PROJECT REPORT

on

Medical Image Enhancer Using GAN’s

School of Computer Science and Engineering
Vellore Institute of Technology, Vellore.

April 13, 2023

Table of Contents

Abstract

Generative Adversarial Networks(GANs) were first proposed to help in image generation using neural networks. With GANs continuing to expand in their scope and ability, there have been vast strides made in improving the resolution of medical imagery. The issue, however, is with the true loss of the extrapolation and how useful it can be in medical imagery. To that end, we introduce a novel SeSiGAN architecture which uses binary pairwise labels coupled with a sequential layer to get images closer to the ground truth from the first epoch. The images obtained from SeSiGAN were then used to build a dataset. The visual quality of these were evaluated by medical professionals and we received positive feedback on the quality of upscaling. It was also used to train a tumor detection model along with three other datasets comprising the normal dataset, a low resolution set and an upscaled dataset made from the low resolution dataset. The upscaled dataset showed significant improvements as compared to the model trained on the normal dataset with 99.89% training accuracy and 93.6% validation accuracy for the upscaled dataset as compared to 97.68% training accuracy and 84.94% validation accuracy for the normal dataset. The PSNR values of the upscaled images also showed significant improvement with the Meningioma Tumor section having a mean of 29.60, the Glioma section having 31.39, the Pituitary section having 31.34 and the no Tumor section having 29.57. Keywords: Sequential Siamese GANs, Medical Imagery, Super Resolution, Tumor Detection.

1. Introduction

In this work, a Siamese GAN paired with a sequential layer in the generator will be used. The Siamese network significantly reduces the labeling cost and increases the scalability of the method to get higher resolution images. After that, these images will be used to train the object detection model as well as make a dataset for future tumor detection models. Further, the output images are analyzed by passing them through an object detection model trained to find tumors. It is then checked against traditional enhancements as a form of comparison.

1.1 Scope

In this work, a Siamese GAN paired with a sequential layer in the generator will be used. The Siamese network significantly reduces the labeling cost and increases the scalability of the method to get higher resolution images. The sequential layer at the end of the generator ensures high quality from the first epoch, thereby giving more growth possibility. Embedding weak binary pairwise label information by a Siamese network without the need of true labels, significantly reduces the labeling cost and increases the scalability of the method. After that, these images will be used to train the object detection model as well as make a dataset for future tumor detection models.

1.2 Aim

In this work, a Siamese GAN paired with a sequential layer in the generator is used. The Siamese network significantly reduces the labeling cost and increases the scalability of the method to get higher resolution images. The sequential layer at the end of the generator ensures high quality from the first epoch thereby giving more growth possibility. Embedding weak binary pairwise label information by a Siamese network without the need of true labels, significantly reduces the labeling cost and increases the scalability of the method. After that, these images will be used to train the object detection model as well as make a dataset for future tumor detection models. We aim to use the proposed new architecture to increase resolution with greater accuracy. Further, the output images are analyzed by passing them through an object detection model trained to find tumours. It is then checked against traditional enhancements as a form of comparison. Professional medical opinions would also be taken on the output of the GAN and it's possible usage in real world practices.

2. Literature Study

2.1. Existing Work

Image super-resolution was initially done using crude algorithms such as Optical or Diffractive superresolution, Geometrical or image-processing super-resolution and Aliasing. It was initially used in radar and sonar imaging applications. Simply put, SR is an algorithm that aims to provide details finer than the sampling grid of a given imaging device by increasing the number of pixels per unit area in an image Protter et al. (2009). MJOLSNESS (1986) was the first paper to use neural networks to improve resolution. Inspired by many super-resolution studies using GAN, a generative adversarial network, for image superresolution(SR) (SRGAN) Ledig et al. (2017) applied a realistic super-resolution process in the quality of photo-shooting to each image. Perceptual loss and content loss functions were used together instead of pixel-based similarity. SRGAN also benefited from the deep residual network and achieved higher mean opinion score (MOS) than the state-of-the-art techniques in the literature. The next application, by Johnson et al. Johnson et al. (2016), applied magnification in style transfer. The proposed system consists of two parts; an image transformation network and a loss network which is a convolutional neural network. Next, Perceptual GANs Li et al. (2017a) were developed to solve the issue of detecting the small objects in traffic signals. This was necessary to avoid silhouettes creeping in and leading to bad detections. In order to overcome the limitations of pixel-based loss methods, loss functions have been designed for both the generative and the discriminative network. SRPGAN Wu et al. (2017) implemented this methodology having a perceptual GAN.In Computer Vision, GANs showed remarkable data augmentation performance, such as SimGANs Jiang et al. (2021) 21% performance improvement in eye-gaze estimation.

