

# Brain MRI Images Generating Method Based on CycleGAN

Hinh Van Nguyen, Thanh Han Trong\*

**Abstract**—Generating tumor images on a brain MRI image at random locations can help medical researchers and medical students in predicting possible tumor possibilities. However, MRI imaging with brain tumors is uncommon in practice, therefore the collecting of MRI images with brain tumors databased takes a lot of time. In this study, we propose to apply CycleGAN to create MRI images with brain tumors from MRI images without brain tumors, thereby increasing the number of MRI images with brain tumors. The received results will be evaluated and compared to others studied based on FID score.

**Index Terms**—Brain Tumor, Artificial Intelligence, Convolution Neural Networks, Machine Learning, Generative Adversarial Network.



## 1. Introduction

**M**EDICAL imaging is important for clinical analysis and medical interventions because it provides important insights into a number of diseases whose structures may be hidden by the skin or by bone. One of the most common techniques used today is Magnetic Resonance Imaging (MRI) [1]. This is a common technique in hospitals and medical centers. In this type of imaging, many different sequences (or modalities) can be obtained, and each sequence can provide useful and different insights into a particular patient problem.

Brain tumor is a global public health problem and is increasingly appearing due to the adverse effects of the current social environment. There are many types of brain tumors, among which there are malignant tumors and benign brain tumors. Brain tumors arise very quickly, causing serious consequences and even death. In such a context, the early and prompt identification of the patient's brain tumor leads to timely treatment when the brain tumor is not at a bad stage.

A large enough number of brain MRI images is an important issue for machine learning models to improve performance and apply in practice. Brain MRI images with brain tumors are rare in practice and data collection with this type of imaging takes a long time. Therefore, generating additional data for the use of machine learning models such as segmentation or classification models for brain tumor detection is

essential.

The application of artificial intelligence and image processing technologies to the diagnosis of diseases from medical images is an area that is mentioned a lot today, including the classification of diseases based on brain MRI images. From brain MRI scans, it is possible to diagnose and recognize many different types of brain tumors and offer appropriate treatment methods. A more advanced data enhancement technique, generative adversarial networks (GAN) [2], uses two convolution neural networks (CNN). The most obvious application of GANs in medical imaging is to generate training data. This study focuses on using the CycleGAN algorithm [3] to extract the brain tumor area features of MRI images with brain tumors, thereby assigning brain tumor features to MRI images without brain tumors, this is for the purpose of creating a richer volume of data, which is widely used in image classification or segmentation algorithms. This is an urgent problem today, the application of deep learning algorithms can help doctors find MRI images with brain tumors quickly, helping patients receive timely treatment.

The article is organized as follows. Part 2 provides an overview of MRI brain images and the GAN algorithm models used. Section 3 presents the results of the implementation and makes an assessment. Conclusion and development direction in section 4.

## 2. Materials and methods

### 2.1. Brain tumor MRI image

The commonly used standard for MRI images today is DICOM - an acronym for Digital Imaging and Communications in Medicine Standards [4] is an industry standard system developed to meet the needs of manufacturers as well as users in connecting, storing, exchanging and printing medical images. Data in MRI images include demographic information, patient information, parameters obtained for image study,

Hinh Van Nguyen is with School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Hanoi 100000, Vietnam .

Thanh Han Trong is with School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Hanoi 100000, Vietnam (E-mail: thanh.hantrong@hust.edu.vn).

\*Corresponding author: Thanh Han Trong (E-mail: thanh.hantrong@hust.edu.vn)

Manuscript received July 13, 2022; revised October 01, 2022; accepted November 05, 2022.

Digital Object Identifier 10.31130/ud-jst.2022.310ICT

image size, ..., patient information displayed includes: full name, gender, age, date of birth.

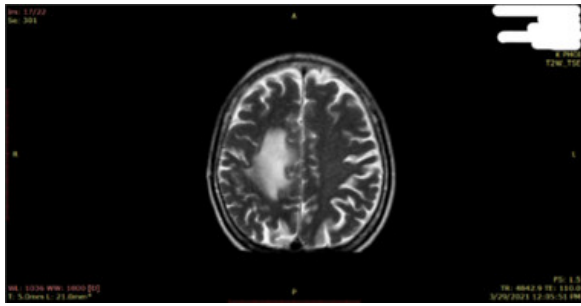


Fig. 1: The MRI image has been masked with the patient's name.

