MRI-GAN: Generative Adversarial Network for Brain Segmentation

Afifa Khaled[1(✉)] and Taher A. Ghaleb[2]

[1] School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, China
afifakhaied@tju.edu.cn

[2] School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Canada
tghaleb@uottawa.ca

Abstract. Segmentation is an important step in medical imaging. In particular, machine learning, especially deep learning, has been widely used to efficiently improve and speed up the segmentation process in clinical practices of MRI brain images. Despite the acceptable segmentation results of multi-stage models, little attention was paid to the use of deep learning algorithms for brain image segmentation, which could be due to the lack of training data. Therefore, in this paper, we propose $MRI-GAN$, a Generative Adversarial Network (GAN) model that performs segmentation MRI brain images. Our model enables the generation of more labeled brain images from existing labeled and unlabeled images. Our segmentation targets brain tissue images, including white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). We evaluate the performance of the $MRI-GAN$ model using a commonly used evaluation metric, which is the Dice Coefficient (DC). Our experimental results reveal that our proposed model significantly improves segmentation results compared to the standard GAN model while taking shorter training time.

1 Introduction

The significant growth of medical imaging applications in the last decade has witnessed a matching increase in image segmentation and classification. Such growth has encouraged researchers in clinical fields to develop models that make segmentation work similar to the human process in clinical practices [1,2,28,30]. To this end, machine learning-based brain segmentation, in which brain images are divided into multiple tissues, has emerged as it makes brain image segmentation more accurate [3,4].

Many brain image segmentation models have been proposed in the literature. A common technique is to use two-stage models, which involves fusing global information with local information generated in two subsequent stages, to achieve acceptable segmentation results. The design of multi-stage models, in

B. Sheng et al. (Eds.): CGI 2023, LNCS 14495, pp. 246–256, 2024.
https://doi.org/10.1007/978-3-031-50069-5_21

general, allows achieving better results, since it helps solve the information loss problem [5–8].

There have been many studies [10–12,15,26,27] proposing techniques to improve the accuracy of brain image segmentation to reach results that are close enough to manual reference. Recently, the use of deep learning algorithms for brain image segmentation started to emerge. However, there is still a lack of available data to train deep learning models. To address such an issue, adversarial learning and few-shot learning techniques have been developed to perform well in cases where only a few labeled images are available [9,13]. For example, Mondal et al. [9] proposed a few-shot 3D multi-modal image segmentation using a GAN model, which consists of U-net, a generator, and an encoder [9]. Fake images were first generated using the generator, then used along with labeled and unlabeled data to train the discriminator, which in turn distinguishes between generated and true data. The encoder was used to compute the predicted noise mean and log-variance. Despite the merits of such a model, its achieved results were not significantly higher than previous state-of-the-art models.

While previous techniques enabled neural networks to produce acceptable segmentation output, there were very few models that address the segmentation of infant brain images into White Matter (WM), Grey Matter (GM), and Cerebrospinal Fluid (CSF). As an example, Dolz et al. [14] proposed a model to segment infant brain images, which was evaluated using the iSEG Grand MICCAI challenge dataset. The model utilized the direct connections between layers from the same and different paths, which were used to improve the learning process. However, that model did not take into consideration deeper networks with fewer filters per layer. Moreover, individual weights from dense connections were not investigated.

Therefore, in this paper, we propose $MRI-GAN$, a novel Generative Adversarial Network (GAN) model that performs segmentation of MRI brain images, particularly WM, GM, and CSF. Our model enables the generation of more labeled data from existing labeled and unlabeled data. To do this, we employ an MRI encoder with a ground truth encoder to compress the features and convert them into low-dimensional MRI and tissues vectors. Each encoder is capable of compressing one or more inputs. In summary, this paper makes the following contributions:

- Novel MRI-GAN Model: Introduces a new GAN model for segmenting brain MRI images into WM, GM, and CSF tissues.
- Data Augmentation: Enables data modeling from labeled and unlabeled data, addressing limited annotated datasets.
- Integrated Encoders: Uses MRI and ground truth encoders for efficient feature compression and vector conversion.
- Improved Accuracy: Outperforms existing methods in accurate tissue segmentation.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 presents the $MRI-GAN$ model. Section 4 presents our setup

materials and methods. Section 5 presents and discusses our experimental results. Finally, Sect. 6 concludes the paper and suggests possible future work.