2.2. Research Gap Identification

While all these applications of GANs in image resolution prove to be highly important, it is also needed to consider specific research that looks at the implementation of these GANs for increasing the resolution of medical imagery. A recently published paper Shende et al. (2019) showed various types of GAN networks and their applications in reconstructing MRI images with improved edges and textures for lower resolution images. Another study done by Gupta et al. (2020a) showed the ability of normal GANs to obtain resolution matching the original images after pixelation. Yamashita and Markov (2020) enhanced images of the optic nerve heads obtained by Optical Coherence Tomography. This paper used an Enhanced Super Resolution GAN(ESRGAN) which was able to improve the images based on two criterion; peak signal to noise ratio(PSNR) and the structural similarity index(SSIM). The issue however, was the quality of image was increased using the GAN although the metrics defined for improvement were not satisfied.

Title	Technique	Drawback	Remarks
Learning More with Less: GAN-based Medical Image Augmentation Han et al. (2019b) (2019)	Used traditional GAN architecture to increase resolution	Low output accuracy, higher number of untrue tumor detections	Binary pairwise label information (siamese) can be used to increase accuracy.
SiGAN: Siamese Generative Adversarial Network for Identity-Preserving Face Hallucination Hsu et al. (2019) (2019)	Siamese GAN for increasing resolution of images of human faces	Initial epochs output only random data as GAN outputs noise	Need to optimise speed of training and hence increase accuracy after same training time
Super-Resolution using GANs for Medical Imaging Gupta et al. (2020b) (2020)	GAN-CPCE and GAN-CIRCLE architecture to detect with tumors	Low accuracy and true label loss	CPCE and CIRCLE can be imbibed into other architectures with better accuracy
Medical Image Synthetic Data Augmentation Using GAN Zhang et al. (2020) (2020)	PG-ACGAN; a PG-GAN and ACGAN	Lack of modal data	PG-ACGAN gives high accuracy

3. Problem Statement and Objectives

3.1 Problem Statement

The literature study shows that the gap that has to be bridged is increasing resolution while retaining high true accuracy. True accuracy, as opposed to relative accuracy for model creation, can be defined as the actual presence of tumors. Two main approaches are utilized to do this; first, the data collection is translated into a usable form using binary paired labels. This enables us to employ a Siamese network generator rather than a conventional generator. The rationale for this is to instantly verify inaccuracy in created pixels using a comparator image (the second image passed through).

3.2 Objectives

- In this work, a Siamese GAN paired with a sequential layer in the generator is used. The Siamese network significantly reduces the labeling cost and increases the scalability of the method to get higher resolution images.

- We aim to use the proposed new architecture to increase resolution with greater accuracy. Further, the output images are analyzed by passing them through an object detection model trained to find tumors.

4. Design

4.1. Introduction

As can be seen from the literature survey, the gap that needs to be filled is to increase resolution while maintaining high true accuracy. True accuracy can be defined as the real presence of tumours and not a relative accuracy for model building. To do these, two basic methodologies are used; first, the data set is converted into a usable form such that it has binary pairwise labels. This allows us to use a Siamese network generator instead of a traditional generator. The reason to do this is to check error in generated pixels immediately with a comparator image(the second image passed through). Binary pairwise labels also significantly reduces the labeling cost.

