International Journal of Scientific Research and Engineering Development-– Volume 6 Issue 2,2023
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Generation of Synthesized Brain MRI and Segmentation using GANs and UNET

Atharva Bendre
Computer Science Dept, Sikkim Manipal Institute of Technology, Sikkim Manipal University, Sikkim
bendre.ath@gmail.com

--************************----------------------------------

Abstract:
The algorithms for deep learning have been applied in many different sectors, however, they continue to be modified before being used in sensitive domains like medical imaging. The degree of precision ensures reliability since the employment of technological advancements in the healthcare industry is required due to time constraints. Machine learning algorithms used in the field of medicine cannot use medical data due to privacy issues. For instance, it is challenging to separate brain tumours using image segmentation due to the dearth of brain MRI images. Implementation of the generative adversarial network (GAN)-based enhancement methods led to the resolution of this problem. This paper aims to a structure for image segmentation to separate out the tumour done using UNET on brain MRI's that have been synthesized using a DCGAN.

Keywords **—brain MRI images, image generation, DCGANs, image segmentation, UNET, deep learning.**

--************************------------------------------

I. Introduction

The interpretation of medical images is a vital component in the detection, therapy, and management of several neurological illnesses. Magnetic resonance imaging (MRI) may be used to provide precise and high-resolution pictures of the brain, but these images can be difficult to obtain because of things like motion artefacts, a poor signal-to-noise ratio, and lengthy acquisition periods. Thus, the use of "Generative Adversarial Networks" (GANs)for producing high-quality brain MRI pictures was madein order to get around these difficulties. An effective deep learning architecture called GANs has been demonstrated to produce realistic images that are challenging to differentiate from genuine photographs. GANs may also be used to supplement data, which is advantageous for medical picture analysis because it enhances the quantity of data that is accessible in deep learning model training.

In this study, we concentrate on using GANs to create high-resolution brain MRI pictures and then using the U-Net architecture to segment these images. U-Net, a "convolutional neural network" created exclusively for biomedical picture segmentation, has shown to be quite successful at this task. We show that high-quality brain MRI pictures can be produced using the GANs and U-Net combo, and that these images may then be divided into several regions of interest. With this method, it may be possible to diagnose and plan treatments for a variety of neurological disorders more quickly and accurately.

II. Literature Review

The proposal of GANs by Goodfellow et al [1], presented a novel algorithm which was aimed at solving problems which required generative modelling using adversarial processes. The authors achieved the estimation of generative models by developing an architecture to train a discriminator model and a generator model at the same time.

The generator model was responsible of capturing the image data and the discriminator model to check its similarity to the real data. The designed algorithm used a minmax setting where the generative model minimized the distribution meanwhile the discriminator was used to maximize the same. The algorithm was successfully able to generate synthesised images using the TFD, MNIST and CIFAR-10 datasets.

To implement GANs to medical images Iqbal, Talha, and Hazrat Ali H [2] proposed a new GAN model called the "MI-GAN". The new model was used to generate retinal images using the segmented images of the mask and the filamentary structures. The authors used 2 datasets namely "DRIVE" and "STARE" each consisting of 10 images each. The model utilized and encoder decoder technique to introduce the noise code. This new architecture was able to achieve a dice coefficient of 0.832 and 0.837 on the two datasets respectively thereby outperforming the pre-existing models.

There are multiple GANs which can be used to generate medical images. The experimentation by Han, Changhee, et al [3] MRIs was used to understand the conditions for selection of DC-GANs versus WGANs while generating multi-sequence brain MRIs. It also presented ideas on the exploitation of medical images which have "intrinsic intra-sequence variability towards GAN-based data augmentation". The authors made use of the 2016 "BRATS" dataset created for "Multimodal Brain Tumour Image Segmentation Benchmark Challenge". By getting the images tested by a professional physician the experimentation concluded that WGANs are better suited for generation of such images when the training time is not an important metric or when there are multiple classes of images available.

It is the lack of datasets which has created this motivation of using GANs to generate synthetic datasets. The experimentation by C. Bowles et al [4] shows the viability of adding GAN-derived manufactured data to the training datasets. The authors create the synthetic data using the "Progressive Growing of GANs (PGGAN) network" as it is stable for images with huge sizes and because it is robust why selecting hyperparameters. This PGGAN was trained on a dataset consisting of 80,000 images after a 10-90% reduction in percentage. The generated images were then used as input to train the UNET and UResNET for segmentation of these images.

