ENHANCEMENT OF BRAIN MRI IMAGE DATABASE USING CYCLEGAN

Thanh Han Trong, Kien Le Trung, Phuong Nguyen Anh
School of Electrical and Electronic Engineering, Hanoi University of Science and Technology

Abstract: Generating tumor images on brain MRI scans at random locations can assist medical researchers and medical students in predicting the likelihood of tumors. However, brain MRI images with tumors are rare in practice, making the collection of MRI image data with brain tumors a time-consuming process. In this study, we propose the application of CycleGAN to generate T2 pulse sequence MRI images of the human brain from T2 Flair pulse sequence images of the same type and vice versa, thereby increasing the number of MRI images of various types. The returned results will be evaluated and compared with other studies based on FID, SSMI, and PSNR metrics.

Keywords: Brain Tumor, Artificial Intelligence, Convolutional Neural Networks, Machine Learning, Generative Adversarial Networks.

I. INTRODUCTION

Medical imaging, a vital component for clinical analysis and medical interventions, provides a clear insight into various pathologies and hidden injuries concealed by the skin or bones. Magnetic Resonance Imaging (MRI) [1] is one of the most popular techniques used today. It is widely employed in hospitals and medical centers. Within this modality, multiple signal sequences (or modalities) can be obtained, each offering distinct valuable information regarding specific patient conditions.

Diseases related to brain tumors pose a global healthcare concern, and their prevalence continues to rise due to adverse influences from the current social environment. There are numerous types of brain tumors, including both malignant and benign tumors. Brain tumors exhibit rapid growth, leading to severe consequences for human health, and in some cases, even fatality. In such a context, the prompt and early identification of brain tumors in patients plays a pivotal role in timely treatment, particularly before the tumors progress to a deteriorating stage.

The acquisition of a sufficiently large number of brain MRI images is an important issue for machine learning models to improve performance and practical application. Brain MRI images with tumors are rare in practice, and collecting data with this image type is time-consuming. Therefore, generating additional data for training machine learning models, as well as segmentation or classification models for brain tumor detection, is necessary. The application of artificial intelligence and image processing techniques to medical image diagnosis is a widely discussed field, including the classification of pathologies based on brain MRI images. From each brain MRI scan, machine learning models can diagnose and identify various types of brain tumors and propose appropriate treatment methods. A more advanced technique for data augmentation in biomedical images is Generative Adversarial Networks (GAN) [2], which utilizes two convolutional neural networks (CNNs). The most evident application of GAN in medical imaging is data generation for training purposes. This study focuses on utilizing the CycleGAN algorithm [3] to extract features from specific brain regions in T2 MRI images and transfer these features to T2 Flair MRI images, aiming to create a rich dataset widely used in image classification or segmentation algorithms. This is an urgent issue, and the application of deep learning algorithms can assist doctors in quickly searching for brain MRI images with tumors, facilitating timely treatment for patients. The paper is organized as follows: Section 2 provides an overview of brain MRI images and the GAN algorithm models used. Section 3 presents the implementation and evaluation results. The conclusion and future directions are discussed in Section 4.

II. METHODOLOGY

A. Brain MRI images

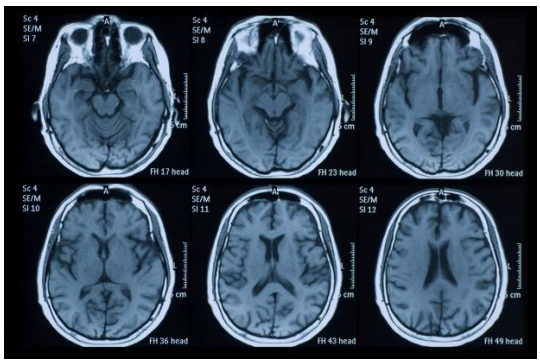


Figure 1. Standard MRI images have been filtered to remove patient names.

The commonly used standard for MRI images today is DICOM, which stands for Digital Imaging and Communications in Medicine [4]. It is an industry standard system developed to meet the needs of manufacturers and users in connecting, storing, exchanging, and printing medical images. The data within MRI images includes demographic information, patient information, research-specific parameters, image dimensions, and more. Patient information displayed includes name, gender, age, and date of birth. Brain MRI images encompass basic types such as T1W, T2W, FLAIR, and DWI. Figure 1 above is an example of a widely used MRI image that has been filtered

Contact author: Thanh Han Trong
Email: Thanh.hantrong@hust.edu.vn
Manuscript received: 6/2023, revised: 7/2023, accepted: 8/2023.

to remove patient-related information. In the T2W phase, the acquired signals have been transformed comprehensively, forming a uniform entity. This image is also useful in assessing hemorrhage and abnormalities. Furthermore, the role of the T2W phase is to reflect the homogeneity of soft tissue masses. This is particularly evident in the diagnosis of meningeal tumors as well as malignant tumors in general. Overall, MRI images are highly effective in diagnosing brain tumors and related pathologies. MRI has proven to be superior in determining the tumor's location and its relationship with surrounding structures.