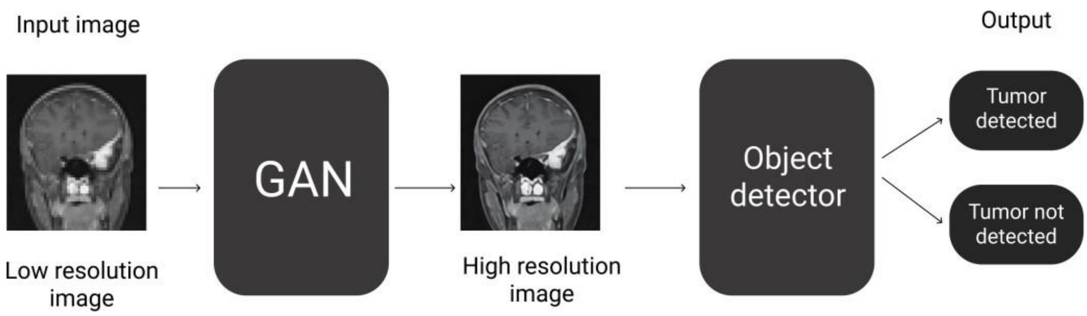
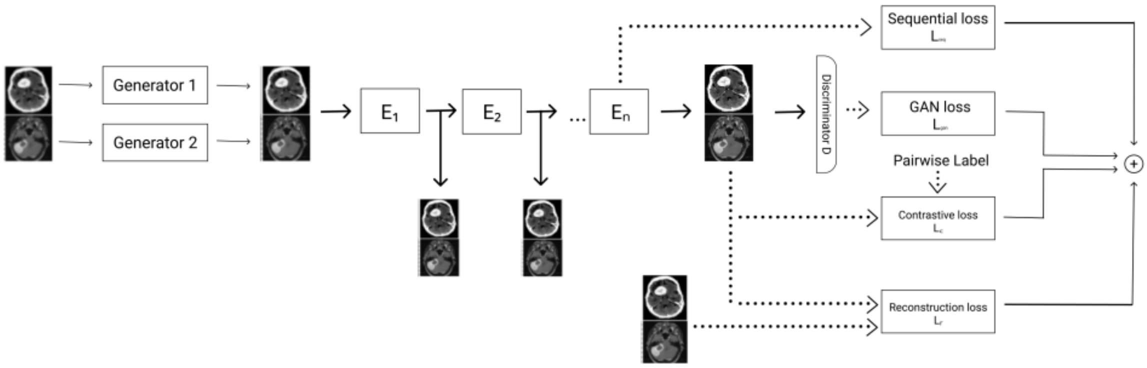


Figure 1: Workflow; Image goes through the SeSiGAN and then through an Object Detection Model

Figure 1 shows the overall workflow of our paper. We send in a low-resolution scan through our GAN. The high resolution image is then sent through a SOTA object detector which predicts the presence of a tumor

5. Implementation

We have two Siamese Generators which output into a sequential model which upscales the image. The output of this is sent to the Discriminator, who's prediction is added to the other three loss functions to get an overall loss which back propagates through the model



The basic SeSiGAN Architecture

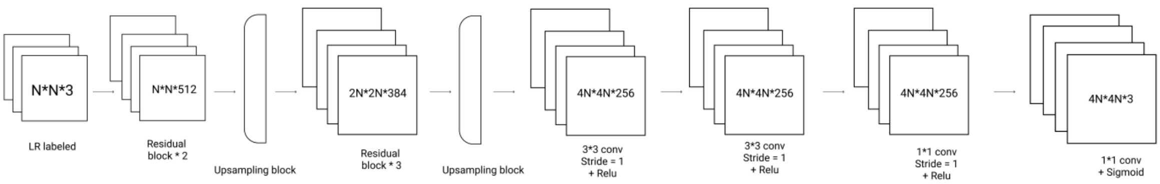
5.1 High level design

5.1.1 Dataset

Kaggle Dataset Bhuvaji et al. (2020) This dataset goes a step further and classifies the tumours under 4 different categories; no tumour, benign tumour, pituitary tumour and malignant tumour. It has 3264 images of which 394 are testing and 2870 are training.