Brain MRI images are of basic types: T1W phase, T2W phase, FLAIR, and DWI. In the T2W phase image, the gain signal has changed completely, which is a fairly homogeneous gain block. Imaging is also helpful in evaluating hemorrhages and cysts. Furthermore, the role of T2W stage is to reflect the homogeneity of soft tumors. This is seen more clearly in meningiomas, malignancies in general. Overall, MRI imaging is very effective in diagnosing brain tumors and brain-related diseases. MRI has been shown to be superior in localizing the tumor and its relationship to surrounding structures.

**2.2. Convolutional Neural Networks**

Convolutional Neural Networks (CNN) [5] is one of the most popular and most influential deep learning models in the computer vision community. CNN is used in many problems such as image recognition, video analysis, MRI images or for problems in the field of natural language processing and most of them solve these problems well.

CNN includes a set of basic layers such as: Convolution layer, nonlinear layer, pooling layer, fully connected layer, layers linked together in a certain order. Basically an image will go through the convolution layer and nonlinear layer first, then the computed values will go through the pooling layer to reduce the number of operations while keeping the data features. The convolution layer, nonlinear layer and pooling layer can appear one or more times in the CNN network. Finally, the data is passed through fully connected and softmax to calculate the probability of object classification.

**2.3. Generative Adversarial Networks**

The generative adversarial networks was proposed in 2014 by Ian J. GoodFellow [2] and represents a new framework for estimating patterns arising in adversarial contexts. GAN is composed of two networks, Generator and Discriminator. While the Generator generates realistic data, the Discriminator tries to distinguish between the data generated by the Generator and the actual data.

The architecture of the congruent network is depicted in Fig. 2. As we can see, there are two

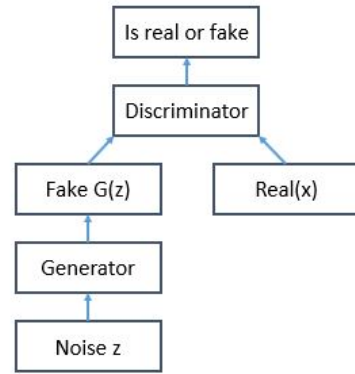


Fig. 2: Generative Adversarial Networks.

components to the architecture of the GAN - first, we need a device that is capable of generating lifelike data. If we are working with images, the model needs to generate the image. If we are working with speech, the model needs to be able to generate audio sequences, etc. We call this model a generator network. The second component is the discriminator network. It tries to distinguish fake and real data. Both of these networks will compete with each other. The life net will try to deceive the discriminator. At the same time, the discriminator network will adapt to the newly generated dummy data. The information obtained will be used to improve lives, and so on.

The discriminator network is a binary classifier that distinguishes whether the input  $x$  is real (from the real data) or fake (from the generator network). Usually, the output of the discriminator network is a predicted  $x$  scalar for the input  $o \in \mathbb{R}$ , such as using a fully connected layer with hidden size 1 and then passed through the sigmoid function to get the probability predict  $D(x) = \frac{1}{(1+e^o)}$ . Suppose using label  $y$  for real data is 1, for real data is 0, we will train the discriminator network to minimize the cross-entropy loss, that is:

$$\min_D \left( -y \log D(x) - (1 - y) \log (1 - D(x)) \right) \quad (1)$$

For the generator network, it first generates a few random parameters  $z \in R^d$  from a source, for example, a normal distribution  $z \sim N(0, 1)$ . We often call  $z$  it a latent variable. The goal of the generator network is to trick the discriminator network into classifying it  $x' = G(z)$  as real data, that is, we want  $D(G(z)) \approx 1$ . In other words, given a discriminator network  $D$ , we will update the parameter of the generator network  $G$  to maximize the cross-entropy loss when  $y = 0$ , that is:

$$\begin{aligned} \max_G \left( - (1 - y) \log \left( 1 - D(G(z)) \right) \right) = \\ \max_G \left( - \log \left( 1 - D(G(z)) \right) \right) \end{aligned} \quad (2)$$

