Transfer Learning using Generative Adversarial Networks for MRI Brain Image 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 practice. 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 a Generative Adversarial Network (GAN) model that performs transfer learning to segment 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 our GAN model using a commonly used evaluation metric, which is Dice Coefficient (DC). Our experimental results reveal that our proposed model significantly improves segmentation results compared to the standard GAN model. We observe that our model is 2.1–10.83 minutes faster than stat-of-the-art-models.

Keywords: Transfer Learning; Generative Adversarial Networks; MRI Brain Images

1 Introduction

The great 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 practice [1,2]. 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 was to use two-stage models, by fusing global information

2   Afifa Khaled et al.

with local information generated in two subsequent stages, to achieve acceptable segmentation results. The design of multi-stage models, in general, allows to achieve better results, since it helps solve the information loss problem [5,6,7,8].

There have been many studies [10,11,12,15] proposing techniques to improve the accuracy of brain image segmentation to reach results that are close enough to manual reference. Unlike conventional machine learning models, the use of deep learning algorithms for brain image segmentation received little attention, which is due to the lack of training data. 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 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, it did not achieve promising results compared to previous state-of-the-art models.

While previous techniques enabled neural networks to produce acceptable segmentation output, there were very limited models that address the segmentation of infant brain images to 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 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 a novel Generative Adversarial Network (GAN) model that performs transfer learning to improve the segmentation of MRI brain images, particularly (WM), (GM), and (CSF). Our model enables generating 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. We design the layers of our model to pass image information to the decoder, which in turn transfers tissues information into MRI images using a GT layer where GT layer stands for ground truth.

The remainder of this paper is organized as follows. Section 2 reviews related work. Section 3 presents our GAN model. Section 4 presents our setup materials and methods. Section 5 presents and discusses our experimental results. Finally, Section 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 (GAN) for brain segmentation

GANs have shown promising results in both medical image diagnostics [20] and brain image segmentation [19,23]. Fig. 1 shows GAN has two parts: The generator is to generate the data, whereas 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 $BraTS\ 2018$ dataset. Their model outperformed previous models with an overall accuracy of 66%.

Delannoy et al. [22] proposed a super-resolution and segmentation framework using GAN to neonatal brain MRI images. Their framework composed of two networks (a) a generating network trained to estimate 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%.

Yuan et al. proposed an end-to-end GAN, called $SegAN$ [28] by employing a fully convolutional neural network, as the segmentor, along with an adversarial critic network with a multi-scale $L1$ loss function. Using these two networks enables to learn both global and local features.

Ding et al. [29] utilized a coarse-to-fine to improve brain segmentation using a two-stage GAN, where, in the first stage the model generates a coarse outline for global information, and in the second stage the model generates an outline for specific information. An improved segmentation result was obtained by fusing the coarse and refined outlines.

Kamran et al. [30] proposed a multi-scale GAN for retinal vessel segmentation, called $RV - GAN$, which employed two generators and two multi-scale discriminators and used a weighted feature matching loss from the discriminator's decoder over the encoder. Two metrics (i.e., the mean intersection-over-union ($Mean - IOU$) and the structural similarity measure ($SSIM$)) were used to evaluate the effectiveness of their proposed model.

Wu et al. proposed an unsupervised brain tumor segmentation, called $D - GAN$[31], which learns a non-linear mapping between the left and right brain images and then learns to reconstruct normal brains and to segment brain tumors. Their model was evaluated on the ($BRATS$) dataset (2012 and 2018) and showed an improvement in segmentation results.

Khaled et al. proposed two brain tissues segmentation models, one using FCN+MIL+G+K [15] and another using a multi-stage GAN model [24]. They evaluated their models on two infants and adults brain images and obtained improved segmentation results, expressed by dice coefficients of up to 94% for segmenting GM and WM.
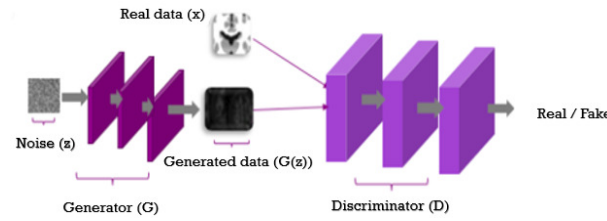
Fig. 1: Illustration of Generative Adversarial Networks (GAN)

