

# Generative Adversarial Networks-based solution for improving medical data quality and insufficiency

**Chenxi Gao**

Department of Computer Science, University of Wisconsin Madison, Madison, 53715, United States

cgao65@wisc.edu

**Abstract.** As big data brings intelligent solutions and innovations to various fields, the goal of this research is to solve the problem of poor-quality and insufficient datasets in the medical field, to help poor areas can access to high quality and rich medical datasets as well. This study focuses on solving the current problem by utilizing variants of generative adversarial network, Super Resolution Generative Adversarial Network (SRGAN) and Deep Convolutional Generative Adversarial Network (DCGAN). In this study, OpenCV is employed to introduce fuzziness to the Brain Tumor MRI Dataset, resulting in a blurred dataset. Subsequently, the research utilizes both the unaltered and blurred datasets to train the SRGAN model, which is then applied to enhance the low-quality dataset through inpainting. Moving forward, the original dataset, low-quality dataset, and the improved dataset are each used independently to train the DCGAN model. In order to compare the difference between the produced image datasets and the real dataset, the FID Score is separately computed. The results of the study found that by training DCGAN with SRGAN repaired medical dataset, the naked eye can observe that the medical image dataset is significantly clearer and there is a reduction in Fréchet Inception Distance (FID) Score. Therefore, by using SRGAN and DCGAN the current problem of low quality and small quantity of datasets in the medical field can be solved, which increase the potential possibilities of big data in artificial intelligence filed of medicine.

**Keywords:** Generative Adversarial Networks, Brain Tumor, Deep Learning.

## 1. Introduction

In this era of big data, the application of data spans across nearly every aspect of life, bringing more intelligent solutions and innovations to various scenarios and achieving unexpected results. For instance, in the field of telecommunications, these data can facilitate operators in predicting and balancing network load; in the field of energy, meteorologists can predict the production of wind, light, and other renewable energy based on data such as weather and seasons [1]. The abundance of data allows individuals to glean fresh insights and observations that contribute to the advancement of societal infrastructure. At the same time, high-standard data quality is crucial under most scenarios, and low-quality data can result in poor decisions and the loss of significant commercial prospects [2]. According to the KPMG 2017 Global CEO Outlook, 56% of CEOs were concerned about the reliability and caliber of the data supporting their decision-making processes [3].

Compared with other fields, the medical adheres to a more stringent and solemn standard. Datasets of low quality are likely to lead to wrong judgments for applications using these datasets. Also, the low-quality data and misleading information may directly affect the health and life of patients. Now artificial intelligence has achieved medical support, enhanced doctors' diagnosis, and improved the efficiency of hospital visits [4-6]. If these AI applications can use higher quality and richer datasets, it would undoubtedly result in a mutually beneficial arrangement for both patients and healthcare institutions. However, there are two common problems in the datasets used for medical treatment, namely poor quality and small datasets. Poor dataset quality tends to occur in poor areas. Medical resources in impoverished areas are very limited. The equipment in poor areas is relatively backward, and image quality is often degraded or lost [7]. Due to the expensive expense of getting medical images [8] and the fact that most hospitals do not share data [7], there is a dearth of data, which prevents some large models from being trained. In terms of the two common issues currently faced, Generative Adversarial Networks (GANs) in the field of artificial intelligence have gained significant attention and popularity due to their capacity to produce accurate data. Among them, one of GAN's variants called Super Resolution Generative Adversarial Network (SRGAN) can repair the image and make the image clearer; and a variant called DCGAN is often applied in the field of image generation to generate more image data [9].

Due to the rise of machine learning in various fields, researchers are more likely to study the model and develop its functionality but ignore the enhancement of the dataset [10]. However, the dataset is the cornerstone of further model training. The model iteration is very limited without a high-quality and sufficient dataset as the research basis.