2 Related Work

This section reviews the work related to our study.

2.1 Generative Adversarial Network for Brain Segmentation

GANs have shown promising results in both medical image diagnostics [20] and brain image segmentation [19,23]. The standard GAN has two parts: The generator is to generate the data and the discriminator is to distinguish between the generated data and real data. Much research on brain image segmentation has been conducted using GANs. For example, Cirillo et al. [21] proposed a 3D volume-to-volume (GAN) to segment the images of brain tumors. Their model achieved 94% result when the generator loss was weighted five times higher than the discriminator loss. The proposed model was evaluated on the *BraTS 2013* dataset. Their model outperformed previous models with an overall accuracy of 66%. Delannoy et al. [22] proposed a super-resolution and segmentation framework using GANs to neonatal brain MRI images. The framework composed of (a) a training of a generating network that estimates the corresponding high resolution (HR) image for a given input image and (b) a discriminator network D to distinguish real HR and segmentation images. Their model outperformed previous models with an overall accuracy of 83%.

2.2 Encoder/Decoder

The encoder/decoder model emerged more than a decade ago as a concept to describe an image [5]. A well-known study of encoder/decoder was the auto encoder/decoder [17], which has investigated the encoder and decoder model based on pixel-wise classification. In addition, this model enabled the use of nonlinear upsampling and a smaller number of parameters for training, which requires higher computational power than any other deep learning architectures. However, many studies that performed encoding/decoding considered mapping a dense block into a standard encoder/decoder model. We expect that applying encoder/decoder models in a GAN model will provide more accurate segmentation results for brain images. To achieve this, we first develop a new encoder-decoder model that compresses the feature of the inputs and also maps the tissues' information to the decoder. Results show that our $MRI - GAN$ model exhibits results that are fairly close to the manual reference, and a significant reduction in training time compared to the state-of-the-art models. Furthermore, the Dice coefficient is applied to better demonstrate the significance of the $MRI - GAN$ model.

3 Proposed Model

This section describes the structure of our proposed GAN model.

3.1 Encoder/Decoder

Our $MRI - GAN$ model consists of generator and discriminator. Fig. 1 shows our proposed GAN model. All the MRI encoder, ground truth encoder, tissues mapping, boundary detection network, and decoder together represent the generator of the $MRI - GAN$ model. MRI encoder and ground truth encoder take MRI image and ground truth then convert them to MRI and ground truth vectors. The detection network provides more information about the boundary. The output of the decoder is a GT image where GT denotes the image generated from the generator.

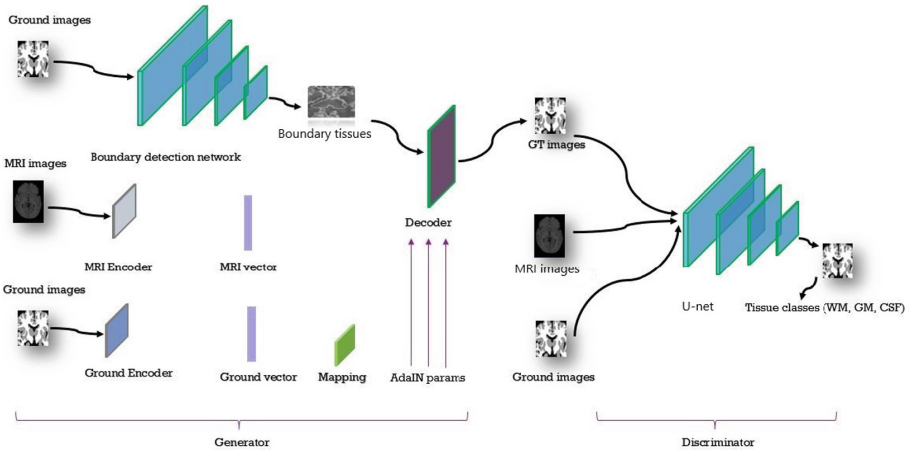

Fig. 1. Illustration of our proposed GAN model

3.2 Mapping

The decoder upscales the MRI code into a 3D geometry using SpiralBlocks that are conditioned by the ground code using Adaptive Instance normalization ($AdaIN$) [29]. Given a sample x that is passing through the network, $AdaIN$ first normalizes the activations in each channel of x to a zero μ and unit σ. The activations are then scaled on a per-channel basis. We use a mapping function R that maps a ground code y into (μ, σ) parameters for every channel of each $AdaIN$ layer. Hence the following equation:

$$AdaIN(x, y) = R_\sigma(y)\frac{x - \mu(x)}{\sigma(x)} + R_\mu(y) \tag{1}$$

where R is a learned affine function composed of multiple fully connected layers, taking the ground latent code as input. Since the AdaIN transformation operates on whole channels, the ground code alters global appearance information while the local features are determined by the MRI code.

