Research Proposal-1.docx

by Amitkumar Maheshwari

Handling class imbalance by GAN based Data Augmentation in Medical Images

Amitkumar M Maheshwari

Research Proposal

Master of Science in Machine Learning and Artificial Intelligence

April 2022

Abstract

Deep learning based models have proven their strength in medical fields, especially working with medical images. In recent times, many open source platforms collaborated with medical institutes and experts had attempted to address the fundamental obstacle of the lack of reliable training datasets by making the data available to the community with proper annotation. However, this attempt doesn't solve the other significant problem which is the lack of particular class(es) in the available training dataset. It is generally observed in medical images that some anomaly/abnormality/condition would occur very rarely in comparison with other cases. Such class imbalance impacts the performance of the models by leading the output to be biased towards the dominating class(es). The class imbalance issue isn't hidden from the research community and there has been fair enough research has been done to address the lack of training image by synthetically augmenting. Although in many cases of radiographic image datasets, successful image augmentation has been presented still in the case of camera-based or natural medical images that contain a high degree of variance in visual appearance and colors, the performance of synthetical augmentation is still not satisfactory. This research is aimed to further improve image augmentation for camera-based medical images by using GAN-based image synthesis. This research will utilize skin lesion dermoscopic images to train and validate image augmentation carried out using GAN variants like DC-GAN and Style-GAN. The augmented dataset will be independently evaluated as well as the classification models trained on the dataset.

Table of content

3

List of Figures

List of Tables

List of Abbreviations

GAN	Generative Adversarial Nets
AC GAN	Auxiliary GAN
DC GAN	Deep Convolutional GAN
PG GAN	Progressive GAN
TMP GAN	Texture-constrained Multichannel Progressive GAN
CNN	Convolutional Neural Network
VGG NET	Visual Geometry Group Net
YOLO	You Only Look Once
ISIC	International Skin Imaging Collaboration
BraTS	Brain Tumor Segmentation
CBIS	Curated Breast Imaging Subset
DDSM	Digital Database for Screening Mammography
CT	Computed Tomography
MRI	Magnetic resonance imaging
VAE	variational autoencoders

1. Background

Machine learning, especially deep learning based models and AI is continuously making their prominent place in modern-day medical science. From routine checks, to assisting in complex surgical operations AI solutions have been established as digital assistance to doctors and other medical staff. However, for better-performing models, a better training dataset is needed. An ideal training dataset should have sufficient and diverse enough training data. But in the medical domain, there are often cases of unavailability of training data, or even if the data is available, the number of positive cases of rare anomalies is very less in comparison with the number of negative cases which results in either overfitted or extremely biased detection/classification model. Often misclassification of any medical condition can be as bad as fatal, so it is important to develop an unbiased and reliable classification model. Additionally, medical experts are required to get the training data reviewed to label them. This process is manual, time-consuming, and cost inefficient. On top of that, it is highly dependent on the expertise of the medical professional and prone to human error.

In this research, ISIC 2020 skin lesion images (International Skin Imaging Collaboration. SIIM-ISIC 2020 Challenge Dataset., 2020) are used to demonstrate the issue of class imbalance. (Table 1.1: Number of images per class in the ISIC 2020 dataset and Figure 1.1: Class distribution in ISIC 2020 dataset show) the distribution of different cases of skin lesions.

Diagnosis	Count of diagnosis
atypical melanocytic proliferation	1
cafe-au-lait macule	1
lentigo NOS	44
lichenoid keratosis	37
melanoma	584
nevus	5193
seborrheic keratosis	135
solar lentigo	7
unknown	27124
Total images	33126

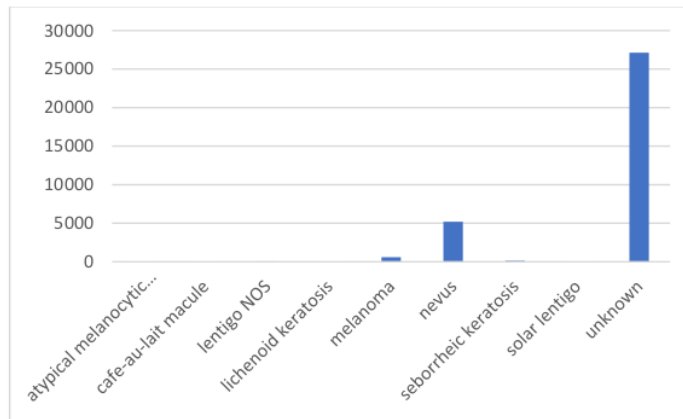Table 1.1: Number of images per class in the ISIC 2020 dataset

Figure 1.1: Class distribution in ISIC 2020 dataset

Class 'unknown' and class 'nevus' are highly dominating the entire distribution and it is obvious if this dataset is used to train the skin lesion classification model as is, the resultant model will be biased towards these two classes. The condition becomes too dangerous given the fact that 'melanoma' type skin lesion is critical to be detected especially when dermoscopy is the only reliable source of traditional detection as naked eye examination is proven to be less accurate (M E Vestergaard et al., 2008).

Two general approaches are there to handle class imbalance, under sampling and over sampling. Oversampling, the process of increasing the training data using data augmentation techniques (or just duplicating the data) is a more appropriate approach as just like the most cases of medical images, under-sampling of the two dominant classes to balance class distribution can't be the possible approach as it is observed in the Table , availability of the images in other classes are extremely less and an attempt to under-sample the dataset will result in underfitted model.