If the generator network does well, then  $D(x') \approx 1$  for the loss to be close to 0, the resulting gradients will become too small to make any significant progress for the discriminator network. Therefore, we will minimize the loss as follows:

$$\min_G \left( -y \log \left( D(G(z)) \right) \right) = \min_G \left( -\log \left( D(G(z)) \right) \right) \quad (3)$$

where only  $x' = G(z)$  the discriminator network is introduced but given the label  $y = 1$ . It can be said that D and G are performing a "minimax" game with a comprehensive objective function as follows:

$$\min_D \max_G \left( -E_{x \sim Data} \log D(x) - E_{z \sim Noise} \log \left( 1 - D(G(z)) \right) \right) \quad (4)$$

## 2.4. CycleGan

Image-to-image translation [6] is a class of computer vision problems whose goal is to learn a mapping between input and output images. This problem can be applied to a number of areas such as style transfer, image coloring, image sharpening, data generation for segmentation, face filter, ... Typically, to train an Image-to-image translation model, we will need a large number of input and label image pairs. Since pairwise datasets are almost non-existent, there is a need to develop a model capable of learning from unpaired data. More specifically, any two sets of unrelated images and common features extracted from each collection can be used and used in image translation. This is called the unpaired image-to-image translation problem. A 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 an approach to training an image generation model consisting of two neural networks: a generator network and a discriminator network. Generator takes a random vector taken from latent space as input and generates new image and Discriminator takes an image as input and predicts whether it is real (taken from dataset) or fake (generated by generator). Both models will compete against each other, the Generator will be trained to generate images that can fool the Discriminator and the Discriminator will be trained to better distinguish the generated images.

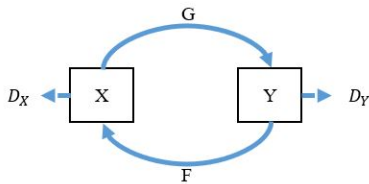


Fig. 3: Generative Adversarial Networks.

CycleGan is an extension of the classic GAN architecture consisting of 2 Generators and 2

Discriminators as shown in Fig. 3. The first generator, called G, takes as input an image from domain X and converts it to domain Y. The other generator called Y, is responsible for converting images from domain Y to X. Each Generator network has a Discriminator corresponding to it:

- $D_Y$ : Distinguish images taken from domain Y and translated images  $G(x)$ .
- $D_X$ : Distinguish images taken from domain X and translated images  $F(y)$ .

During training, the generator G tries to minimize the adversarial loss function by translating the image  $G(x)$  (with  $x$  the image taken from the domain X) so that it is most similar to the image from the domain Y, otherwise the Discriminator  $D_Y$  tries to maximize the adversarial loss function by analyzing separate image  $G(x)$  and real image  $y$  from domain:

$$L_{adv}(G, D_Y, X, Y) = \frac{1}{n} \left[ \log D_Y(y) \right] + \frac{1}{n} \left[ \log \left( 1 - D_Y(G(x)) \right) \right] \quad (5)$$

Adversarial loss is similarly applied to generator F and Discriminator:

$$L_{adv}(F, D_X, X, Y) = \frac{1}{n} \left[ \log D_X(x) \right] + \frac{1}{n} \left[ \log \left( 1 - D_X(F(y)) \right) \right] \quad (6)$$

With adversarial loss alone, it is not enough for the model to give good results. It will hybridize the generator in the direction of producing any output image in the target domain but not the desired output. For example, with the problem of turning a zebra into a normal horse, the generator can turn a zebra into a very beautiful ordinary horse but has no features related to the original zebra.

To solve this problem, cycle consistency loss is introduced. In paper [3], the author thinks that if image  $x$  from domain X is translated to domain Y and then translated back to domain X by 2 generators G, F respectively, we will get the original  $x$  image:

$$x \rightarrow G(x) \rightarrow F(G(x)) \approx x \quad (7)$$

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

From 2 losses on full loss of CycleGan is represented by the formula:

$$L = L_{adv}(G, D_Y, X, Y) + L_{adv}(F, D_X, X, Y) + \lambda L_{cycle}(G, F) \quad (9)$$

where  $\lambda$  is the hyperparameter and is chosen as 10.