3.3 Loss Function

Discriminator Loss Function. The discriminator in the $MRI-GAN$ model has labeled data loss, unlabeled data loss, and GT images loss (fake loss). We formulate the overall loss function of $MRI-GAN$ as follows:

$$\mathrm{l_{discriminator}} = \lambda_{\mathrm{labeled}} l_{\mathrm{labeled}} + \lambda_{\mathrm{unlabeled}} l_{\mathrm{unlabeled}} + \lambda_{\mathrm{fake}} l_{\mathrm{fake}}, \tag{2}$$

where $\lambda_{\mathrm{labeled}}$, $\lambda_{\mathrm{unlabeled}}$, and λ_{fake} are hyper-parameters. We set the hyper-parameters in Equation (2) to $\lambda_{\mathrm{labeled}} = 1.0$, $\lambda_{\mathrm{unlabeled}} = 1.0$, and $\lambda_{\mathrm{fake}} = 2.0$.

We used the proposed loss functions from Mondal et al. [9], where $Pmodel$ refers to the probability distribution of the data. More details about loss functions can be found in [9].

$$l_{\mathrm{labeled}} = -E_{\mathrm{x,y}\ \sim \mathrm{pdata(x,y)}} \sum_{\mathrm{i}=1}^{H\times W\times D} \log Pmodel(y_{\mathrm{i}}|x) \tag{3}$$

$$l_{\mathrm{unlabeled}} = -E_{\mathrm{x}\sim \mathrm{pdata(x)}} \sum_{\mathrm{i}=1}^{H\times W\times D} \log \frac{Z_i(x)}{Z_i(x)+1} \tag{4}$$

$$l_{\mathrm{fake}} = -E_{\mathrm{z}\ \sim \mathrm{noise}} \sum_{\mathrm{i}=1}^{H\times W\times D} \log \frac{1}{(Z_i(G_{\Theta G})(z))+1} \tag{5}$$

$$Z_{\mathrm{i}}(x) = \sum_{\mathrm{k}=1}^{K} exp[l_{\mathrm{i,k}}(x)] \tag{6}$$

Generator Loss Function. We propose a novel generated loss to induce G to generate real data. Let x and z denote real data and noise, respectively.

$$C = E_{\mathrm{x}\sim \mathrm{pdata(x)}} f(x) - log(1 - D(G(z))), \tag{7}$$

In our paper, we consider $f(x)$ to contain the activation of the last layer.

$$L(G) = || \ C - x \ ||_2^2, \tag{8}$$

By minimizing this loss, we force the generator to generate real data in order to match our data and the corresponding K classes of real data, which are defined as $classes = 1, ..., K$.

4 Setup Materials and Methods

This section describes the setup materials and methods used in our paper.

4.1 Datasets

MICCAI iSEG Dataset. The $MICCAIiSEG$ organizers[1] introduced a publicly available evaluation framework to allow comparing different segmentation models of WM, GM, and CSF on T1-weighted ($T1$) and T2-weighted ($T2$). The MICCAI iSEG dataset contains: 10 images (i.e., subject-1 up to subject-10), subject $T1$: $T1$-weighted image, subject $T2$: $T2$-weighted, and a manual segmentation label. All these images are used as a training set. The dataset also contains 13 images (i.e., subject-11 up to subject-23), which are used as a testing set. An example of the MICCAI iSEG dataset ($T1$, $T2$, and manual reference contour) is shown in Fig. 2.

Table 1 shows the parameters used to generate $T1$ and $T2$. The dataset has two different times: the longitudinal relaxation time and the transverse relaxation time, which are used to generate $T1$ and $T2$. The dataset has been interpolated, registered, and the images are skull-removed by the MICCAI iSEG organizers.