B. Convolutional Neural Networks

Convolutional Neural Networks (CNNs) [5] are among the most popular and influential deep learning models in the field of computer vision. CNNs are utilized in various tasks such as image recognition, video analysis, medical image processing, and natural language processing. CNN models perform well in addressing a wide range of problems.

A CNN consists of a set of fundamental layers, including convolutional layers, non-linear layers, pooling layers, and fully connected layers, which are interconnected in a specific order. Initially, an image passes through the convolutional and non-linear layers, followed by the pooling layer to reduce the computational complexity while preserving the data features. The convolutional, non-linear, and pooling layers can appear once or multiple times within the CNN. Finally, the data flows through fully connected layers and the softmax activation function is applied to compute the probability of object classification.

C. Generative Adversarial Networks

Generative Adversarial Networks (GANs) were proposed in 2014 by Ian J. Goodfellow [2] and represent a new generation of frameworks for estimating models in adversarial contexts. GANs consist of two networks: the Generator network and the Discriminator network. While the Generator network generates new data based on real data, the Discriminator network is designed to differentiate between the data generated by the Generator and real data.

The architecture of a GAN is depicted in Figure 2. There are two main components in GAN architecture. In the first component, the network needs a device capable of generating new data based on real data. In the case of image generation, for example, or voice synthesis, the model needs the ability to produce audio sequences. Similar principles are applied in other cases. The authors refer to this model as the Generator network. The second component is the Discriminator network. It aims to distinguish between fake and real data. Both networks compete. The Generator network tries to deceive the Discriminator network, while the Discriminator network adapts to the newly generated fake data. The information obtained is used to improve the Generator network and vice versa.

The Discriminator network is a binary classifier that determines whether the input x is real (from real data) or fake (from the Generator network). Typically, the output of the Discriminator network is a predicted value x for the input $i \in R$, for example, by using a fully connected layer with a hidden size of 1, followed by a sigmoid activation function to obtain the predicted probability.

$$D_{(x)} = \frac{1}{1+e^{i}}. \quad (1)$$

Assuming the real data is assigned as label $y = 1$ for real data and $y = 0$ for fake data generated by the Generator network, the Discriminator network will be trained to minimize the cross-entropy loss calculated using the following formula:

$$L_r = min_D\big(-ylogD(x) - (1-y)\,log\big(1 - D(x)\big)\big) \quad (2)$$

The Generator network, during the initialization phase of the data, generates random parameters $z \in R^d$ from a fixed source, in some cases following a standard normal distribution $z \sim N(0,1)$. z is commonly referred to as the latent variable. The goal of the Generator network is to deceive the Discriminator network by making it mistake the generated data $x' = G(z)$ for real data. This is equivalent to having $\big(G(x)\big) \approx 1$. In other words, with the Discriminator network D, the parameters of the Generator network G will be updated to maximize the cross-entropy loss with $y = 0$. Thus, the loss function of G can be computed using the following formula:

$$L_{g1} = max_G(-(1-y)\log\Big(1 - D\big(G(x)\big)\Big) = max_G(-log\Big(1 - D\big(G(x)\big)\Big) \quad (3)$$

In the scenario where the Generator network performs exceedingly well compared to the Discriminator network, resulting in $D(x') \approx 1$, the loss approaches zero, and the gradient of the parameters becomes too small to make any significant progress for the Discriminator network. Therefore, the loss will be minimized as follows:

$$L_{g2} = min_G\Big(-ylog\Big(D\big(G(z)\big)\Big)\Big) = min_G\Big(-\log\Big(D\big(G(z)\big)\Big)\Big) \quad (4)$$

In this case, the Discriminator network only approves elements when $x' = G(z)$ and assigns them the label $y = 1$. It can be said that D and G are playing a "best response" game with the comprehensive objective function defined as follows:

$$Lt = L_{total} min_D max_G(-E_x \sim DatalogD(x) - E_z \sim Noiselog\Big(1 - D\big(G(z)\big)\Big) \quad (5)$$