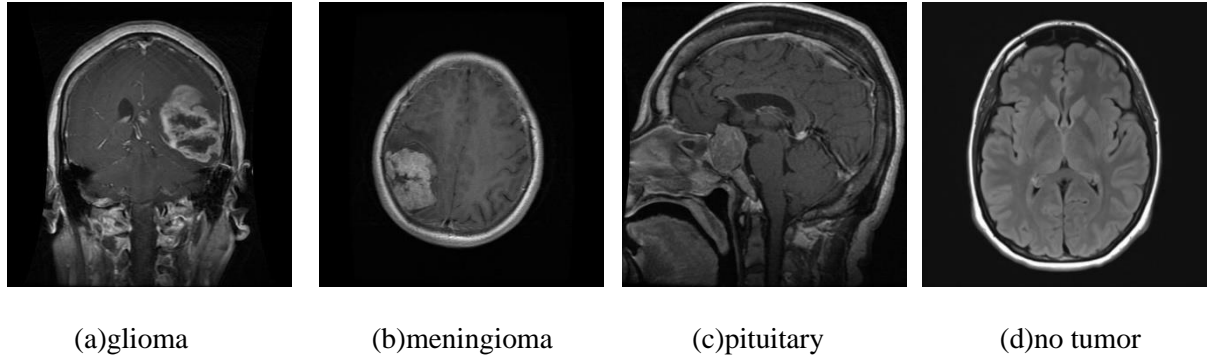
Therefore, this paper focuses on improving the two aspects of poor dataset quality and insufficient data volume of medical imaging datasets. The research process will be divided into two phases. The first phase will utilize SR-GAN to optimize poor-quality datasets. The second stage uses DC-GAN to verify the improvement of the dataset's quality and increase the amount of data. The study will obtain a lower-quality dataset by blurring and downscaling the original high-quality dataset.

Subsequently, the SRGAN technique is used to optimize this poor-quality dataset. The research will use these three datasets: the original high-quality dataset, the fuzzed dataset, and the optimized dataset, which are used to train the DCGAN model. Compare the three DCGAN training outcomes to determine if the optimized dataset can considerably improve the model's ability to generated high-quality images.

## 2. Method

### 2.1. Dataset preparation

The research will use the Brain Tumor MRI Dataset in Kaggle [11] for the main training task. This dataset contains different images of brain tumors, which are abnormal cell aggregates or masses that arise inside the brain. Any different forms of brain tumors can cause increased pressure within the skull, which can cause brain damage [11]. The Brain Tumor MRI Dataset [11] consists of the training dataset and the testing dataset. The test dataset contains 1,311 images while the training dataset contains 5,712 images, each with a size of 512×512 pixels. These include four different image data of glioma, meningioma, pituitary, and no tumor in both the testing and training datasets. Below are examples of images from different brain tumor datasets. Figure 1 displays a few sample images.



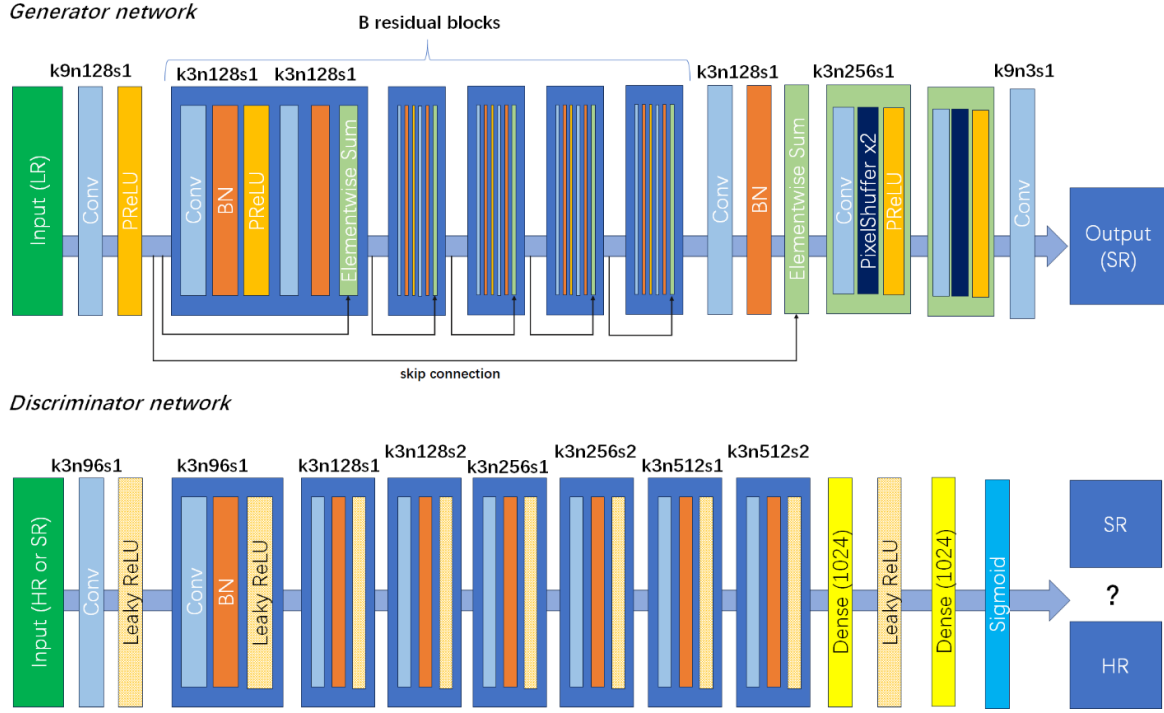
**Figure 1.** The sample images in four datasets [11].