2.2 Transfer Learning

Transfer learning in medical imaging can be used in image labeling process, which results in augmented images. Typical, the number of images available in medical applications' dataset is very small. Therefore, to solve such a problem, transfer learning can be utilized to transfer labeled images to unlabeled images. Instead of mixing the unlabeled images with labeled images during model training, transfer learning can help generate more labeled images, which assumes images are smooth and look real during the transfer. Wacker et al. [25] used ResNet with a decoder transfer learning for brain tumor segmentation. Moreover, transfer learning has not widely been used in medical imaging applications. Hence, we expect that transfer learning can bring notable improvement to medical imaging datasets in general, and image segmentation in particular.

3 GAN Transfer Model

This section describes the structure of our proposed GAN model, including the encoder/decoder and a GT layer.
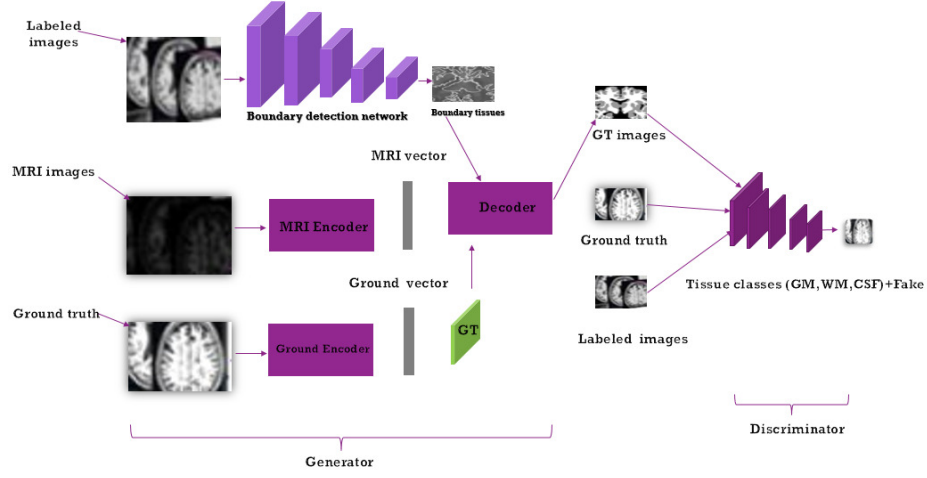
3.1 Encoder/Decoder

Our GAN model consists of encoder and decoder networks, followed by a GT layer. Fig. 2 shows our proposed GAN model. All the MRI encoder, ground truth encoder, tissues mapping, boundary detection network, and decoder together represent the generator of our GAN model. MRI encoder and ground truth encoder take an MRI image and a ground truth to convert them into MRI vector and ground truth vector, respectively. The detection network gives more information about the boundary. Using GT layer, the output of the ground vector is transferred to the decoder. Finally, the output of the decoder is the GT image where GT denotes to the ground truth.

3.2 GT Layer

The GT layer employs an R mapping function to map a tissue y into (μ, σ) parameters.

$$GT(x,y) = R_\sigma(y)x - \mu(x)/\sigma(x) + R_\mu(y) \qquad (1)$$

Fig. 2: Illustration of our proposed GAN model

where R is composed of multiple fully connected layers and takes the latent code of tissues as input.

3.3   Loss Functions

Discriminator loss function The discriminator in our GAN model has labeled data loss, unlabeled data loss, and GT images loss. We formulate the overall loss function of our GAN model is as follows:

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

where λ_{labeled}, $\lambda_{\text{unlabeled}}$, and λ_{fake} are hyper-parameters. We set the hyper-parameters in Equation (2) to $\lambda_{\text{labeled}} = 1.0$, $\lambda_{\text{unlabeled}} = 1.0$, and $\lambda_{\text{fake}} = 2.0$. For labeled data, we use the same loss function in the standard segmentation network. Mondal et al. [9] showed that using $l_{i,k+1}$ as a subtracted function changes the softmax function as follows:

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

$$l_{\text{unlabeled}} = -E_{\text{x} \sim \text{pdata(x)}} \sum_{\text{i}=1}^{H \times W \times D} \log((Z_{\text{i}}(x)/Z_{\text{i}}(x)) + 1), \tag{4}$$

$$l_{\text{fake}} = -E_{\text{z} \sim \text{noise}} \sum_{\text{i}=1}^{H \times W \times D} \log[((1/Z_{\text{i}}(G_{\Theta G}(z) + 1)], \tag{5}$$

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

where labeled loss is the cross-entropy. Our main goal here is to introduce unlabeled loss and fake loss, which are analogues to the two components of discriminator loss in the standard GAN. More information can be found in [9].

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.

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

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 define as $Classes = 1, ..., K$.

4 Setup Materials and Methods

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

4.1 Datasets

In our experiments, we use two datasets to evaluate our model: the $MICCAI\ iSEG$ dataset (containing infant brain images) and the $MRBrainS$ dataset (containing adult brain images). These dataset have been used as a benchmark to evaluated brain segmentation models previous studies [32,33].

MICCAI iSEG Dataset. The $MICCAIiSEG$ organizers[3] introduced a publicly available evaluation framework to allow the comparison of different segmentation models of WM, GM, and CSF on T1-weighted ($T1$) and T2-weighted ($T2$). The MICCAI iSEG dataset contains: 10 infant brain 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
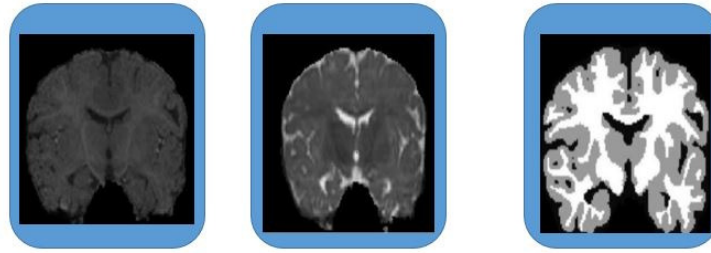
Fig. 3: An example of the MICCAI iSEG dataset ($T1$, $T2$, and manual reference contour)

used as a testing set. An example of the MICCAI iSEG dataset ($T1$, $T2$, and manual reference contour) is shown in Fig.3.

Table 1 shows the parameters used to generate $T1$ and $T2$. The dataset has two different times: the longitudinal relaxation time and 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 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 a 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 proposed model were conducted using $Python$ on a PC with $NVIDIAGPU$ running $Ubuntu$ 16.04. Training our GAN model took 30 hours in total, whereas testing took 5 minutes.

---

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

4.3   Segmentation Evaluation

Dice Coefficient (DC) To better highlight the significance of our GAN model, we use the Dice Coefficient (DC) metric to evaluate the performance of our 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:

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

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 our GAN model on two datasets of different ages: adults and infants. Table 2 presents the results of our GAN model to segment CSF ,GM, and WM using the MICCAI iSEG dataset. Our 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 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 our GAN model compared to the standard GAN.

Table 2: Dice Coefficient (DC) results of the segmentation achieved on the MICCAIiSEG 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 GAN	93%	**94%**	**92%**

Table 3 presents the results achieved using the MRBrains dataset. We observe that our 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.

Fig. 4 shows a sample visualized result of our GAN model on a subject used as part of the validation set. As the images show, we observe that the segmentation

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.

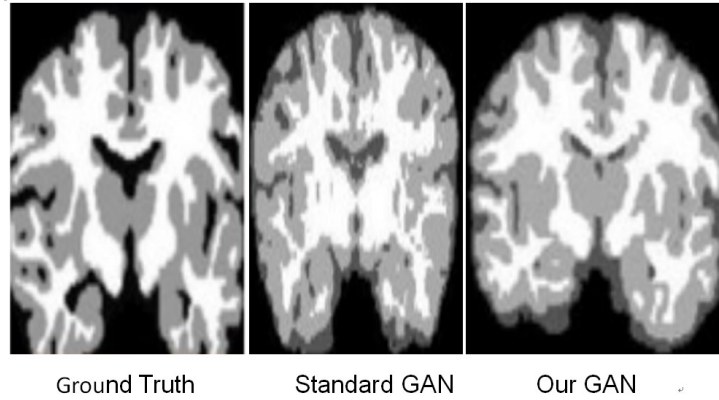| Model	Dice Coefficient (DC) Accuracy		
	CSF	GM	WM
Standard GAN	87%	87%	85%
3D, FCN + MIL+G+K [15]	87.4%	90.6%	90.1%
Multi-stage [24]	**93%**	**93%**	88%