D. CycleGAN

1) General issues

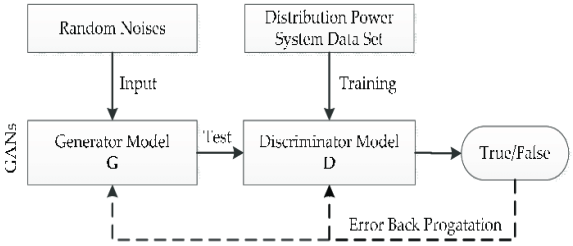


Figure 2. Generative Adversarial Networks GAN

Image-to-image translation [6] is a group of issues in the field of computer vision that aims to learn a mapping between input images and output images. This problem can

be applied to various domains such as format conversion, image colorization, image enhancement, data augmentation for segmentation, face filtering, etc. Typically, to train an image-to-image translation model, a large number of paired input-output image samples are required. However, paired datasets are often scarce or unavailable, necessitating the development of models capable of learning from unpaired data. Specifically, any two unrelated sets of images and their shared features can be used and leveraged for image translation. This is known as the unpaired image-to-image translation problem. One successful approach for unpaired image-to-image translation is CycleGAN [3]. CycleGAN is designed based on Generative Adversarial Networks (GAN) [2]. The GAN architecture is a method for training image generation models consisting of two neural networks: the generator network and the discriminator network. The generator network takes a random vector from the latent space as input and generates new images, while the discriminator network takes an image as input and predicts whether it is real (from the dataset) or fake (generated by the generator network). Both models compete, where the generator is trained to produce images that can deceive the discriminator. Consequently, the discriminator is trained to better distinguish between real and generated images.

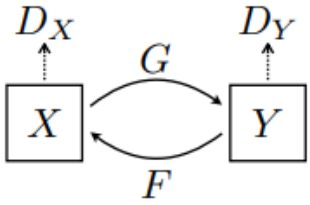


Figure 3. CycleGAN structure

CycleGAN is an extension of the classical GAN architecture, consisting of 2 Generator Networks and 2 Discriminator Networks as illustrated in Figure 3. The first generator network, called G, takes an image from domain X as input and transforms it into domain Y. Another generator network, called F, is responsible for converting images from domain Y to domain X. Each generator network has a corresponding discriminator network:

• DY: Discriminates images taken from domain Y and translated images $G(x)$.

• DX: Discriminates images taken from domain X and translated images $F(y)$.

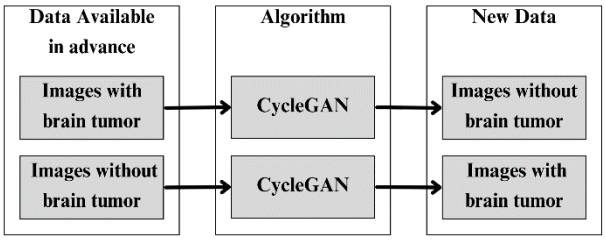


Figure 4. System structure

During the training process, the generator network G tries to minimize its adversarial loss by transforming the image $G(x)$(with x being an image taken from domain X) to resemble the image from domain Y, while the discriminator network DY attempts to maximize its adversarial loss by distinguishing between the individual images $G(x)$and the real images y from domain Y. The adversarial loss function is formulated as follows:

$$Ladv(G, DY, X, Y) = \frac{1}{n}[log\ DY(y)] + \frac{1}{n}\left[log\left(1 - DY\left(G(x)\right)\right)\right] \quad (6)$$

The adversarial loss function is applied similarly to the generator network F and the discriminator network is computed by:

$$Ladv(F, DX, X, Y) = \frac{1}{n}[log\ DX(x)] + \frac{1}{n}\left[log\left(1 - DX\left(F(y)\right)\right)\right] \quad (7)$$

With only the adversarial loss function, the model cannot produce satisfactory results. It would confuse the generator network in generating any output within the target domain, but not the desired output. For example, in the task of transforming a zebra into a horse, the generator could transform a zebra into a beautiful horse but without any characteristics related to the original zebra.

To address this issue, the cycle consistency loss function has been employed. In the paper [3], the authors argue that if an image x from domain X is translated to domain Y and then translated back to domain X using the two generators G and F, respectively, the original image x should be obtained. The cycle consistency loss function in this case is formulated as follows:

$$Lcycle(G, F) = \frac{1}{n}\sum\left|F\left(G(x_i)\right) - x_i\right| + \left|G\left(F(y_i)\right) - y_i\right| \quad (8)$$

From the combination of the two individual loss functions, the overall loss function of the CycleGAN network can be expressed as follows:

$$L = Ladv(G, DY, X, Y) + Ladv(F, DX, X, Y) + \lambda Lcycle(G, F) \quad (9)$$

where λ is a parameter chosen to be 10 in this case to ensure that the weight of the adversarial loss is much smaller compared to the cycle consistency loss. This helps stabilize the consistency of the images after each training cycle.

2) Optimizing ReLU activation function

Conventional GANs commonly use the default ReLU activation function. In some cases, when the learning rate is improperly set or when gradients propagate from the final layers back to the initial layers, they tend to decrease exponentially with the number of layers in deep learning models. This phenomenon makes the gradients in the initial layers very small, leading to difficulties in convergence (vanishing gradient [7]). Initializing the weights according to a certain standard helps the model converge more easily and achieve better results. To address the problem of vanishing gradient in generating 2D medical images and enhance the performance of the ReLU activation function, this paper proposed to apply the Kaiming He initialization method. Kaiming He initialization [8], also known as "He initialization," is an initialization method for deep neural networks that considers the non-linearity of activation functions, such as ReLU. An appropriate initialization method avoids diminishing or amplifying the intensity of the input signal exponentially.