The most recent work done by H.H.NAlrashedy et al [5] presented a novel automated framework called the "BrainGAN". It was created with the aim to generate brain MRIs and the perform classification of real images using CNN models by training them on the manufactured dataset. The experiment made use of two GAN models namely "Vanilla GAN" and a "DCGAN" to generate two different sets of brain MRI images. These generated datasets consisted of two classes, one with brain tumours and one without. The framework then utilized three models ("Simple CNN", "MobileNetV2", "ResNet152V2"). These models were able to successfully classify the images with accuracies of 96.63%, 95.85%, and 99.09%. It was then concluded that the "ResNet152V2" was the most suitable model for the task.

After the generation of if the synthetic images they are utilized for either classification or segmentation. O. Ronneberger, P. Fischer, and T. Brox[6] proposed a new network to perform segmentation on biomedical images. The proposed network was a "U-Shaped" architecture made up of two paths which were symmetric to each other. The first path was contractive in nature i.e., performed downsampling and was responsible for capturing the context of the images passed through it. The second path was expansive i.e., performed upsampling in nature and was responsible for localization with precision. The experiment made use of "DIC-HeLa" and "PhC-U373" datasets achieving an average IOU of 77.5% and 92% on the respective datasets.

III. PROPOSED METHOD

A. *Basic Approach*

The algorithm is divided into two major processes, namely image generation and image segmentation.

The first step is the synthesis of images. Here we use a Deep Convolutional Generative Adversarial Network (DC GAN) to generate brain MRI images. In this implementation the algorithm only generates MRI images which have a positive indication of a brain tumour. The brain MRIs without a tumour are negated as the next step in the algorithm is the segmentation of the tumour from the brain MRI. Thus, for simplicity's sake the dataset used to train the generator and the generated dataset is made up of MRIs which contain a tumour.

The training requires the original image and its segmented image the current dataset can be called incomplete. Hence To complete this generated dataset so that it can be used further, binary segmented images or masks for the tumour were generated using an online software called "lablebox". The generated dataset along with its segmented images is then used as input to train the UNET architecture. Upon completion of training the model is then tested using a new MRI dataset to test its performance on real MRIs after being trained on a synthesised dataset.

IV. MODEL AND LEARNING

A. *DCGANs and UNET*

The experimentation makes use of a simple DCGAN to generate synthesised images. Just like vanilla GANs DCGANs are also made up of two parts (generator and discriminator). Where the generator is responsible for generating the augmented image and the discriminator to check the authenticity of the generated image. The model for the generator and the discriminator used comprise of multiple convolutional layers with leaky ReLU used as the activation function between each of these layers. In addition to the above configuration the discriminator also consists of a combination of a flattened layer, a dropout layer and finally a dense layer towards the end of its structure.

The process begins when a random normal distribution is sent to the generator. Since the generator does not have a point of reference it produces another random normal distribution.

Simultaneously, an authentic instance or empirical evidence is used as inputfor the discriminator, thereby allowing it tograsp the distribution of the true input. The discriminator then asses the distribution of the outputof the generator.

If the created sample's distribution closely resembles that of the sample that was originally used, the discriminator will produce a result that is near to '1' or authentic. The discriminator produces a value that is near to '0' or faux if the distributions don't match or are not close enough to appear the same. The distance between the real and the generated image is calculated using a loss function in a minmax setting. This means that the generator tries to minimize the loss function to produce the most authentic image possible whereas the discriminator tries to maximize this loss function to differentiate between the real and the created image.

The minmax setting used is given by:

$$\nabla_{\theta d}(1/m) \sum [logD(x^i)+ .log(1-D(G(z^i)))] \tag{1}$$

UNET is the alternative model in use. An encoder network and a decoder network make up the U-shaped architecture that it uses to function. To create a segmentation map, the decoder employs the high-level characteristics that the encoder extracts from the input picture. The accuracy of segmentation is increased by the U-shaped architecture's ability to integrate high level as well as the low level characteristics. Utilising a pixel-wise cross-entropy loss function, the network is tuned during training.