Ground Truth   Standard GAN   Our GAN

Fig. 4: A sample visualized result from the MRBrainS dataset

achieved by our GAN model is fairly close to the manual reference (ground truth) contour provided by the MICCAI iSEG organizers.

Our evaluation results show that the proposed model only outperforms two baselines (Standard GAN and 3D,FCN+MIL+G+K [15]) on the three tissues, but also outperforms multi-stage [24] on two tissues. Adopted the transfer learning approach using GT layer improved the results in GM and WM d on the MICCAIiSEG dataset and WM on the MRBrains dataset compared with Multi-stage [24].

6   Conclusion

In this paper, we proposed a novel Generative Adversarial Network (GAN) model that performs transfer learning for the segmentation of MRI and tissue 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 designed layers to pass the information to the decoder, which in turn transfers tissues information into MRI images using a GT layer. Our experimental results show that our GAN model is a viable solution for tissues brain segmentation as it achieves a significant improvement in the accuracy of brain segmentation

compared to the standard GAN model. We observe that our model is 2.1˘10.83 minutes faster than stat-of-the-art-models.

Future work. We aim in the future to investigate our model performance on segmenting more brain tissues and consider pathological brain images, such as those with tumor or edema. Pathological brain images are not considered in this work due to the lack of the datasets.

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. 2303-2314 (2020)
2. Xin, Y., Ekta, W., Paul, B.: Generative adversarial network in medical imaging: A review. Medical Image Analysis. (2019)
3. Hadeer, H., Mahmoud, B., Amira, H.: Toward deep MRI segmentation for Alzheimer's disease detection. Neural Computing and Applications. (2021)
4. Salome, K., Christoph, B., Arjan, K., Bram, v., Ginneken, N., Shadi, A., Anirban M.: GANs for medical image analysis. (2020)
5. Talha, I., Hazrat, A.: Generative Adversarial Network for Medical Images (MI-GAN). Journal of Medical Systems. (2018)
6. Dinggang, S., Guorong, W., Heung-Il, S.: Deep Learning in Medical Image Analysis. Annu Rev Biomed Eng. (2017)
7. Muralikrishna, P., Ravi, S.: Medical image analysis based on deep learning approach. Multimedia Tools and Applications. 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 Transactions on Big Data. 750-758 (2021)
9. Mondal, A., Jose, D., Christian, D.: Few-shot 3D Multi-modal Medical Image Segmentation using Generative Adversarial Learning. arXiv:1810.12241v1. (2018)
10. Yanmei, L., Dong, N., Bo, Z., Zhian, L., Xi, W., Jiliu, Z., Yan, W., Dinggang, S.: Edge-preserving MRI image synthesis via adversarial network with iterative multi-scale fusion. Neurocomputing. 63-77. (2021)
11. Yandi, G., Yang, P. Hongjun, L.: AIDS Brain MRIs Synthesis via Generative Adversarial Networks Based on Attention-Encoder. 2020 IEEE 6th International Conference on Computer and Communications. (2020)
12. Hajar, E., Ming, D., Siamak, N., Carri, G.: SA-GAN: Structure-Aware GAN for Organ-Preserving Synthetic CT Generation. International Conference on Medical Image Computing and Computer-Assisted Intervention. 471-481 (2021)
13. Rishav, S., Vandana, B., Vishal, P., Abhinav, K., Amit, Kumar K.: MetaMed: Few-shot medical image Classification using gradient-based meta-learning. Pattern Recognition. 471-481 (2021)
14. Dolz, J., Ismail, A., Jing, Y., Christian, D.: ISOINTENSE INFANT BRAIN SEGMENTATION WITH A HYPER-DENSE CONNECTED CONVOLUTIONAL NEURAL NETWORK. International Symposium on Biomedical Imaging (ISBI). (2018)