Assuming,

$$y^k = W^k x^k + b^k \qquad (10)$$
$$x^{k+1} = f(y^k)$$

where k is the number of layers and f is the activation function. Here, y, x, and b are column vectors, and W is a matrix. This also applies to feedforward neural networks (NNs) as well as convolutional neural networks (CNNs) (since convolutional operations can be expressed as matrix multiplication). Each component yi of y is computed as follows:

$$y_i = W_{i,1}x_1 + W_{i,2}x_2 + W_{i,3}x_3 + \cdots + W_{i,n}x_n + W_{i,1}x_1 + b_i \qquad (11)$$

where n is the size of the input matrices of the activation function at each layer of the model. This leads to:

$$Var(y_i) = n * Var(W_{i,j}) * E[x_j^2] \qquad (12)$$

where

$$E[x_j^2] \qquad Var(x_j)$$

, as the ReLU activation function does not have a mean value of 0.

The condition for using this initialization type is:

$$\frac{n^l}{2} Var(W^l) = 1 \qquad (13)$$

leading to W following a Gaussian distribution centered at the origin with a standard deviation of $\sqrt{\frac{2}{n^l}}$

$$W \sim N\left(0, \frac{2}{n^l}\right) \qquad (14)$$

E. Fréchet Inception Distance

The Inception Score (IS), proposed by Salimans et al. [9], is one of the popular methods to evaluate the image quality and diversity of GANs by utilizing a pretrained network (InceptionNet [10], trained on the ImageNet dataset [11]) to extract desired attributes from the generated images. In this study, the generated images are MRI brain images, which do not belong to any of the classes in the ImageNet dataset. Therefore, to assess the image quality and performance of the generated MRI brain images by CycleGAN, the Fréchet Inception Distance (FID) [12] has been employed. FID is one of the most used metrics to evaluate GANs nowadays, and a lower FID value is considered better. FID embeds a set of images into a feature space. Viewed as a continuous multivariate Gaussian distribution, this feature space is used to compute the mean and covariance of the generated and real images. The Fréchet distance between these two distributions is then used to evaluate the quality of the generated samples, with a lower FID indicating a smaller distance between the real and generated distributions. The FID score is calculated using the following formula:

$$FID(r,g) = \|\mu_r - \mu_g\|_2^2 + \mathrm{Tr}\,(\sum\_r + \sum\_g - 2\sqrt{(\sum\_r \sum\_g)}) \qquad (15)$$

where $(\mu r, Pr)$ and $(\mu g, Pg)$ are the mean and covariance of the real images and the generated images, respectively, beside that $\sum_r$ and $\sum_g$ are covariance matrix of vectors [18]. The matrix $Tr(\ )$ is a trace matrix of size $n * n$, defined as follows:

$$Tr(A) = \sum_{i=1}^{n} a_{ii} \qquad (16)$$

F. Structural Similarity Index (SSIM)

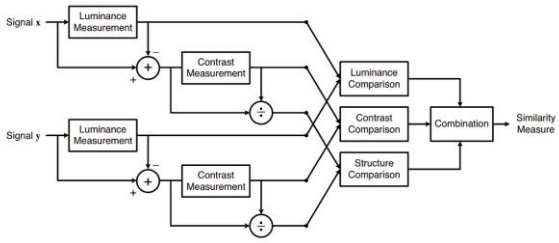


Figure 5. The structure of the SSIM evaluation

The Structural Similarity Index Measure (SSIM), proposed by Zhou Wang et al. [13], is a metric used to quantify the similarity between two given images. The SSIM metric extracts three main features from an image: luminance, contrast, and structure. The comparison between two images is based on these three features. Figure 5 below illustrates the arrangement and flow of the structural similarity measurement system. The signals *X* and *Y* refer to the Reference Image and the Sample Image, respectively. This process computes the structural similarity index between the two given images, with values ranging from -1 to $+1$. A value of $+1$ indicates that the two given images are very similar or identical, while a value of -1 indicates a significant difference between the two images. Typically, these values are adjusted to fall within the range $[0, 1]$, where the maximum values still carry the same meaning as the original scale.

G. The Peak Signal-to-noise Ratio (PSNR)

The Peak Signal-to-Noise Ratio (PSNR) is the ratio between the maximum power of a signal and the power of the noise in that signal. Engineers often use PSNR to measure the quality of reconstructed images that have been compressed. Each pixel in an image can have a varying color value when the image is compressed and then decompressed. Signals can have a wide dynamic range, so PSNR is commonly expressed in decibels.

