiel, K.A., O'Connor, P.W., et al.: Ventricular enlargement in MS. Neurology **66**(5), 693–698 (2006). https://doi.org/10.1212/01.wnl.0000201183.87175.9f
9. Ghasemi, N., Razavi, S., Nikzad, E.: Multiple Sclerosis: Pathogenesis, Symptoms, Diagnoses and Cell-Based Therapy Citation: Ghasemi N, Razavi Sh, Nikzad E. Multiple sclerosis: pathogenesis, symptoms, diagnoses and cell-based therapy. Cell Journal **19**(191), 1–10 (2017). https://doi.org/10.22074/cellj.2016.4867
10. Guptha, S.H., Holroyd, E., Campbell, G.: Progressive lateral ventricular enlargement as a clue to Alzheimer's disease. The Lancet **359**(9322), 2040 (2002). https://doi.org/10.1016/S0140-6736(02)08806-2
11. Huang, H., Yu, P.S., Wang, C.: An introduction to image synthesis with generative adversarial nets (2018). https://doi.org/10.48550/ARXIV.1803.04469
12. Isensee, F., Jaeger, P.F., Kohl, S.A.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 (Feb 2021). https://doi.org/10.1038/s41592-020-01008-z
13. Jin, Q., Cui, H., Sun, C., Meng, Z., Su, R.: Free-form tumor synthesis in computed tomography images via richer generative adversarial network. Knowledge-Based Systems **218**, 106753 (2021). https://doi.org/10.1016/j.knosys.2021.106753
14. Karras, T., Laine, S., Aila, T.: A style-based generator architecture for generative adversarial networks. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 4396–4405 (2019). https://doi.org/10.1109/CVPR.2019.00453
15. Li, Q., Yu, Z., Wang, Y., Zheng, H.: Tumorgan: A multi-modal data augmentation framework for brain tumor segmentation. Sensors **20**(15) (2020). https://doi.org/10.3390/s20154203
16. Luxenberg, J.S., Haxby, J.V., Creasey, H., Sundaram, M., Rapoport, S.I.: Rate of ventricular enlargement in dementia of the Alzheimer type correlates with rate of neuropsychological deterioration. Neurology **37**(7), 1135 (1987). https://doi.org/10.1212/WNL.37.7.1135
17. Mao, X., Li, Q., Xie, H., et al.: Least squares generative adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV). pp. 2813–2821 (2017). https://doi.org/10.1109/ICCV.2017.304
18. Reed, S., Akata, Z., Yan, X., et al.: Generative adversarial text to image synthesis. In: Proceedings of The 33rd International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 48, pp. 1060–1069. PMLR, New York, New York, USA (20–22 Jun 2016)
19. Reinhold, J.C., Carass, A., Prince, J.L.: A Structural Causal Model for MR Images of Multiple Sclerosis. In: Medical Image Computing and Computer Assisted Inter-

vention – MICCAI 2021. pp. 782–792. Springer International Publishing, Cham (2021). https://doi.org/10.1007/978-3-030-87240-3_75
20. Salem, M., Valverde, S., Cabezas, M., et al.: Multiple Sclerosis Lesion Synthesis in MRI Using an Encoder-Decoder U-NET. IEEE Access **7**, 25171–25184 (2019). https://doi.org/10.1109/ACCESS.2019.2900198
21. Styner, M., Lee, J., Chin, B., et al.: 3d segmentation in the clinic: A grand challenge ii: Ms lesion segmentation (11 2008). https://doi.org/10.54294/lmkqvm
22. Sun, L., Wang, J., Huang, Y., et al.: An adversarial learning approach to medical image synthesis for lesion detection. IEEE Journal of Biomedical and Health Informatics **24**(8), 2303–2314 (2020). https://doi.org/10.1109/JBHI.2020.2964016
23. Tang, H., Xu, D., Sebe, N., Yan, Y.: Attention-guided generative adversarial networks for unsupervised image-to-image translation. In: 2019 International Joint Conference on Neural Networks (IJCNN). pp. 1–8 (2019). https://doi.org/10.1109/IJCNN.2019.8851881
24. Xia, T., Chartsias, A., Tsaftaris, S.A.: Pseudo-healthy synthesis with pathology disentanglement and adversarial learning. Medical Image Analysis **64**, 101719 (2020). https://doi.org/10.1016/j.media.2020.101719
25. Yun, S., Han, D., Chun, S., et al.: Cutmix: Regularization strategy to train strong classifiers with localizable features. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 6022–6031 (2019). https://doi.org/10.1109/ICCV.2019.00612
26. Zeng, C., Gu, L., Liu, Z., Zhao, S.: Review of Deep Learning Approaches for the Segmentation of Multiple Sclerosis Lesions on Brain MRI. Frontiers in Neuroinformatics **14** (2020). https://doi.org/10.3389/fninf.2020.610967
27. Zhang, X., Liu, C., Ou, N., et al.: CarveMix: A Simple Data Augmentation Method for Brain Lesion Segmentation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. pp. 196–205. Springer International Publishing, Cham (2021)
28. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV). pp. 2242–2251 (2017). https://doi.org/10.1109/ICCV.2017.244