B. *Architecture*

The generator consists of an input layer, 3 convolution layers (64 filters of 128X128, 128 filters of 64X64, 128 filter of 32X32, 256 filter of 32X32), with leaky ReLu used as the activation function.The dimensions of each layer are represented in fig.1.

```
Model: "generator"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense_1 (Dense)             (None, 262144)            26476544

 leaky_re_lu_4 (LeakyReLU)   (None, 262144)            0

 reshape (Reshape)           (None, 32, 32, 256)       0

 conv2d_transpose (Conv2DTra  (None, 64, 64, 128)      524416
 nspose)

 leaky_re_lu_5 (LeakyReLU)   (None, 64, 64, 128)       0

 conv2d_transpose_1 (Conv2DT  (None, 128, 128, 128)    262272
 ranspose)

 leaky_re_lu_6 (LeakyReLU)   (None, 128, 128, 128)     0

 conv2d_4 (Conv2D)           (None, 128, 128, 1)       2049

=================================================================
Total params: 27,265,281
Trainable params: 27,265,281
Non-trainable params: 0
```

Fig. 1. Generator

The discriminator consists of 4 convolution layers (128 filters of 64X64, 128 filters of 128X128, 1 filter of 128X128), with leaky ReLu used as the activation function. Followed by a flatten and a dense layer of 1X65536. The dimensions of each layer are represented in fig.2.

```
Model: "discriminator"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 128, 128, 64)      640

 leaky_re_lu (LeakyReLU)     (None, 128, 128, 64)      0

 conv2d_1 (Conv2D)           (None, 64, 64, 128)       73856

 leaky_re_lu_1 (LeakyReLU)   (None, 64, 64, 128)       0

 conv2d_2 (Conv2D)           (None, 32, 32, 128)       147584

 leaky_re_lu_2 (LeakyReLU)   (None, 32, 32, 128)       0

 conv2d_3 (Conv2D)           (None, 16, 16, 256)       295168

 leaky_re_lu_3 (LeakyReLU)   (None, 16, 16, 256)       0

 flatten (Flatten)           (None, 65536)             0

 dropout (Dropout)           (None, 65536)             0

 dense (Dense)               (None, 1)                 65537

=================================================================
Total params: 582,785
Trainable params: 582,785
Non-trainable params: 0
```

Fig. 2. Descriminator

Fig.3 describes the overall architecture of the DCGAN used.

```
Model: "gan_model"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 input_1 (InputLayer)        [(None, 100)]             0

 generator (Sequential)      (None, 128, 128, 1)       27265281

 discriminator (Sequential)  (None, 1)                 582785

=================================================================
Total params: 27,848,066
Trainable params: 27,265,281
Non-trainable params: 582,785
_________________________________________________________________
```

Fig. 3. DCGAN

The UNET architecture is the same one proposed in [6]. However, the avgpooling layers are replaced with maxpooling layers to provide a better result.

V. EXPERIMENT

A. *Dataset*

The experimentation uses 3 datasets, out of which the first dataset consists of 155 brain MRI images sourced from an open source Kaggle library [7]. Only those MRI images are used to construct this database which test positive for a brain tumour. The second dataset consists of images that have been generated using the DCGAN and the segmented/mask images. It consists of 200 images of each the aforementioned categories used for training the UNET. The third and final dataset is consisting of both the MRI images with brain tumour and its masks. It was sourced from [8]. This dataset consists of 3000 images of each class out of which only 50 were used for testing.