15. Khaled, A., Chungming, O., Wenyuan, T, Taher, G.: Improved Brain Segmentation Using Pixel Separation and Additional Segmentation Features. 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. ArXiv:1511.00561v3. (2016)
17. Arshia, R., Saeeda, N., Usman, N., Imran, R., Ibrahim, H.: Deep Auto Encoder-Decoder Framework for Semantic Segmentation of Brain Tumor. ICONIP. (2019)
18. L. Wang, D. Nie, G. Li, É. Puybareau, J. Dolz, Q. Zhang, F. Wang, J. Xia, Z. Wu, J. Chen, K. Thung, T. D. Bui, J. Shin, G. Zeng, G. Zheng, V. S. Fonov, A. Doyle, Y. Xu, P. Moeskops, J. P. W. Pluim, C. Desrosiers,I. B. Ayed, G. Sanroma, O. M. Benkarim, A. Casamitjana, V. Vilaplana, W. Lin, G. Li, and D. Shen. Benchmark on Automatic Six-Month-Old InfantBrain Segmentation Algorithms: TheiSeg-2017 Challenge. IEEE transactions on medical imaging. (2019)
19. Marco, C., David, A., Anders, E.: Vox2Vox: 3D-GAN for Brain Tumour Segmentation. International MICCAI Brainlesion Workshop. 274-284 (2020)
20. Niyaz, U., Sambyal, S.: Advances in deep learning techniques for medical image analysis. 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC). 271-277 (2018)
21. Yi, S., Chengfeng, Z., Yanwei, F., Xiangyang, X.:Parasitic GAN for Semi-Supervised Brain Tumor Segmentation. IEEE International Conference on Image Processing (ICIP). (2019)
22. Delannoy, Q., Chi-Hieu, P., Clément, C., Carlos, T., Guillaume, D., Hélène, M., Nathalie, B., Ronan, F., Nicolas, P., François, R.: 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. 104011-104024 (2019)
24. Khaled, A., Han, H., Taher, G.: Multi-Model Medical Image Segmentation Using Multi-Stage Generative Adversarial Networks. IEEE Access. (2022)
25. Jonas W., Marcelo L., Jos N.: Transfer Learning for Brain Tumor Segmentation. arXiv preprint arXiv:1912.12452.
26. Chen T., Song X., Wang C., Preserving-texture generative adversarial networks for fast multi-weighted. IEEE Access. 71048–71059.
27. Iqbal T., Ali H., Generative adversarial network for medical images (MI-GAN). Journal of medical systems. 1–11.
28. Yang P., Yan Y., Zhang H., Yu Y., Shi X., Mou M., Kalra Y., Zhang L., Sun G., Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss.1348-1357. IEEE Transactions on Medical Imaging. (2018)
29. Wu X., Bi M., Fulhama D., Fenga L., Zhoue J., Unsupervised brain tumor segmentation using a symmetric-driven adversarial network. Neurocomputing. (2021)
30. Kamran K., Hossain A., Tavakkoli S., Zuckerbrod K., M. Sanders, S. A. Baker., RV-GAN: Segmenting Retinal Vascular Structure in Fundus Photographs using a Novel Multi-scale Generative Adversarial Network. International Conference on Medical Image Computing and Computer Assisted Intervention. (2021)

12   Afifa Khaled et al.

31. Wu X., Bi L., Fulhama D., Fenga L., Zhoue J., Unsupervised brain tumor segmentation using a symmetric-driven adversarial network. Neurocomputing. (2021)
32. Dolz, J., Ayed, I-B., Yuan, J.: Isointense Infant Brain Segmentation with Hyper Dense Connected Convolutional Neural Network. IEEE 15th International Symposium on Biomedical Imaging (ISBI). 616-620. (2018)
33. Dolz, J., Desrosiers, C., Ayed, I.B.: 3D fully convolutional networks for sub-cortical segmentation in MRI: A large-scale study. NeuroImage. 170, 456–470 (2018)