5.1.2 Generator

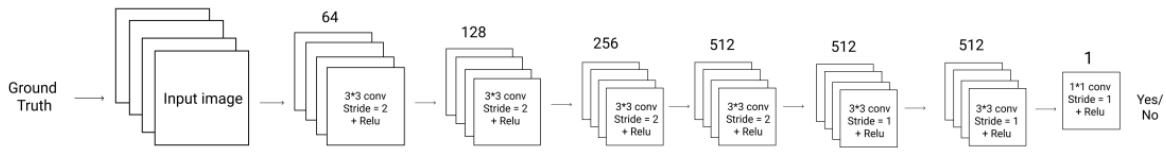
The generator consists of 2 residual blocks and upsampling blocks. Next there are 3 convolutional layers followed by a sigmoid. This outputs the generator output for an epoch. Specifically, we send in 2 images to upscale by a factor of 4



Generator Network Model

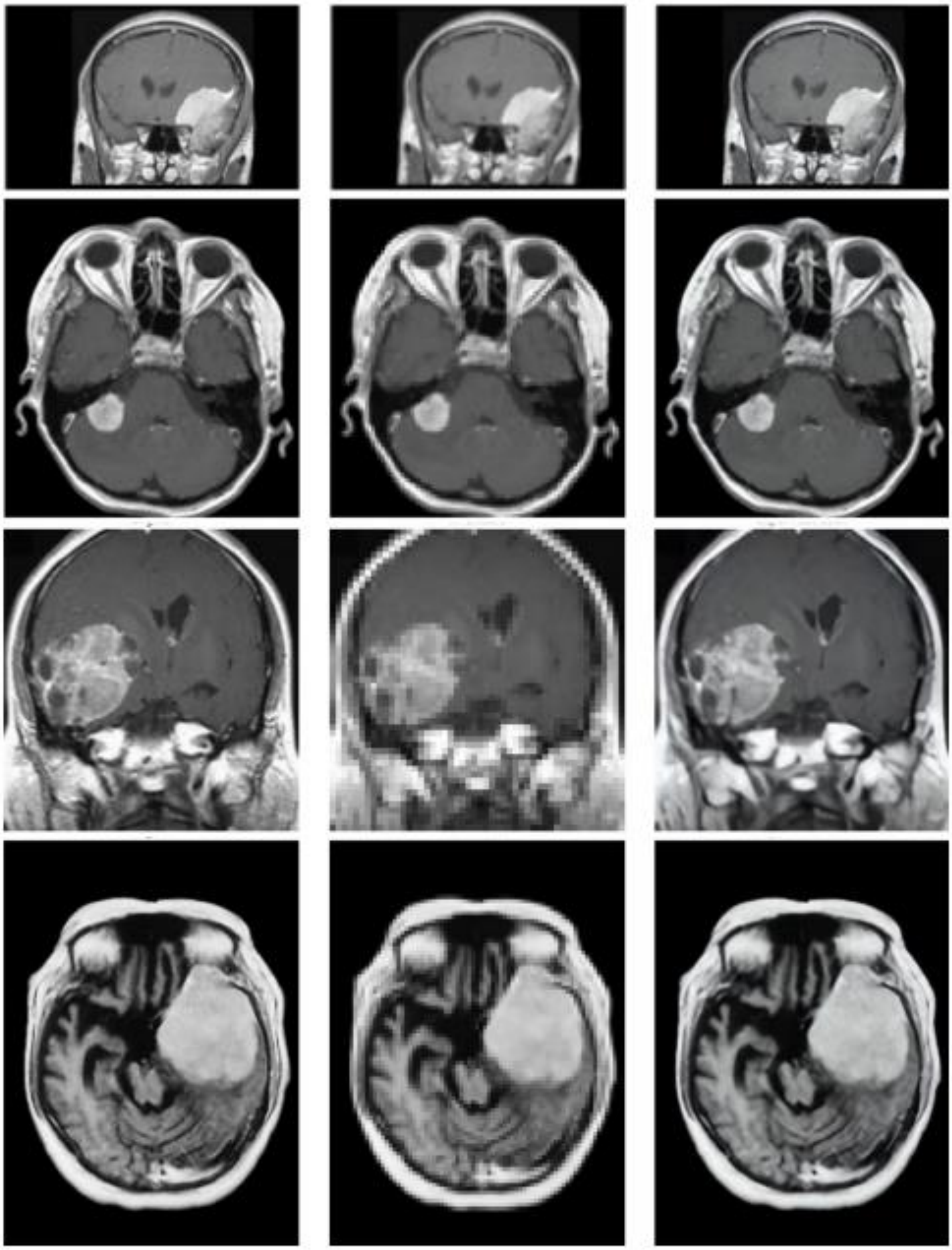
5.1.3 Generator

The discriminator model is a fully connected convolutional network as shown in Figure 6. It has 7 convolutional layers followed by a polling layer. The first four layers have a stride of 2 and the last three have a stride of 1. They all use the Relu activation. It is similar in structure to the DCGAN Malhotra et al. (2021).