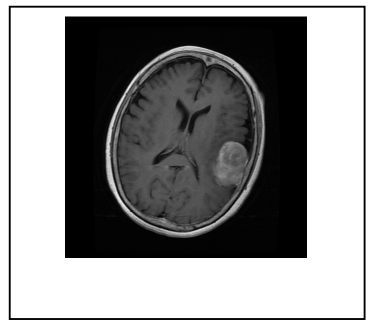

Fig. 4. First Dataset

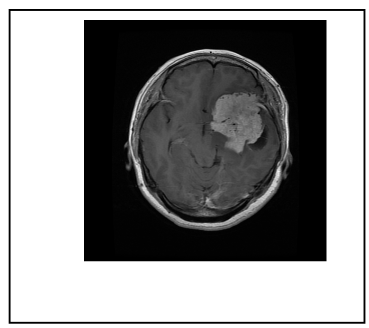

Fig. 5. Generated Dataset

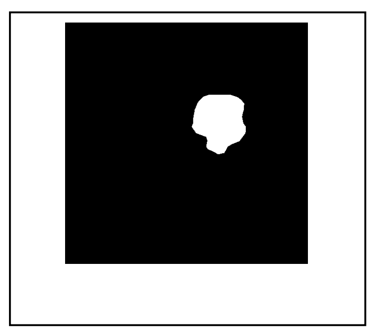

Fig. 6. Masks for generated dataset

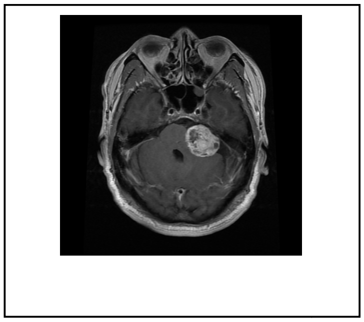

Fig. 7. Third dataset

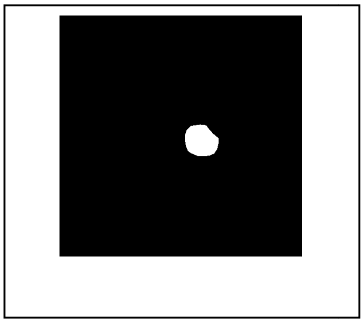

Fig. 8. Predicted Mask for third dataset

A. Process

First the input images from the first dataset are processed to change their dimensions to 128X128. During the training phase these images are fed into the DCGAN for a total of 10 epochs, with 3750 steps per epoch and a batch size of 4 thereafter the noise dimension of the DCGAN was set to 100. The generator generates 200 synthesised images which are tested for authenticity by the discriminator. This generated dataset is then used to create binary segmented images / masks using “lablebox” software. Once the masks for all 200 images are generated the dataset along with the respective masks for the MRI images used as input by UNET for training.

The third dataset obtained (fresh dataset which consist of MRI images and masks) is then used to test the accuracy of the UNET when it performs segmentation.

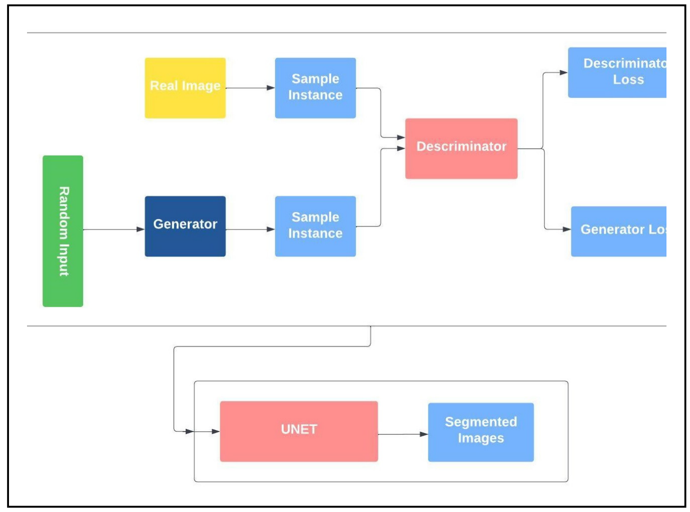

Fig. 9. Flow Diagram

VI. RESULTS

The model was successfully able to produce synthesized images and achieved a generator loss and discriminator loss of 3.4853 and 0.0883 respectively as described in fig.10.

The DCGAN was tested by plotting the distribution of the generated and the rea images displayed in fig.11. As seen in fig.11 the distributions overlap, this indicates the generated samples capture most of the variations of the real images and thus can be considered as the true representations of the real images.

TABLE I

Total Epochs	*Generator loss*	*Discriminator loss*
10	3.4853	0.0883

Fig. 10. Generator and Discriminator Loss

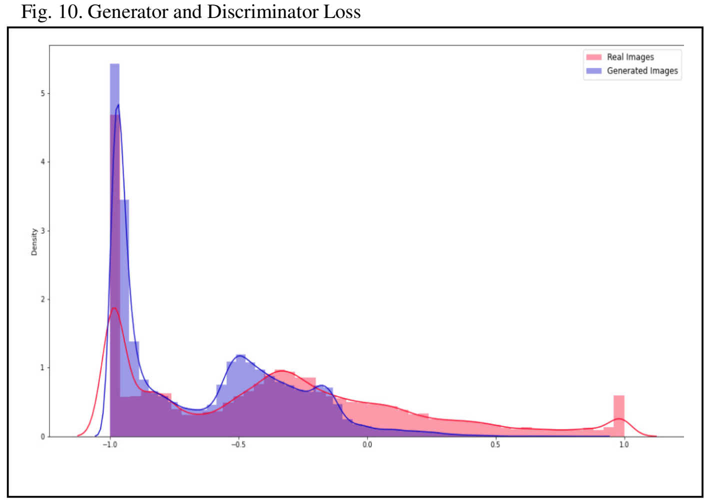

Fig. 11. Distribution Graph

The third dataset consisting of 50 MRI and masks was used to test the UNET. The UNET model achieved an accuracy of 96%. The results of the incorrectly segmented and correctly segmented images are described by fig.12. Fig.13 and Fig.14 represent the loss and accuracy graphs of UNET while training and validation across 8 epochs.