Table 1. Parameters used to generate $T1$ and $T2$

Parameter	TR/TE	Flip angle	Resolution
$T1$	1,900/4.38 ms	7	1×1×1
$T2$	7,380/119 ms	150	1.25×1.25×1.25

MRBrains Dataset. The $MRBrains$ dataset contains 20 adult images for the segmentation of (a) cortical gray matter, (b) basal ganglia, (c) white matter, (d) white matter lesions, (e) peripheral cerebrospinal fluid, (f) lateral ventricles, (g) cerebellum, and (h) brain stem on $T1$, $T2$, and FLAIR. Five images (i.e., 2

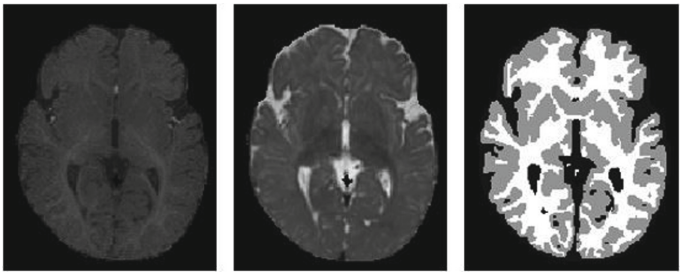

Fig. 2. An example of the MICCAI iSEG dataset ($T1$, $T2$, and manual reference contour)

[1] http://iseg2017.web.unc.edu.

male and 3 female) are provided as a training set and 15 images are provided as a testing set. For segmentation evaluation, these structures merged into gray matter ($a-b$), white matter ($c-d$), and cerebrospinal fluid ($e-f$). The cerebellum and brainstem were excluded from the evaluation.

4.2 Experimental Setup

The experiments of the proposed model were conducted using *Python* on a PC with $NVIDIAGPU$ running *Ubuntu 16.04*. Training $MRI-GAN$ took 30 hours in total, whereas testing took 5 minutes.

4.3 Segmentation Evaluation

Dice Coefficient (DC). To better highlight the significance of our proposed $MRI-GAN$ model, we use the Dice Coefficient (DC) metric to evaluate the performance of the $MRI-GAN$ model. Dice Coefficient (DC) has been used to compare state-of-the-art segmentation models. We use V_{ref} for reference segmentation and V_{auto} for automated segmentation. The DC is given by the following equation:

$$DC(V_{\text{ref}}, V_{\text{auto}}) = \frac{2V_{\text{ref}} \bigcap V_{\text{auto}}|}{|V_{\text{ref}}| + |V_{\text{auto}}|}[18], \tag{9}$$

where DC values range between [0, 1], where 1 indicates a perfect overlap and 0 indicates a complete mismatch.

5 Result and Discussion

We train and test the $MRI-GAN$ model on two datasets of different ages: adults and infants. Table 2 presents the results of the $MRI-GAN$ model to segment CSF ,GM, and WM using the MICCAI iSEG dataset. Our $MRI-GAN$ model achieves a DC value of 93% in CSF segmentation. In contrast, the DC values achieved from segmenting CSF by Standard GAN is 86%, which is 7% less accurate. In addition, our $MRI-GAN$ model achieves DC values of 94% and 92% in segmenting GM and WM, respectively. The Standard GAN model, in contrast, achieves a DC value of 80% (14% lower) for GM segmentation and 81% (11% lower) for WM segmentation. These results highlight the remarkable efficiency achieved by the $MRI-GAN$ model compared to the standard GAN.

Table 2. Dice Coefficient (DC) results of the segmentation achieved on the MICCAI-iSEG dataset. The best performance for each tissue class is highlighted in bold.

Model	Dice Coefficient (DC) Accuracy		
	CSF	GM	WM
Standard GAN	86%	80%	81%
3D, FCN + MIL+G+K [15]	94.1%	90.2%	89.7%
Multi-stage [24]	**95%**	94%	92%
Our MRI-GAN	93%	**94%**	**92%**

Table 3 presents the results achieved using the MRBrains dataset. We observe that our $MRI-GAN$ model achieves a DC value of 91% on CSF segmentation, 90% on GM segmentation, and 95% on WM segmentation. Such results are superior to the results achieved by the Standard GAN model.

Figure 3 shows a sample visualized result of our $MRI-GAN$ model on a subject used as part of the validation set. As the images show, we observe that the segmentation achieved by the $MRI-GAN$ model is fairly close to the manual reference (ground truth) contour provided by the MICCAI iSEG organizers.

Table 3. Dice Coefficient (DC) results of the segmentation achieved on the MRBrains dataset. The best performance for each tissue class is highlighted in bold.