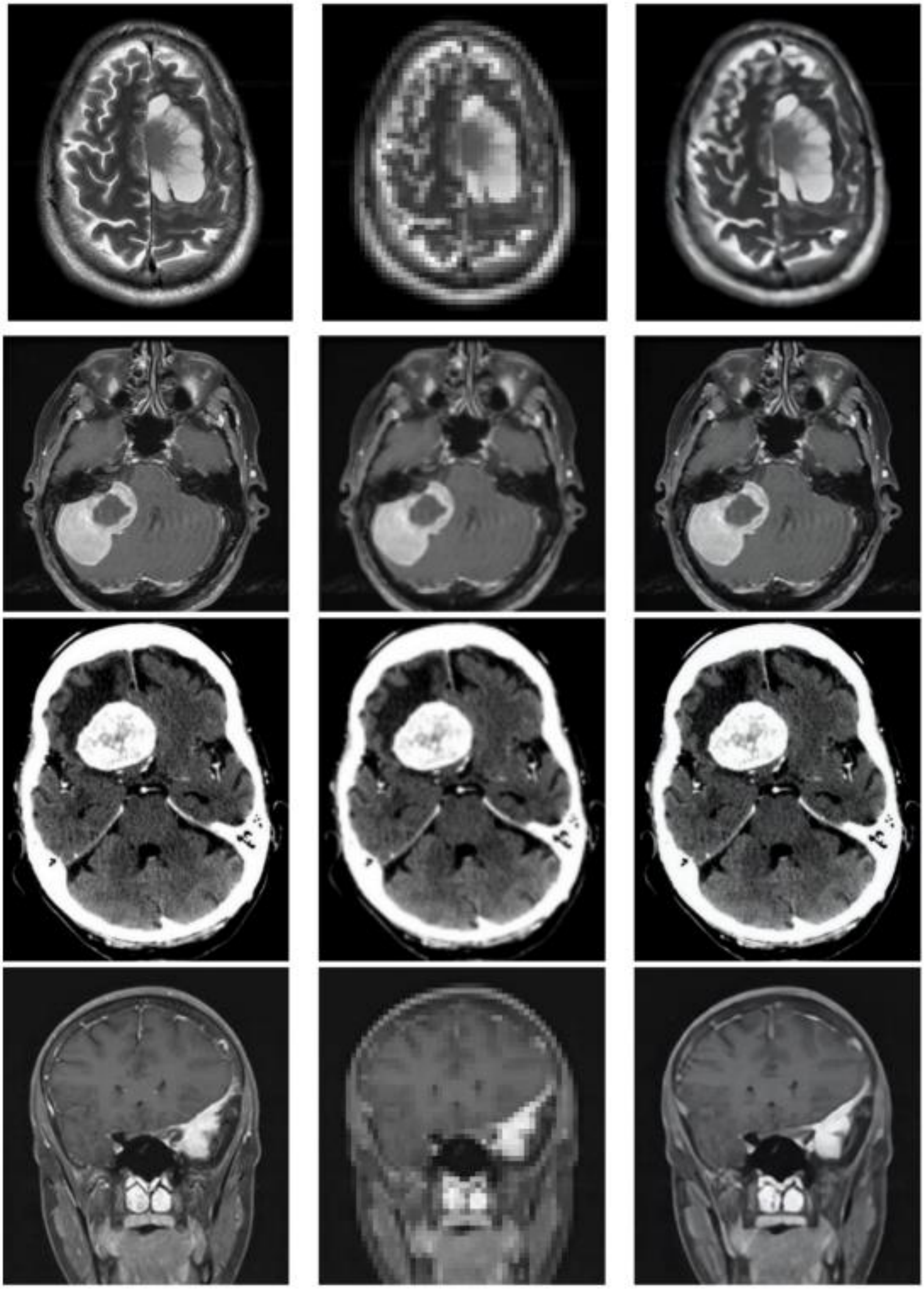


Discriminator Network Model

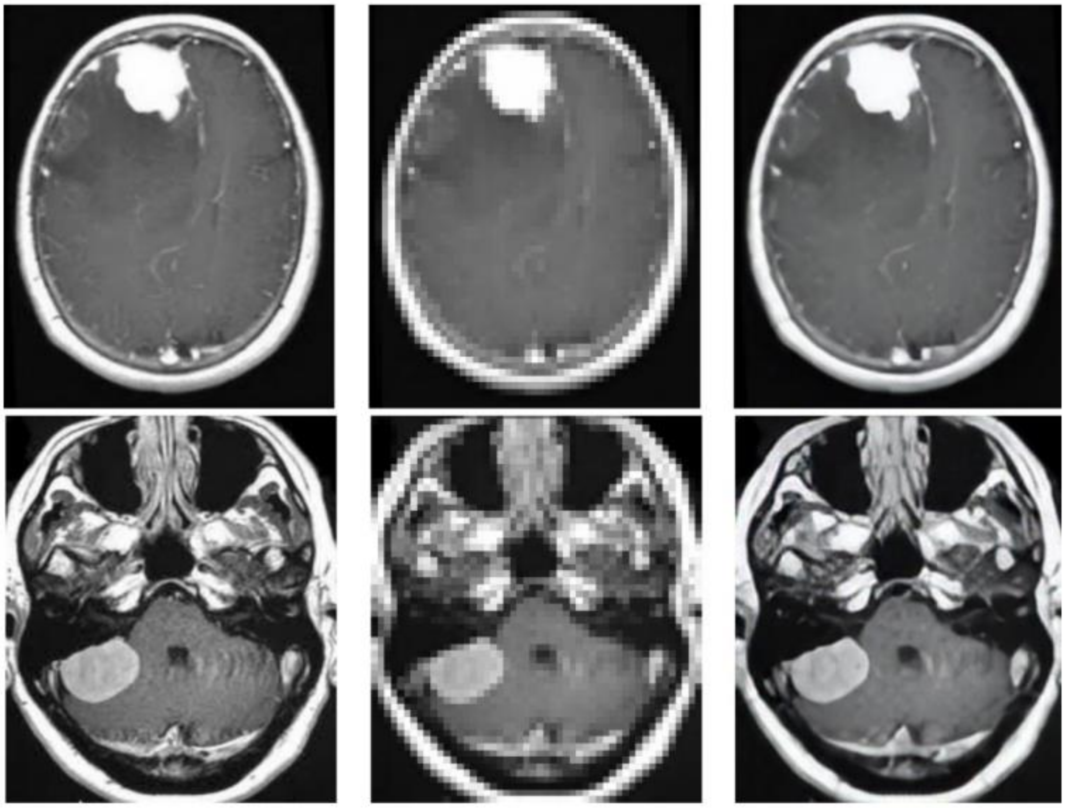
6.1 Demo Screenshots



Comparison of Images after Super resolution



Comparison of Images after Super resolution

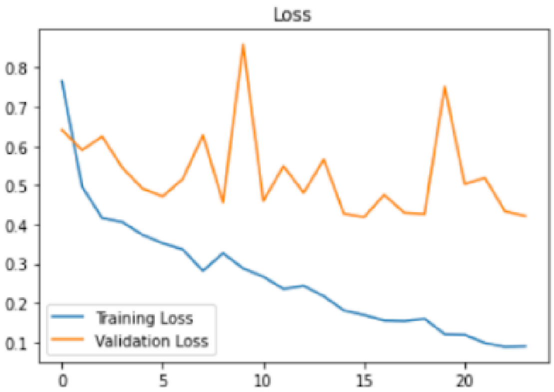


Comparison of Images after Super resolution

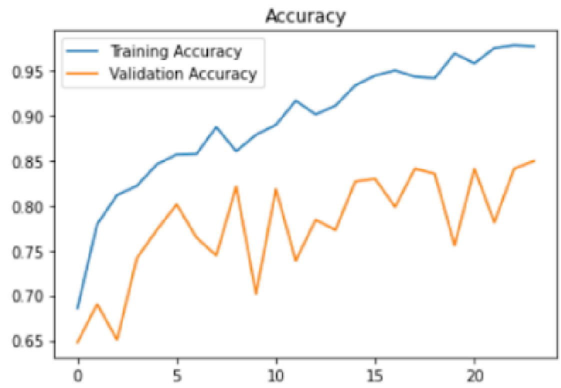
6.2 Graphs

6.2.1 Low resolution Dataset

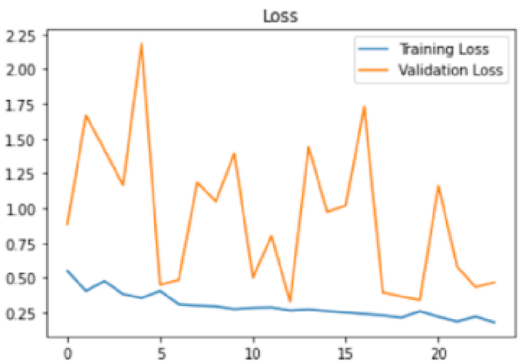
The training dataset contained 70% of the images in the dataset while the testing dataset contained 30% of the dataset. The model when trained with the low resolution dataset for 24 epochs each having a batch size 32,running 51 