PSNR is calculated based on the Mean Squared Error (MSE) between two color images, where one image is considered an approximation of the other. MSE can be described as the average squared difference in pixel values between corresponding pixels of the two images.

The mathematical expression of Peak Signal-to-Noise Ratio (PSNR) is presented as follows:

$$PSNR = 20\,log\left(\frac{MAX}{\sqrt{MSE}}\right) \qquad (17)$$

with MSE calculated according to the formula:

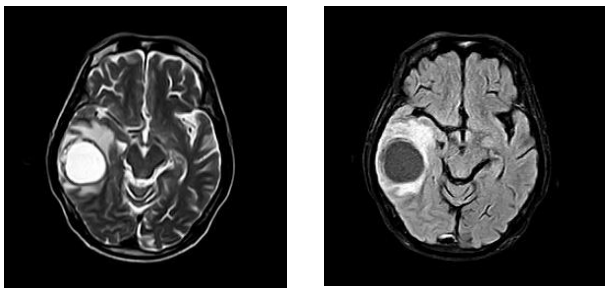
$$MSE = \frac{1}{mn}\sum_{o}^{m-1}\sum_{o}^{n-1}\|f(i,j) - g(i,j)\|^2 \qquad (18)$$

where f represents the matrix data of the original image; g represents the matrix data of the generated image; m is the number of rows of the image pixels and i represents the index of the row; n is the number of columns of the image

pixels and j represents the index of the column; $MAXf$ is the maximum signal value existing in the original image.

III. EXPERIMENT

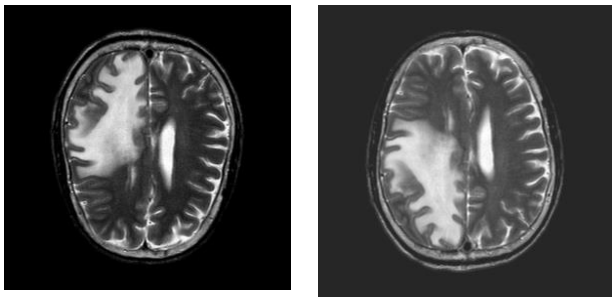
A. Data preprocessing – data cleaning



A B

Figure 6. T2 pulse sequence (A) and T2 Flair pulse sequence (B) images depicting a brain tumor in the patient.

In this study, the dataset used consisted of imaging data from 123 patients with brain cancer at Bach Mai Hospital, encompassing all age groups. Initially, the MRI images were stored in DICOM format. To remove patient-specific information from the DICOM images and convert them into a suitable format for machine learning, the DICOM format was transformed into JPEG format with a size of 256x256 pixels.



A B

Figure 7. Pre-processing images (A) and post-processing images (B).

The images used during the training process were T2 pulse sequence images and T2 Flair pulse sequence images. The signal intensity with the T2 phase of these two types of images correlates well not only with tissue homogeneity but also with cellular profiles.