Since the minimum pixel size of the input image for SRGAN in the following experiment is  $384 \times 384$ , in the preparation of data preprocessing, this study converts the  $512 \times 512$ -pixel images in the original dataset to  $384 \times 384$  pixel images. Then, the blur function in OpenCV was employed to blur the images by setting the filter kernel to  $5 \times 5$  and  $2 \times 2$  respectively to get the blurred and low-resolution dataset: the dataset obtained by the filter kernel of  $5 \times 5$  will be abbreviated as SRGAN-Train Dataset in the following, and filter kernel of  $2 \times 2$  will be abbreviated as Primary Dataset.

## 2.2. SR-GAN

SRGAN is a generative adversarial network for image super-resolution, which has a remarkable effect in improving image quality, adding details, and recovering missing information [12]. The SRGAN model used in this study was first proposed by researcher Christian Ledig and other researchers [12]. The code used for implementation was obtained from the GitHub repository provided by seven contributors [13]. The experiment followed the instructions provided in the repository's README file to set up and adapt the code for the specific dataset in this research.

The structure of the generator in the SRGAN used in this study consists of two convolutional layers and two pre-trained sub-pixel convolutional layers shown in Figure 2. In the convolutional layer, the kernel size is  $3 \times 3$ . And the size of the feature maps is 128. Each convolutional layer is followed by a bn layer, and the activation function used after the bn layer is ParametricReLU. The structure of the discriminator contains eight convolutional layers and two dense layers. The size of the kernels in the convolutional layer is  $3 \times 3$ , the size of the feature maps is 128, and LeakyReLU is used as the activation function after each convolutional layer. The final layer of the discriminator uses sigmoid as the activation function to obtain a classification likelihood.



**Figure 2.** The network structure of SRGAN (Photo/Picture credit: Original).

In this study, the input size of the feature map of the first layer in the original generator was modified from 64 to 128, and the feature map of the first layer in the original discriminator was changed from 64 to 96. This change allowed the model to capture finer image details. During the training process, to allow the trained SRGAN to repair images better, the experiment's batch size was set to 24 and the epochs was set to 360. In addition, this study also used the Momentum optimizer, and the learning rate was set to 0.02, and the step size, gamma, and Momentum was separately set to 1000, 0.1, and 0.9. These are to overcome obstacles like local minima and noisy gradients, thereby accelerating the optimization process and leading to faster convergence [14].

The following experiment will use the original Brain Tumor MRI Dataset and the SRGAN-Train Datasets modified based on the original dataset to train SRGAN. The training process mainly includes putting the images in the SRGAN-Train Datasets into the SRGAN generator to perform image super-resolution processing to make the images clearer. Subsequently, both the high-quality image dataset and the original dataset were fed into the discriminator, which assessed whether these images were produced by the generator. This study expects that using SRGAN-Train Datasets to train SRGAN will make SRGAN more capable of repairing blurred images. After the SRGAN training finished, the Primary-Dataset will be passed into the trained SRGAN to repair it, and get the repaired Primary-Deblurred Dataset, which will be used for subsequent DCGAN training.