iterations for each instance, produced an accuracy of 93.05%. The model had a loss of 0.1818, a validation loss of 0.4346 and a validation accuracy of 0.8121. As Figure 14 shows, the loss for the training dataset starts high and the loss for the validation dataset starts slightly lower than the training dataset. Finally, with the 24th epoch the loss of the training dataset is 0.1891 and the loss of the validation dataset is 0.8116.



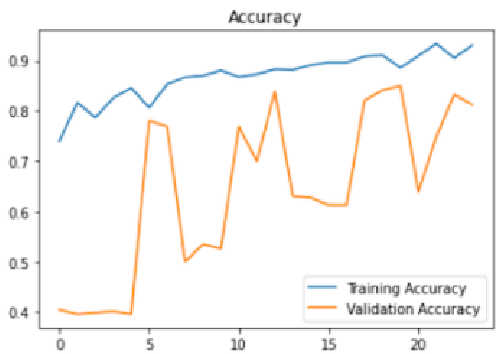
(a) Loss: Normal Dataset



(b) Accuracy: Normal Dataset



(a) Loss: Low resolution Dataset



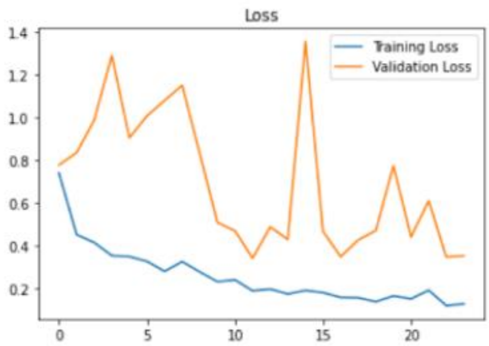
(b) Accuracy: Low resolution Dataset

When tested, it had a testing accuracy of 0.8145 and a testing loss of 0.4341.

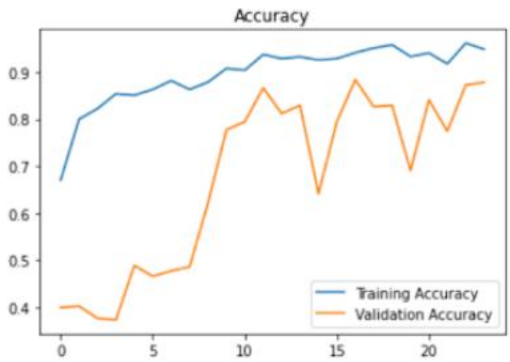
6.2.2 Super Resolution from Low Resolution Dataset

The training dataset contained 70% of the images in the dataset while the testing dataset contained 30% of the dataset. The model when trained with the super resolution from low resolution dataset for 24 epochs each having a batch size of 32 running with 51 iterations, produced an accuracy of 94.91%. The model had a loss of 0.1277, a validation loss of

0.3524 and a validation accuracy of 0.8786. Finally, with the 24th epoch the loss of the training dataset is 0.032 and the loss of the validation dataset is 0.2237.