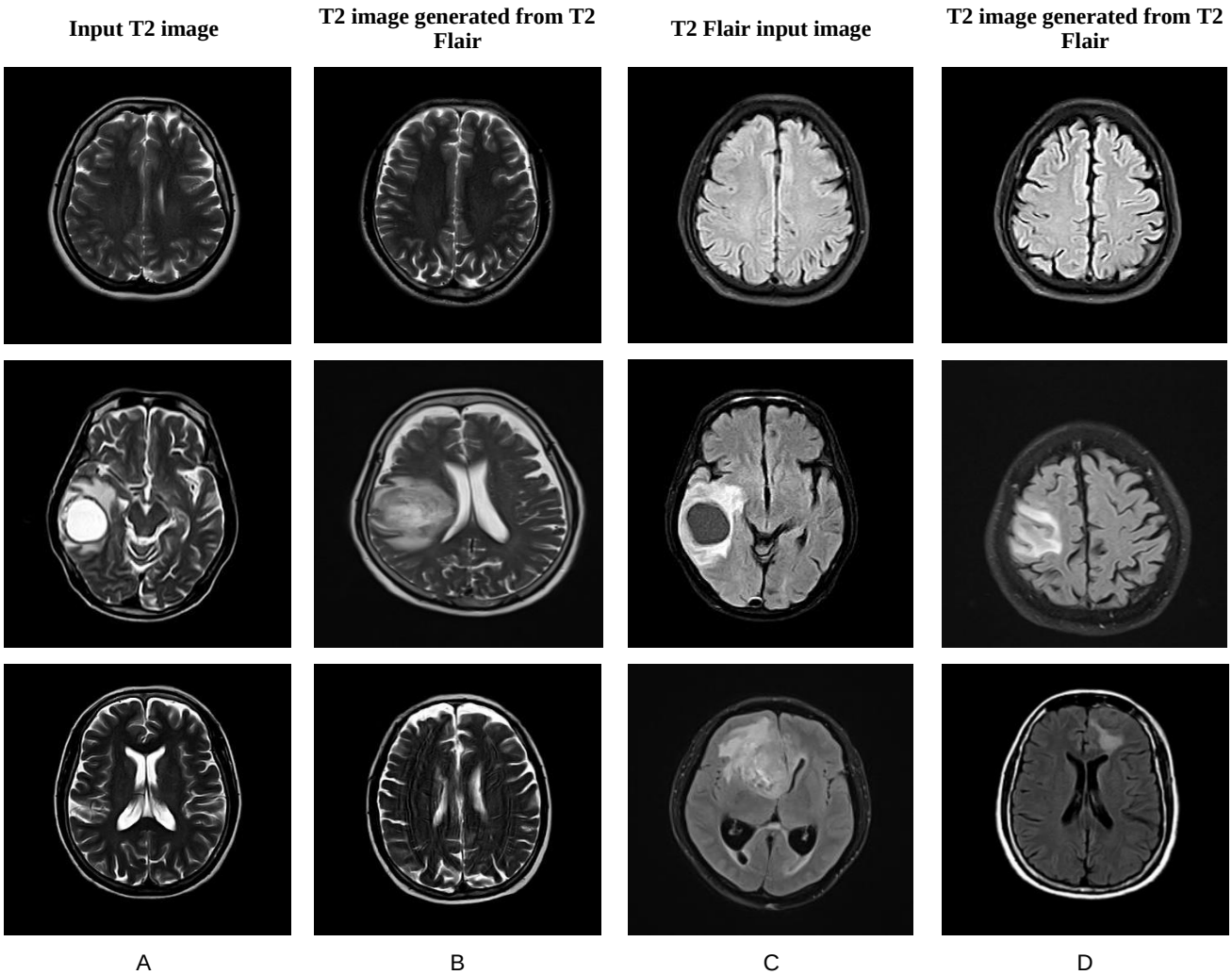


Figure 8. MRI images of T2 and T2 Flair sequences (A, C); Generated images of T2 and T2 Flair sequences in different cases (B, D).

Specifically, areas with low signal intensity represent more fibroblastic tissue, while regions with higher signal intensity indicate a softer characteristic, such as a vascular tumor. This makes the T2 pulse sequence and T2 Flair pulse sequence images the most reliable methods for determining the presence of brain tumors in patients. Out of the 123 patients with brain tumors, a total of 1307 T2 pulse sequence and T2 Flair pulse sequence images were selected, with 647 images displaying the T2 pulse sequence and 660 images displaying the T2 Flair pulse sequence.

The images in the database were preprocessed with several steps: brightness adjustment, contrast enhancement, and random image flipping, which improved the ability to detect image features. For example, images will be randomly flipped at angles of 30, 60 90 degrees and increased and decreased by 50% brightness to serve for data augmentation.

B. Result

The quantitative evaluation:

Figure 8 shows the results achieved by the CycleGAN algorithm. It can be observed that Figure 8B is generated from the original Figure 8C and produced by the CycleGAN model with the characteristics of T2 Flair image sequence. Similarly, Figure 8D is generated from the original Figure 8A. The images in each row represent different cases: generating non-tumor images from non-tumor image sequences, generating tumor images from tumor image sequences, and generating tumor images from non-tumor image sequences. Therefore, during the model generation process, each original image can produce more than one new MRI image. With the initial dataset consisting of 660 images of each type of T2 and T2 Flair, the model has generated a new dataset of over 800 images for each type. In MRI images, the T2 brain is characterized by the presence of cerebrospinal fluid with the highest signal intensity, resulting in a bright white appearance, pale gray matter, light gray color for white matter, and bright color for tumor cell masses. Usually, abnormal brain tissue appears white. Qualitative assessment using visual inspection of the images indicates that the generated tumor brain images (Figure 8B) from non-tumor images (Figure 8A) exhibit similar characteristics to the T2 image sequence. Figure 8B clearly shows the distinct-white-colored abnormal cells in the T2 tomography slice.

During the model training process, the loss function is an important factor to consider whether the model performs well or not. The smaller the loss function, the more accurate the similarity between the generated images and the original images. Figures 9 and 10 show the loss functions of the CycleGAN model when using the T2 MRI image sequence to generate the T2 Flair image sequence and vice versa. It can be observed that the loss function of the discriminator tends to decrease over epochs (here, 100 epochs are selected as the loss function reached a saturation point and could not decrease further), indicating that the discriminator of the GAN model becomes less capable of detecting the difference between the generated images and the original images, meaning that the generated images closely resemble the original images in terms of their characteristic.