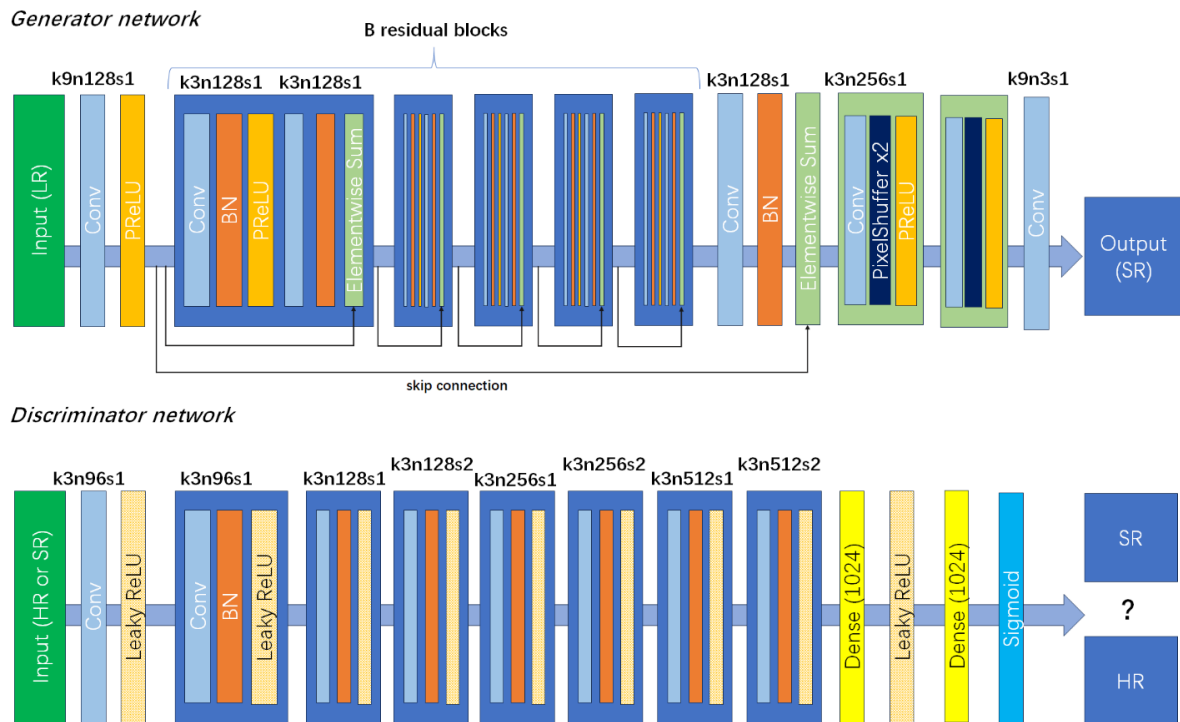
### 2.3. DCGAN

DCGAN is a variant of GAN optimized using deep neural network CNN, which is important in the field of image generation [15]. This research will implement DCGAN based on the study [15], then trained with the Brain Tumor MRI Dataset, Primary-Dataset, and Primary-Deblurred Dataset to demonstrate that the optimized dataset using SRGAN results in a higher quality of the images generated by DCGAN.

The generator's structure in the DCGAN used in this research contains five convolutional layers shown in Figure 3. After the first four convolutional layers, there is a bn layer. After the bn layer, ReLU will be used as the activation function, and the last layer of the generator will use Tanh as an activation function. The kernels size of the first four convolutional layers is 4×4, and the kernels size of the last

convolutional layer is set to  $10 \times 10$  to generate an image of  $256 \times 256$  pixels. The feature maps from the first to the fifth convolutional layer correspond to 512, 256, 128, 64, and 1, respectively. The feature map size of the fifth layer is set to 1 to generate a grayscale image with only one-color channel.

Additionally, the structure of the discriminator contains five convolutional layers. The second, third, and fourth convolutional layers are followed by a bn layer. Leaky ReLU will be used as the activation function after the first convolutional layer and the bn layers of the second, third, and fourth convolutional layers. The kernels of the first convolutional layer are  $10 \times 10$ , and the subsequent four convolutional layers are  $4 \times 4$ . The corresponding feature maps of these five convolutional layers are 64, 128, 256, 512, and 1. The last layer of the discriminator, sigmoid will be used as an activation function to obtain a likelihood value for classification.