## 2.5. Fréchet Inception Distance (FID)

The Inception Score (IS) was proposed by Salimans et al. [7] is one of the popular methods to evaluate the image quality and image diversity of GANs using a pre-trained network (InceptionNet [8], trained on the ImageNet dataset [9]) to capture the properties of the desired GAN in the generated image. In this study, the generated images are brain MRI images, which do not belong to one of the classes of the ImageNet dataset. Therefore, to evaluate the image quality and efficiency of the brain MRI images generated by CycleGAN, Fréchet Inception Distance [10] (FID) was used. FID is one of the most common metrics used to evaluate GANs today and a lower value of FID is considered better. FID embeds a set of images into a feature space. When viewed as a continuous multivariable Gaussian distribution, this feature space is used to calculate the mean and variance of the generated and real images. The Distance Fréchet between these two distributions is used to evaluate the quality of the generated samples, where a lower FID means that the distance between the real and generated distributions is smaller. FID is calculated using the following formula:

$$FID(r, g) = \|\mu_r - \mu_g\|_2^2 + Tr \left( \sum_r + \sum_g - 2 \sqrt{\sum_r \sum_g} \right) \quad (10)$$

where  $(\mu_r, \sum_r)$  and  $(\mu_g, \sum_g)$  are the mean and covariance of the real image and the generated image, respectively. matrix  $Tr()$  is a trace matrix of size  $n * n$  defined:

$$Tr(A) = \sum_{i=1}^n a_{ii} \quad (11)$$

## 3. Experiments

### 3.1. Data collection

In this study, this dataset is a dataset of MRI brain tumors of 123 patients with brain tumors at Bach Mai Hospital, of all ages. Initially, the MRI image was in DICOM format, to remove the information in the patient's DICOM image and convert the image format for machine learning, the DICOM format has been converted to a JPEG image format with a size of 256x256 pixels.

The image used during training is a T2 pulse sequence image. Signal intensity with T2 phase correlates very well with not only homogeneity but also tissue profile. In particular, with low-intensity signal, the tumor is fibrous and stiffer than normal parenchyma; for example, the tumor has a fibroblastic nature, while the more intense sections show a softer characteristic such as a vascular tumor. This makes the image of the T2 pulse sequence the best assessment of whether the patient has a brain tumor or not. With the above 123 patients with brain tumor pathology, 1307 images of T2 pulse sequence were filtered out, of which 647 images showed brain tumors and 660 images did not show brain tumors.

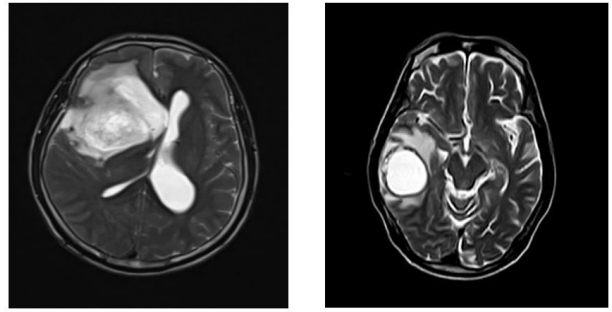


Fig. 4: Image of the T2 pulse sequence showing the patient's brain tumor

### 3.2. Results

To evaluate image quality, we use two evaluation methods: qualitative comparison and quantitative comparison.

#### Quantitative assessment

Fig. 5 shows the results achieved by the CycleGAN algorithm. It can be seen that Fig. 5B is created from the original Fig. 5A and is created by the CycleGAN model with a feature of brain tumors, Fig. 5C is the image that has removed the brain tumor feature from Fig. 5B to reconstruct the brain tumor image. The original is Fig. 5A. So during model generation, each original image generates a new MRI image, this means that each MRI image without a brain tumor produces an MRI image with a brain tumor. So with the initial data set of 660 images without brain tumors, the model created a new data set of 660 images with brain tumors.

With pulse imaging T2 MRI brain is characterized by cerebrospinal fluid with the highest signal intensity, so it is bright white, fat is light in color, gray matter is dark gray, white matter is light gray in color, and tumor cells are light in color. The brain is usually white mutated cells. Qualitative assessment by the method of visual inspection with the naked eye can see that brain tumor images generated (Fig. 5B) from images without brain tumors (Fig. 5A) are all similar in terms of characteristics of T2 image pulses, Fig. 5B shows Clear white mutant cells on the T2 pulsed tomography section.