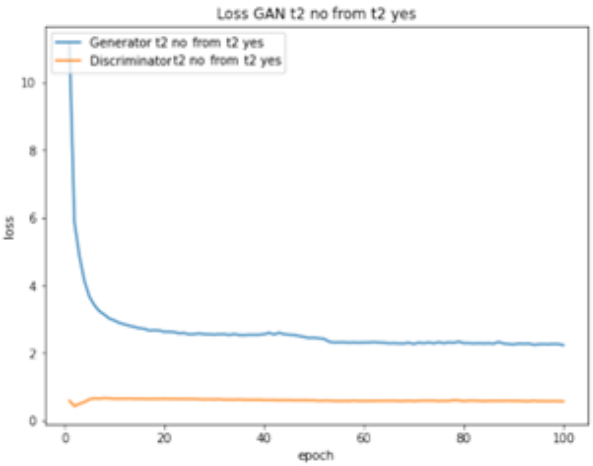


Figure 9. Loss function graph of CycleGAN from T2 to T2 Flair image generation case

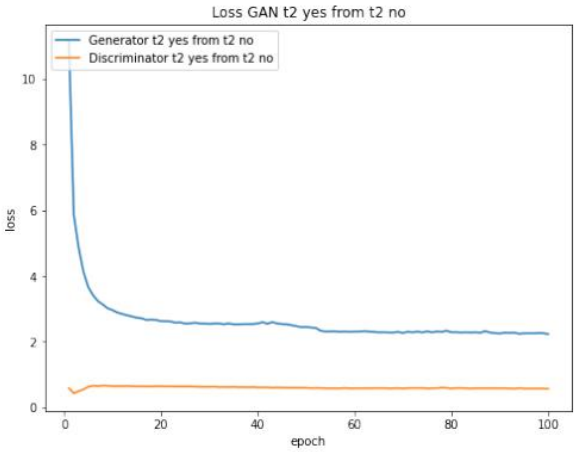


Figure 10. Loss function graph of CycleGAN from T2 Flair to T2 image generation case.

Table 1 displays the FID scores of the two generated image sets: the "Generated T2" set, which consists of MRI T2 images generated from the set of T2 flair images, compared to the original MRI images. A lower FID score indicates a smaller difference between the two datasets. With a dataset containing 660 images for each image sequence, the FID scores are evaluated to be reasonably low, indicating that the generated images can be used for other deep learning algorithms.

TABLE I. Comparison of FID scores with previous studies in using GAN for generating 2D MRI brain images

Research	Algorithm	FID
Kossen, Tabea, et al., 2021 [14]	DCGAN	141.82
Li, Qingyun, et al., 2020 [15]	TumorGAN	77.43
This study	CycleGAN	54.86

TABLE II. Comparison of SSIM and PSNR scores with some methods for 2D medical image synthesis

Method	SSIM	PSNR
PAN [16]	0.9027	25.82
IDC-GAN [6]	0.9039	23.02
MED-GAN [6]	0.9039	23.32
CycleGAN	0.7585	19.05

The compared results of FID, SSIM, and PSNR scores for our proposed system in this study with the most recent published studies are presented in Table I and Table II. The primary application of FID metric is to assess the quality and diversity of images generated by generative models, such as GANs. When FID scores are lower, it signifies higher image quality and greater diversity. Specifically, a lower FID score implies that the generated images closely match the distribution of real images. Based on these tables, it is evident that our proposed system achieves an FID score of 54.86, which is relatively good, compared to other research on brain MRI. It outperforms the DCGAN algorithm with an FID score of 55.61. Furthermore, it surpasses the FID score of 141.82 proposed by Kossen, Tabea et al., 2021 [14], and outperforms the TumorGAN algorithm with an FID score of 77.43 proposed by Li, Qingyun et al., 2020 [15].

While using CycleGAN for image generation is not a novel approach, it does offer greater flexibility for customization compared to the method described in the reference paper. Due to its simplicity, CycleGAN can be adapted to produce improved results. However, it's crucial to acknowledge that due to its straightforward model, CycleGAN may not achieve the same level of precision as other specialized liver-specific methods like MEDGAN and IDCGAN, as discussed in this article. Nonetheless, the brain MRI images generated by the CycleGAN model can still be effectively employed in further scientific research.

Experiments involving different deep learning models on data generated by CycleGAN are in the planning stages. The objective is to demonstrate and enhance the effectiveness of this method.

IV. CONCLUSION

The paper focuses on the application of image processing technologies such as the CycleGAN network to generate new images based on the characteristics of existing image datasets, thereby enriching the dataset for classification and segmentation tasks. After using the CycleGAN model, the generated MRI images of T2 and T2 Flair sequences achieved an FID score of 54.86. Based on the current results, the aim is further refinement and development for the model to be applied to other sequences such as T1, FLAIR, and DWI to increase the number of MRI images available for research purposes.