**Figure 3.** The network structure of DCGAN (Photo/Picture credit : Original).

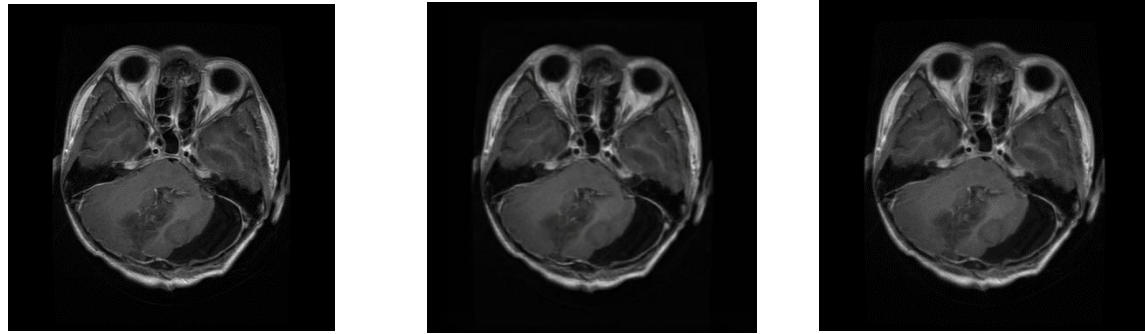
In order to allow DCGAN to generate images better, the experiment set the batch size to 128 and the epoch to 1000. In addition, DCGAN used the Adam optimizer and set the hyperparameters in Adam as follows: Learning rate = 0.0002,  $\beta_1=0.5$ ,  $\beta_2=0.999$ . In the experiment, DCGAN will be trained using three datasets: the Brain Tumor MRI Dataset, Primary-Dataset, and Primary-Deblurred Dataset. After the training of the three datasets is completed, the performance of the three models will be evaluated.

### 3. Results and discussion

#### 3.1. The performance of models

##### 3.1.1. SRGAN result

Observing the samples in Figure 4, it can be observed that the Primary-Deblurred Dataset closely resembles the original dataset, with little discernible difference. but the Primary-Deblurred Dataset has higher resolution than the images in the Primary-Dataset.



(a) Brain Tumor MRI Dataset (b) Primary-Deblurred Dataset (c) Primary-Dataset

**Figure 4.** The sample images in three datasets (Photo/Picture credit : Original).

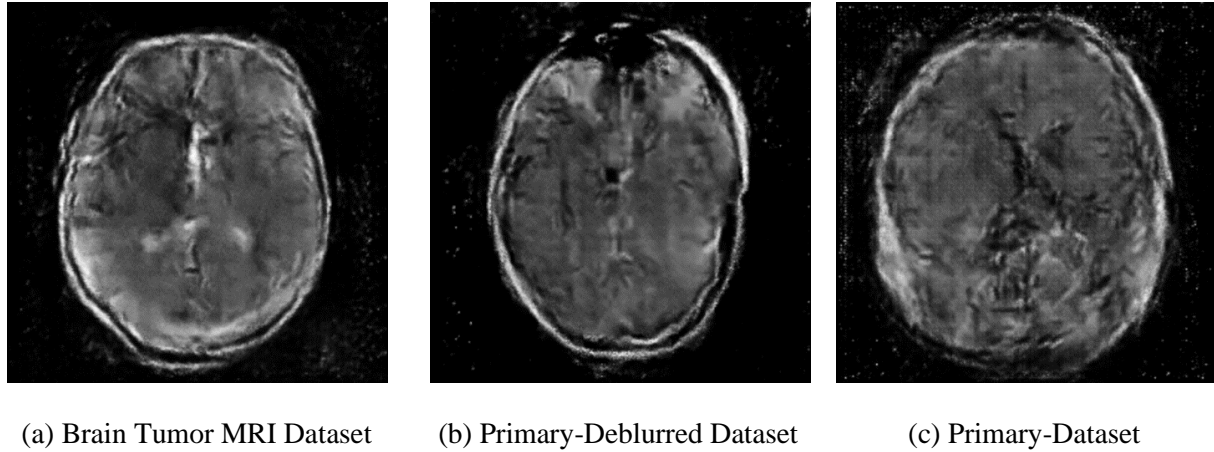
**Table 1.** The SSIM and PSNR of Primary-Deblurred Dataset and Primary-Dataset.

	SSIM	PSNR
<b>Primary-Deblurred Dataset</b>	0.8646	30.5532
<b>Primary- Dataset</b>	0.9493	32.3137

In this study, SSIM and PSNR were calculated using the original dataset Brain Tumor MRI Dataset for Primary-Deblurred Dataset and Primary-Dataset, respectively [16]. The experimental results in Table 1 show that the PSNR of the Primary-Deblurred Dataset is 5% lower than the Primary-Dataset, and the difference in SSIM values is not significant. Although the PSNR and SSIM values become smaller, some PSNR and SSIM values may be sacrificed due to the perceptual loss introduced by SRGAN in pursuit of better visual effects. It indicates that the SRGAN model obtained from the training in this study effectively repairs the images in Primary-Dataset and meets the expectations.