Model	Dice Coefficient (DC) Accuracy		
	CSF	GM	WM
Standard GAN	87%	87%	85%
3D, FCN + MIL+G+K [15]	87.4%	90.6%	90.1%
Multi-stage [24]	**93%**	**93%**	88%
Our MRI-GAN	91%	90%	**95%**

Our evaluation results show that the proposed model not only outperforms two baselines (Standard GAN and 3D, FCN + MIL+G+K [15]) on the three tissues, but also attempts to outperform Multi-stage [24] on two tissues. The proposed $MRI-GAN$ model improved the results in GM and WM on the MICCAIiSEG dataset and WM on the MRBrains dataset compared with Multi-stage [24]. We acknowledge that our model may not perform well for all cases and still has limitations due to the small number of images available, which we aim to improve further in the future.

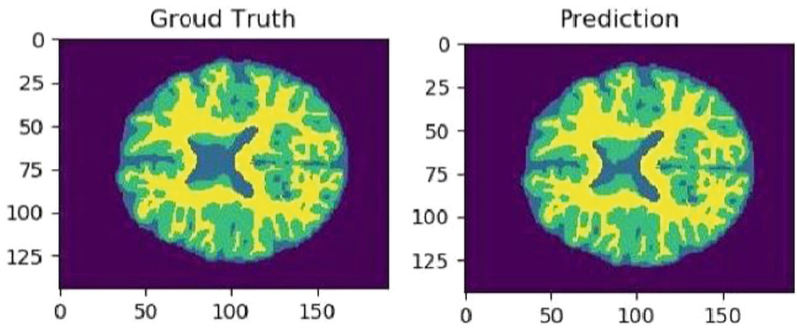

Fig. 3. A sample visualized result from the MICCAI iSEG dataset

6 Conclusion

In this paper, we proposed $MRI - GAN$, a novel Generative Adversarial Network (GAN) model that performs segmentation of MRI brain images. Our model makes segmentation more accurate by applying encoder and decoder algorithms separately, which demonstrated a significant increase in the accuracy of brain image segmentation results. We first extracted and compressed the features of the MRI encoder and ground truth encoder inputs, and then mapped the information to the decoder. Our experimental results show that the $MRI - GAN$ model is a viable solution for brain segmentation as it achieves a significant improvement in the accuracy of brain segmentation compared to the standard GAN model while taking a shorter training time.

Directions for Future Work. Based on our model, we have a number of possible directions for future work. We aim to investigate our model performance in segmenting more brain tissues and consider pathological brain images, such as with tumours or edema. Pathological brain images are not included in this study due to the lack of data.

7 Declarations

7.1 Competing Interests

The authors declare that they have no known competing financial interests.

7.2 Consent for Publication

Not applicable.

7.3 Availability of Data and Materials

The data that support the findings of this study are available from MICCAI grand challenge on 6-month infant brain MRI segmentation [1] and MRBrainS and are publicly available (see Footnote 1).