DATA AVAILABILITY

The data used to support the findings of this study may be requested from the respective authors.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

This research was funded by the Ministry of Education and Training (MOET) Vietnam (Project No. B2021-BKA-09). The authors also acknowledge the support from Hanoi University of Science and Technology (Vietnam) and the Radiology Center of Bach Mai Hospital (Vietnam) in the implementation of this study.

REFERENCES

[1] A. Lashkari, "A Neural Network based Method for Brain Abnormality Detection in MR Images Using Gabor Wavelets," *International Journal of Computer Applications*, vol. 4, Jul. 2010, doi: 10.5120/841-1140.

[2] I. Goodfellow *et al.*, "Generative Adversarial Networks," *Advances in Neural Information Processing Systems*, vol. 3, Jun. 2014, doi: 10.1145/3422622.

[3] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," in *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2242–2251. doi: 10.1109/ICCV.2017.244.

[4] P. Mildenberger, M. Eichelberg, and E. Martin, "Introduction to the DICOM standard.," *Eur Radiol*, vol. 12, no. 4, pp. 920–927, Apr. 2002, doi: 10.1007/s003300101100.

[5] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," *ArXiv e-prints*, Nov. 2015.

[6] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 5967–5976. doi: 10.1109/CVPR.2017.632.

[7] L. R. Grosenbaugh, "Generalization and reparameterization of some sigmoid and other nonlinear functions," *Biometrics*, vol. 21, no. 3, pp. 708–714, 1965, doi: 10.2307/2528551.

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[9] T. Salimans, I. J. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved Techniques for Training GANs," *ArXiv*, vol. abs/1606.03498, 2016, [Online]. Available: https://api.semanticscholar.org/CorpusID:1687220

[10] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, 2015.

[11] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, 2009.

[12] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium," Dec. 2017.

[13] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004, doi: 10.1109/TIP.2003.819861.

[14] T. Kossen *et al.*, "Synthesizing anonymized and labeled TOF-MRA patches for brain vessel segmentation using generative adversarial networks," *Computers in biology and medicine*, vol. 131, p. 104254, 2021.

[15] Q. Li, Z. Yu, Y. Wang, and H. Zheng, "TumorGAN: A Multi-Modal Data Augmentation Framework for Brain Tumor Segmentation.," *Sensors (Basel)*, vol. 20, no. 15, Jul. 2020, doi: 10.3390/s20154203.

[16] J. M. Wolterink, A. M. Dinkla, M. H. F. Savenije, P. R. Seevinck, C. A. T. van den Berg, and I. I\vsgum, "Deep MR to CT Synthesis Using Unpaired Data," *ArXiv*, vol. abs/1708.01155, 2017, [Online]. Available: https://api.semanticscholar.org/CorpusID:35639092

PHƯƠNG PHÁP TẠO ẢNH MRI NÃO BỘ DỰA TRÊN CYCLEGAN CẢI TIẾN

Tóm tắt: Việc tạo hình ảnh khối u trên hình ảnh MRI não tại vị trí ngẫu nhiên có thể giúp các nhà nghiên cứu y tế và sinh viên y học dự đoán khả năng có khối u. Tuy nhiên, hình ảnh MRI với khối u não là hiếm gặp trong thực tế, do đó việc thu thập dữ liệu hình ảnh MRI với khối u não mất rất nhiều thời gian. Trong nghiên cứu này, chúng tôi đề xuất áp dụng CycleGAN để tạo ra hình ảnh MRI chuỗi xung T2 của não người từ hình ảnh chuỗi xung xoá dịch T2 Flair cùng loại và ngược lại, từ đó tăng số lượng hình ảnh MRI các loại. Kết quả trả về sẽ được đánh giá và so sánh với các nghiên cứu khác dựa trên các thông số FID, SSMI và PSNR.

Từ khóa: Khối u não, Trí tuệ nhân tạo, Mạng nơ-ron tích chập, Học máy, Mạng đối nghịch tạo sinh...

Thanh Han Trong received the B.E., M.E., and Dr. Eng. degrees in Electronics and Telecommunications from Hanoi University of Science and Technology, Vietnam in 2008, 2010 and 2015, respectively. In 2019, He was a visiting researcher in The University of Electro - Communication, Japan. He is currently an Assistant Professor at School of Electrical and Electronic Engineering, HUST. His research interests are Software Defined Radio, Advance Localization System and Signal processing for Medical R.

Kien Le Trung is currently a Student at School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Vietnam. His research interests include deep learning, digital image processing, computer vision and signal processing for wireless communications.

Phuong Nguyen Anh is currently a Student at School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Vietnam. Her research interests include deep learning, digital image processing, computer vision and signal processing for wireless communications.