### 3.1.2. DCGAN result

Observation of the sample images in Figure 5 reveals that the DCGAN model trained with the original dataset generates the best quality images, and the DCGAN model trained with the Primary-Dataset generates the worst quality images. This study used the original Brain Tumor MRI Dataset to calculate the fid score for three DCGAN models (original dataset, Primary-Deblurred Dataset, and Primary-Dataset). According to the experimental findings in Table 2, the fid score of the DCGAN model trained with Primary-Deblurred Dataset is 14.3% lower than that of the DCGAN model trained with Primary-Dataset. Still, it is 32.1% higher than the DCGAN model trained with the original dataset. The experimental results are consistent with the expectations, which shows that the DCGAN model trained with the repaired dataset can generate higher-quality images.



**Figure 5.** The sample generated images are in three datasets (Photo/Picture credit: Original).

**Table 2.** The FID Score of Brain Tumor MRI Dataset Model, Primary-Deblurred Dataset Model and Primary-Dataset.

Dataset Name	FID Score
Brain Tumor MRI Dataset	206.0146
Primary-Deblurred Dataset	272.1493
Primary-Dataset	317.5922

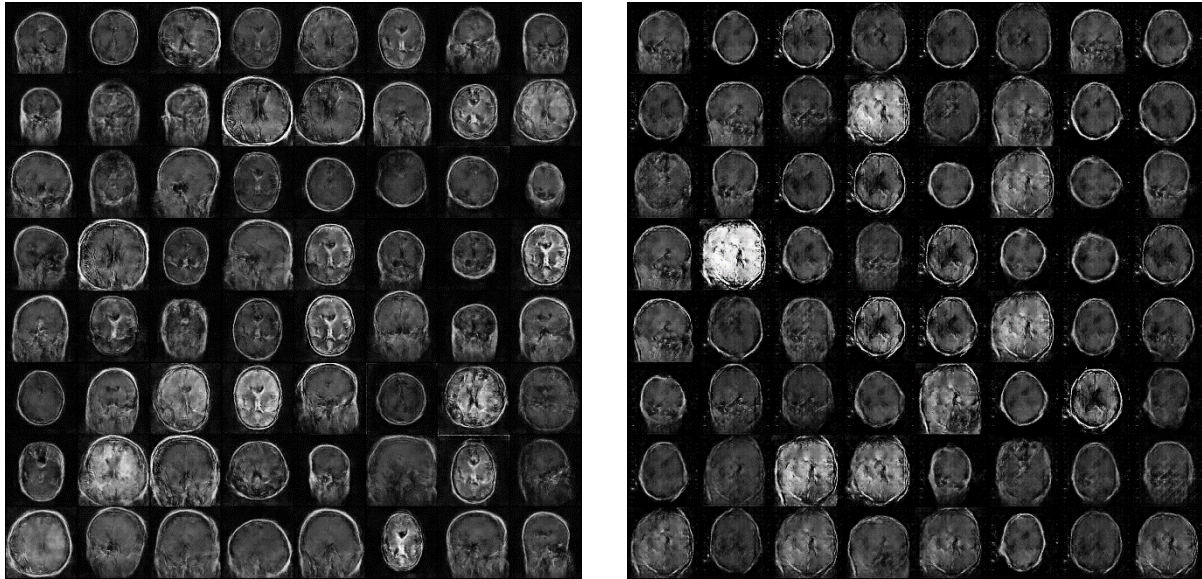
### 3.2. Discussion

Since SRGAN can repair missing details in low-quality images by learning from high-quality image datasets. Therefore, using the trained SRGAN model to optimize the Primary Dataset can effectively enhance the blurred image's details and improve the Primary Dataset's quality.

DCGAN generates new images by learning from the image dataset. If the images in the dataset are clear, DCGAN is able to learn more detailed information, resulting in clearer and more realistic images. On the contrary, if the images in the dataset are blurred, the information learned by DCGAN will be limited, and the generated images will not be clear enough and some details will be lost. Comparing the sample in Figure 5, it can be found that the details and structure of the Primary-Deblurred Dataset's image are clearer than those in the Primary Dataset's. Therefore, using the Primary-Deblurred Dataset to train DCGAN can improve the quality of generated images.

During this research, it was also found that the Primary-Deblurred Dataset allows DCGAN to generate clear images earlier in the training process than using the Primary-Dataset for training.