#### Quantitative assessment

During model training, the loss function is an important issue to see if the model is good or not. The smaller the loss function, the more accurate the similarity between the generated image and the original image. Fig. 6 shows the loss function of the cyclegan model when the model using MRI image sets without brain tumors generates MRI image sets with brain tumors. We see that the loss function of the discriminator tends to decrease over the epochs (here, 100 epochs are chosen because the loss function has reached the saturation level and can no longer decrease), representing the discriminator of The GAN model increasingly fails to detect the difference between the generated image and the original image, in other words, the generated image has nearly the same features as the original image.

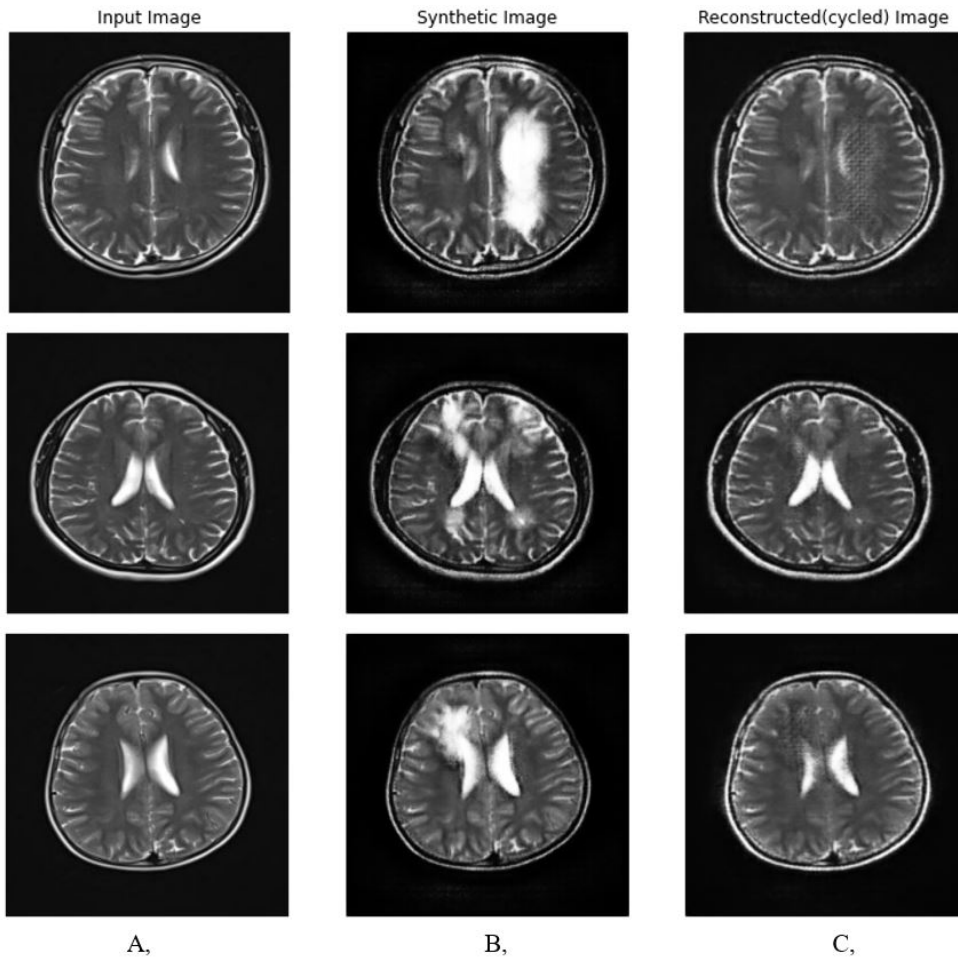


Fig. 5: Brain MRI images (B,) are generated from images without brain tumors (A,) and images restored to baseline (C,) from images with brain tumors are generated.