(a) Loss: Super Resolution from Low Resolution Dataset

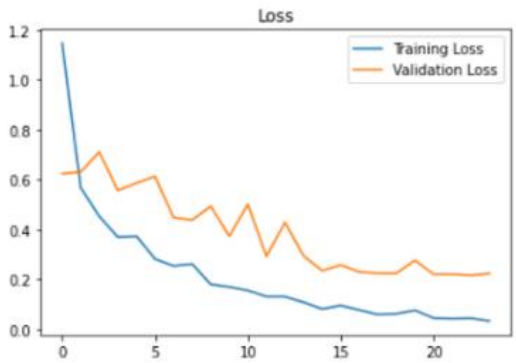


(b) Accuracy: Super Resolution from Low Resolution Dataset

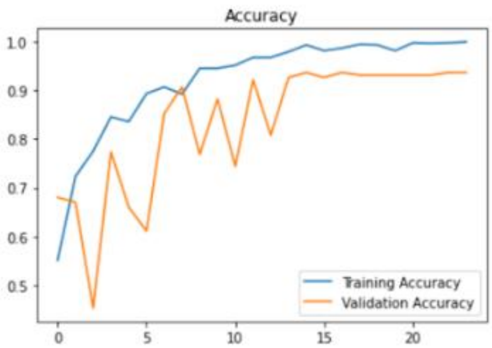
When tested, it had a testing accuracy of 0.8928 and a testing loss of 0.2878.

6.2.3 Super Resolution from Normal Dataset

The training dataset contained 70% of the images in the dataset while the testing dataset contained 30% of the dataset. The model when trained with the super resolution dataset for 24 epochs each having a batch size of 32 and with 51 iterations, produced an accuracy of 99.89%. The model had a loss of 0.0332, a validation loss of 0.2237 and a validation accuracy of 0.9360. Finally, with the 24th epoch the loss of the training dataset is 0.1277 and the loss of the validation dataset is 0.3524.



(a) Loss: Super Resolution from Normal Dataset



(b) Accuracy: Super Resolution from Normal Dataset

When tested, it had a testing accuracy of 0.9455 and a testing loss of 0.2303.

7. Conclusion

We proposed a novel Sequential Siamese GAN architecture to obtain super resolution medical images. The proposed network encompassed current standard architectures while fixing gaps left to achieve true 22 accuracy. The GAN was tested in terms of two metrics; visual quality and objective quality. The survey results from varied medical professionals add value to the visible improvements in visual quality. We showcase the objective quality of the output images in terms of PSNR and by training a tumor detection model on our dataset. The results of the detection are compared with 3 other datasets which are used to train the same model; the original dataset, the low resolution dataset and the super resolution dataset made from the low resolution dataset. The outcome of the comparison was that the super resolution dataset yielded the highest accuracy of 99.89%. The original dataset had comparable loss and accuracy as the super resolution from the low resolution dataset; 0.09 and 0.13, 97.68% and 94.91%. The scale up was hence successful under an objective metric.

7.1. Limitations

The real world loss still remains at a 0.9 and 1.3 which poses some level of issues in diagnosis. This can be solved by taking several other measures to determine the existence of the tumor. Second, the SeSiGAN needs a significant amount of data to train to be efficient even though it trains quicker than a traditional GAN model.

7.2. Future Work

A focus on improving tumor detection models can help tackle the other half of the issue in diagnosis. Also, further scalability can be made given Super resolution images of tumor only scans are developed with better hardware. This can help improve SeSiGANs accuracy and benefits

Bibliography

[1] Changhee, H., Kohei, M., Shin'ichi, S., & Hideki, N. (2019). Learning more with less: GAN-based medical image augmentation. *Med. Imaging Technol, 6*.

[2] Hsu, C. C., Lin, C. W., Su, W. T., & Cheung, G. (2019). Sigan: Siamese generative adversarial network for identity-preserving face hallucination. *IEEE Transactions on Image Processing*, *28*(12), 6225-6236.

[3] Gupta, R., Sharma, A., & Kumar, A. (2020). Super-resolution using gans for medical imaging. *Procedia Computer Science, 173*, 28-35.

[4] Zhang, H., Huang, Z., & Lv, Z. (2020, October). Medical image synthetic data augmentation using gan. In *Proceedings of the 4th International Conference on Computer Science and Application Engineering* (pp. 1-6).