(a) Primary-Deblurred Dataset training results after 122 epochs (b) Primary-Dataset training results after 211 epochs

**Figure 6.** The sample generated images during DCGAN training (Photo/Picture credit : Original).

As can be observed from Figure 6, when the DCGAN was trained using the Primary-Deblurred Dataset in this study, a relatively clear image was generated at 122 epochs, the image background appears entirely black, and the main portion of the image exhibits minimal noise. The approximate effect is only available at 211 epochs when using the Primary Dataset for training. It also shows that using high-quality datasets can allow DCGAN to learn useful details faster, thereby reducing training time and training costs.

The study shows that SRGAN can enhance the resolution. However, some medical imaging images will lose information in practice due to resolution enhancement. In other words, the enhanced image may change the original organs and tissues, so in this case, the enhancement effect of SRGAN is counterproductive. In the future, using SRGAN to improve specific areas is required, notably where lesions are present, to enhance its suitability for practical medical applications. Subsequently, researchers can utilize DCGAN to generate additional medical data that encompasses the necessary information.

#### 4. Conclusion

This study found that training SRGAN and DCGAN variants of the generative adversarial network can solve the problem of using high-quality datasets in areas with inadequate medical resources and insufficient medical datasets. Therefore, in the contemporary era where medical resources are unevenly distributed, promoting SRGAN to enhance substandard medical datasets and using DCGAN to expand medical datasets could enable hospitals in most areas to access high-quality datasets. And sufficient medical datasets can assist physicians in more accurate disease diagnosis and analysis. At a social level, the outcomes of this study will aid in addressing the issue of unequal distribution of medical resources, while at the citizen level, it can facilitate precise comprehensive diagnoses for individuals in areas lacking resources. Finally, from a research standpoint, researchers can train large-scale models using abundant datasets to assist in updating and refining the healthcare system. In the future, this research will center on utilizing SRGAN to rectify targeted lesions, enabling it to have a vital role in medical applications with complex scenarios.



## References

- [1] Lundberg L and Håkan G 2022 Research Trends, Enabling Technologies and Application Areas for Big Data, *Algorithms*, 15(8), 280.
- [2] Cichy C and Stefan R 2019 An Overview of Data Quality Frameworks, *IEEE Access*, 7, 24634-24648.
- [3] Disrupt and Grow 2017 (KPMG)
- [4] DonHee L and Seong N Y 2021 Application of Artificial Intelligence-Based Technologies in the Healthcare Industry: Opportunities and Challenges *International Journal of Environmental Research and Public Health*, 18(1), 271.
- [5] Qiu Y Wang J Jin Z Chen H Zhang M and Guo L 2022 Pose-guided matching based on deep learning for assessing quality of action on rehabilitation training *Biomedical Signal Processing and Control* 72 103323
- [6] Zhou X Li Y Liang W 2020 CNN-RNN based intelligent recommendation for online medical pre-diagnosis support *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 18(3) pp 912-921
- [7] European Society of Radiology (ESR) 2014 Renewal of radiological equipment Insights imaging 5 pp 543-546.
- [8] Zhuoling L et al 2019 CLU-Cnns: Object Detection for Medical Images (*Neurocomputing*) p 53-59
- [9] Dash N et al 2021 A Review of Generative Adversarial Networks (Gans) and Its Applications in a Wide Variety of Disciplines -- from Medical to Remote Sensing, *arXiv preprint arXiv:2110.01442*.
- [10] Sambasivan N et al 2021 everyone Wants to Do the Model Work, Not the Data Work’: Data Cascades in High-Stakes AI, In *proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* pp 1-15
- [11] Nickparvar and Masoud 2021 Brain Tumor MRI Dataset, *Kaggle*, <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>
- [12] Ledig C et al 2017 Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network, In *Proceedings of the IEEE conference on computer vision and pattern recognition* pp 4681-4690
- [13] Hao D et al 2022 SRGAN (GitHub)
- [14] Hao L et al 2020 Rethinking the Hyperparameters for Fine-Tuning, *arXiv preprint arXiv:2002.11770*
- [15] Radford A et al 2016 Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, *arXiv preprint arXiv:1511.06434*.
- [16] Horé A and Ziou D 2010 Image Quality Metrics: PSNR vs. SSIM In *2010 20th international conference on pattern recognition* pp 2366-2369 *IEEE*