References

1. Liyan, S., Jiexiang, W., Yue, H., Xinghao, D., Hayit, G., John, P.: An adversarial learning approach to medical image synthesis for lesion detection. IEEE J. Biomed. Health Inform. **24**(8), 2303–2314 (2020)
2. Xin, Y., Ekta, W., Paul, B.: Generative adversarial network in medical imaging: a review. Med. Image Anal. **58**, 101552 (2019)
3. Hadeer, H., Mahmoud, B., Amira, H.: Toward deep MRI segmentation for Alzheimer's disease detection. Neural Comput. Appl. **34**(2), 1047–1063 (2021)
4. Salome, K., et al.: GANs for medical image analysis (2020)
5. Talha, I., Hazrat, A.: Generative adversarial network for medical images (MI-GAN). J. Med. Syst. **42**, 1–11 (2018)
6. Dinggang, S., Guorong, W., Heung-Il, S.: Deep learning in medical image analysis. Annu. Rev. Biomed. Eng. **19**, 221–248 (2017)
7. Muralikrishna, P., Ravi, S.: Medical image analysis based on deep learning approach. Multimedia Tools Appl. **80** 24365–24398 (2021)
8. Min, C., Xiaobo, S., Yin, Z., Di, W., Mohsen, G.: Deep feature learning for medical image analysis with convolutional autoencoder neural network. IEEE Trans. Big Data **7**(4), 750–758 (2021)
9. Mondal, A., Jose, D., Christian, D.: Few-shot 3D Multi-modal Medical Image Segmentation using Generative Adversarial Learning (2018). arXiv:1810.12241v1
10. Yanmei, L., et al.: Edge-preserving MRI image synthesis via adversarial network with iterative multi-scale fusion. Neurocomputing **452**, 63–77 (2021)
11. Yandi, G., Yang, P. Hongjun, L.: AIDS Brain MRIs synthesis via generative adversarial networks based on attention-encoder. In: 2020 IEEE 6th International Conference on Computer and Communications (2020)
12. Emami, H., Dong, M., Nejad-Davarani, S.P., Glide-Hurst, C.K.: SA-GAN: structure-aware GAN for organ-preserving synthetic CT generation. In: de Bruijne, M., et al. (eds.) MICCAI 2021. LNCS, vol. 12906, pp. 471–481. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87231-1_46
13. Rishav, S., Vandana, B., Vishal, P., Abhinav, K., Amit, Kumar K.: MetaMed: few-shot medical image Classification using gradient-based meta-learning. Pattern Recogn. **120**, 108111 (2021)
14. Dolz, J., Ismail, A., Jing, Y., Christian, D.: Isointense infant brain segmentation with a hyper-dense connected convolutional neural network. In: International Symposium on Biomedical Imaging (ISBI) (2018)
15. Afifa, K., chungming, O., Wenyuan, T, Taher, G.: Improved brain segmentation using pixel separation and additional segmentation features. In: The 4th APWeb-WAIM International Joint Conference on Web and Big Data (2020)
16. Vijay, B., Alex, K., Roberto, C.: SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation (2016). ArXiv:1511.00561v3
17. Arshia, R., Saeeda, N., Usman, N., Imran, R., Ibrahim, H.: Deep auto encoder-decoder framework for semantic segmentation of brain tumor. In: ICONIP (2019)

18. Wang, L., et al.: Benchmark on automatic six-month-old InfantBrain segmentation algorithms: TheiSeg-2017 challenge. IEEE Trans. Med. Imaging **38**(9), 2219–2230 (2019)
19. Cirillo, M.D., Abramian, D., Eklund, A.: Vox2Vox: 3D-GAN for brain tumour segmentation. In: Crimi, A., Bakas, S. (eds.) BrainLes 2020. LNCS, vol. 12658, pp. 274–284. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-72084-1_25
20. Niyaz, U., Sambyal, S.: Advances in deep learning techniques for medical image analysis. In: 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC), pp. 271–277 (2018)
21. Yi, S., Chengfeng, Z., Yanwei, F., Xiangyang, X.: Parasitic GAN for semi-supervised brain tumor segmentation. In: IEEE International Conference on Image Processing (ICIP) (2019)
22. Quentin, D., et al.: SegSRGAN: Super-resolution and segmentation using generative adversarial networks - Application to neonatal brain MRI (2020)
23. Yi, D., Fujuan, C., Yang, Z., Zhixing, W., Chao, Z., Dongyuan, W.: A stacked multi-connection simple reducing net for brain tumor segmentation. IEEE Access **7**, 104011–104024 (2019)
24. Afifa, K., Han, H., Taher, G.: Multi-model medical image segmentation using multi-stage generative adversarial networks. IEEE Access **10**, 28590–28599 (2022)
25. Jonas, W., Marcelo, L., Jos, N.: Transfer Learning for Brain Tumor Segmentation. arXiv preprint arXiv:1912.12452
26. Afifa, K., Ahmed, A.M., Kun, H.: Two Independent Teachers are Better Role Mode. arXiv preprint arXiv: 2306.05745
27. Afifa, K., Jian, J.H., Taher, A.G.: Learning to detect boundary information for brain image segmentation. BMC Bioinform. **23**(1), 332 (2022)
28. Afifa, K., Jian, J.H., Taher, A.G., Radman, M.: Fully convolutional neural network for improved brain segmentation. Arab. J. Sci. Eng. **48**(2), 2133–2146 (2023)
29. Xun H., Serge B.: Arbitrary style transfer in real-time with adaptive instance normalization. In: 2017 IEEE International Conference on Computer Vision (ICCV) (2017)
30. Hazrat, A., et al.: Correction: The role of generative adversarial networks in brain MRI: a scoping review. Insights Imaging **13**(1), 98 (2022)