TABLE III

Total Images	*Correctly Segmented*	*Incorrectly Segmented*
50	48	2

Fig. 12. Table for UNET results

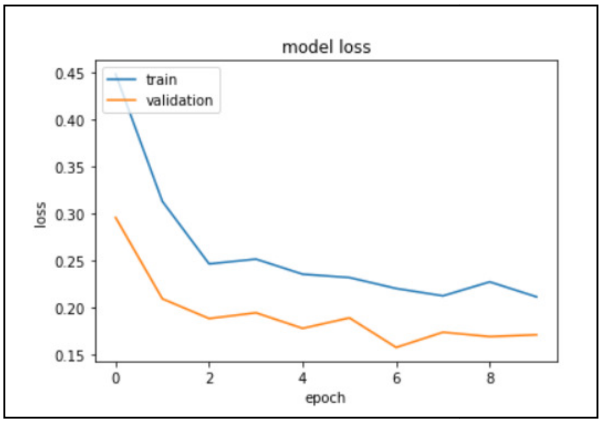

Fig. 13. Loss Graph

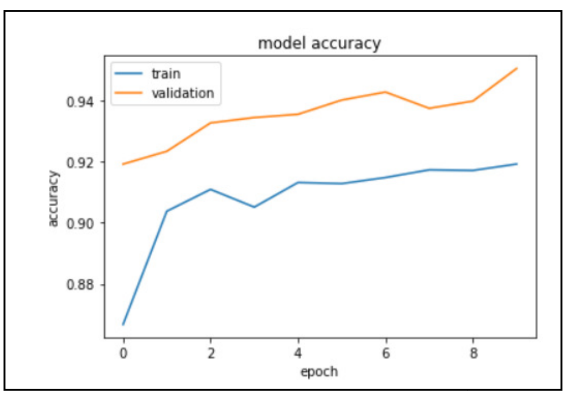

Fig. 14. Accuracy Graph

VII. CONCLUSION

From the experimentation conducted we can conclude that the use "GenerativeAdversarialNetworks" (GANs) has the protentional of absolving the shortage of medical imagery faced by this sphere of technological development as it is capable of producing true representations of the real-world instances.

It was seen that the use of DCGAN is suitable for this setup as it only considers images with a tumour i.e., DCGAN sometimes runs into problems like model collapse when used for multiple classes, hence a WGAN can be used upon increase in the number of classes.

VIII. FUTURE SCOPE

The current framework involves creating the segmented images using a third party software. This is a drawback which can be improved upon by upgrading the DCGAN not only produce the MRI images as shown in this experiment but to also generate their respective segmented images as achieved in [9].

This can not only be applied to brain MRIs but also to other MRIs there prying open the door to technological advancements by ebbing the scarcity of datasets.

Use of WGANs can also be made to improve the distribution as well as the quality of the images obtained.

ACKNOWLEDGMENT

I would like to thank Sikkim Manipal Institute of Technology, Sikkim Manipal University for their support as this research was conducted in affiliation to their specialization program at.

Also, would like to thank the Kaggle community for providing the datasets which were used for the experimentation.

REFERENCE

[1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). "Generative adversarial networks." Communications of the ACM, 63(11), 139-144.

[2] Iqbal, T., & Ali, H. (2018). "Generative adversarial network for medical images (MI-GAN)". Journal of medical systems, 42, 1-11.

[3] C. Han et al., "GAN-based synthetic brain MR image generation," 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 2018, pp. 734-738, doi: 10.1109/ISBI.2018.8363678.

[4] Bowles, C., Chen, L., Guerrero, R., Bentley, P., Gunn, R., Hammers, A., ... & Rueckert, D. (2018). "Gan augmentation: Augmenting training data using generative adversarial networks". arXiv preprint arXiv:1810.10863.

[5] Alrashedy, H. H. N., Almansour, A. F., Ibrahim, D. M., & Hammoudeh, M. A. A. (2022). "BrainGAN: Brain MRI Image

Generation and Classification Framework Using GAN Architectures and CNN Models". Sensors, 22(11), 4297.

[6] Ronneberger, O., Fischer, P., & Brox, T. (2015). "U-net: Convolutional networks for biomedical image segmentation". In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18 (pp. 234-241). Springer International Publishing.

[7] https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection

[8] https://www.kaggle.com/datasets/nikhilroxtomar/brain-tumor-segmentation

[9] Neff, T., Payer, C., Stern, D., & Urschler, M. (2017, May). Generative adversarial network based synthesis for supervised medical image segmentation. In Proc. OAGM and ARW joint Workshop (Vol. 3, p. 4).