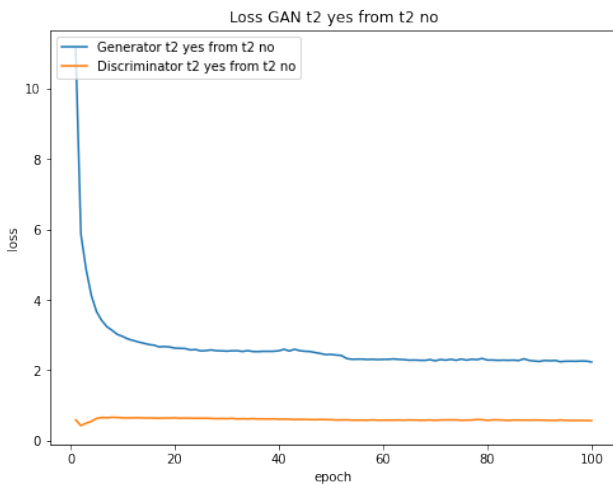


Fig. 6: The loss function of the CycleGan generates images with brain tumors from images without brain tumors.

Table 1 shows the FID score of the two generated image sets, the Generate T2 yes set is the MRI image set with brain tumors generated from the set of images without brain tumors compared with the set of MRI

images with the original brain tumor. The smaller the FID score, the lower the difference between the two data sets. With a dataset of 660 images that do not show brain tumors, the above FID score is evaluated as not too high, showing that the generated images can be used for other deep learning algorithms.

TABLE 1: Comparison of FID score with some previous works when using GAN to generate 2D MRI brain images.

Article	Algorithm	FID
Kossen, Tabea, et al., 2021 [11]	DCGAN	141.82
Li, Qingyun, et al., 2020 [12]	TumorGAN	77.43
This study	CycleGan	53.61

The results of the FID score comparison of the proposed system in our study compared with those most recently published studies are shown in Table 1. The results of that comparison are shown in Table 1. Based on this table, it can be easily seen that the

proposed system gave a FID score of 53.61 which is relatively good compared to other studies with the same subject of brain MRI, better than the DCGAN algorithm with a score of 53.61. FID is 141.82 proposed by Kossen, Tabea, et al., 2021 [11] and better with TumorGAN algorithm with FID score of 77.43 suggested by Li, Qingyun, et al., 2020 [12]. So image Brain MRI with brain tumor born from CycleGan model can completely be applied in further scientific studies.

#### 4. Conclusion

The article focuses on the application of image processing technologies such as Cyclegan network to generate new images based on the characteristics of the available image dataset, thereby enriching the data set for application in image classification and segmentation problems. After using the Cyclegan model, the generated brain tumor MRI image had a FID score of 53.61. From the present obtained results, we aim to perfect and develop the model so that we can apply the algorithm to different pulse sequences such as T1, FLAIR and DWI to increase the number of MRI images with brain tumor.

#### References

- [1] Lashkari and Amirehsan, "A neural network based method for brain abnormality detection in mri images using gabor wavelets," *International Journal of Computer Applications*, vol. 4, no. 7, 2010.
- [2] I. Goodfellow et al., "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.
- [3] J.-Y. Zhu et al., "Unpaired image-to-image translation using cycle-consistent adversarial networks," *Proceedings of the IEEE international conference on computer vision*, 2017.
- [4] P. Mildnerberger, M. Eichelberg, and E. Martin, "Introduction to the dicom standard," *Eur Radiol*, vol. 12, pp. 920–927, 2000.
- [5] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," *arXiv:1511.08458*, 2015.
- [6] P. Isola et al., "Image-to-image translation with conditional adversarial networks," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017.
- [7] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved techniques for training gans," *arXiv:1606.03498*, 2016.
- [8] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.
- [9] J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," *IEEE conference on computer vision and pattern recognition*, pp. 248–255, 2009.
- [10] 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," *arXiv:1706.08500*, 2017.
- [11] T. Kossen et al., "Synthesizing anonymized and labeled tof-mra patches for brain vessel segmentation using generative adversarial networks," *Computers in biology and medicine* 131, 2021.
- [12] Q. Li et al., "Tumorgan: A multi-modal data augmentation framework for brain tumor segmentation," *Sensors*, 2020.



**Hinh Van Nguyen** 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 also signal processing for wireless communications.



**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. From July to September 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 Radar.