A combination of two independent deep learning based networks, one responsible for image generation and the other for image classification, interacting with each other can build an innovative image generation model (Goodfellow et al., n.d.). In their research, they proposed two deep learning models being trained parallelly, a Generative model G which learns the data distribution to produce the image as output and a Discriminative model D that takes the generated image as input and estimates the probability of the input image is from real training dataset rather than generated by G. Together both model can work as one unit that is capable of generating realistic synthetic images and it is known as generative adversarial nets (GAN). Figure 1.2: basic architecture of GAN shows the basic architecture of GAN.
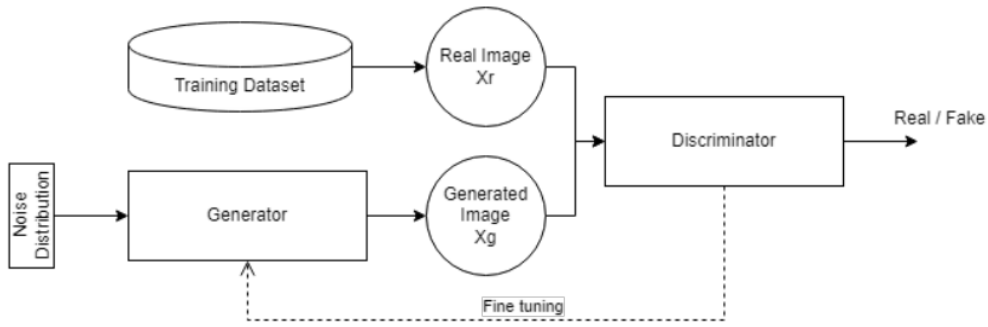
6

Figure 1.2: Basic architecture of GAN

Applications of GANs have a wide range in the computer vision field, there are many cases such as image augmentation, image registration, medical image generation, image reconstructions, and image-to-image translation where GANs are proven to be useful. Basic/Vanilla GAN has issues when working with high resolution images or more complex features like Mode collapse and gradient vanishing. Also, it performs limited on complex tasks such as image-to-image translations. Many researchers extensively worked on GAN to propose different variants of GAN to overcome the limitations of original GAN architecture like, AC-GAN to introduce the conditional operation, Progressive GAN to be able to progressively enhance the resolution of generated images, pix2pix GANs to be able to perform image to image translations and fusing segment of one image (or entire image) on other images to produce out of the box results.

2. Related Research Work

After Goodfellow and his team introduced the concept of Generative Adversarial Nets (GAN) (Goodfellow et al., n.d.), it had soon become an area of interest for many researchers working in the domain of computer vision, and deep learning, and a lot of work has been done in this field so far. Although it was introduced in 2014 a solid trend of using GAN variants to generate synthetic images to be used in other deep learning networks as input can be seen in recent years.

F-CGAN, a two-staged conditional GAN proposed in (Fu et al., 2020) works on image-to-image translation style instead of noised based image generation. F-CGAN showcased a significant improvement in generating fine-grained images when compared with previously acclaimed AC-GAN, and SNGAN and the classification models trained on the dataset generated by F-CGAN showed better accuracy than the standard model and SNGAN model. On the other side,

7

GANs (Dumagpi et al., 2020; Dumagpi and Jeong, 2021), have been put to generate synthetic images of positive threat X-ray images to balance an extremely unbalanced dataset. In (Dumagpi and Jeong, 2021) researchers have used DC GAN for image generation and Cycle-GAN for image translation in addition to traditional image transformation (shown in Figure 2.1: Traditional(left) and Generative(right) techniques of Images Augmentation). While evaluating they noticed that combining all three types of synthesized images can make the classification model generalized enough to bring significant improvement in average precision.
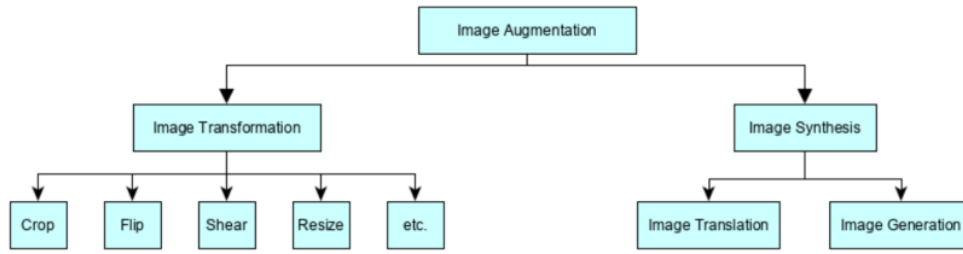

Figure 2.1: Traditional(left) and Generative(right) techniques of Images Augmentation used in (Dumagpi and Jeong, 2021)

This shows that GANs have applications from normal object classification to as critical as subway and airport security by improving the performance of the classification model. Talking about medical images, most research has been done on radiographic images like X-rays, CT, MRI, etc. while on natural or camera images we can see there was comparatively less focus.

There is a fundamental domain difference in medical images in comparison with other images be it camera images or radiographic images. Deep-learning based models like classification model or segmentation model, would, in general, look for certain types of anomalies and in many cases, such anomalies would display very delicate texture or color differences thus Image synthesis for medical images must be sensitive enough to learn such delicate distribution and produce images that contain due features properly. Where traditional GAN may not preserve all the textures of CBIS-DDSM screening images, (Guan et al., 2022) have proposed a method of GAN based image augmentation "texture-constrained multichannel progressive GAN (TMPGAN)". The objective was not to handle class imbalance but to generate synthetic images to overcome the issue of less training images available. TMP-GAN applies a progressive generation mechanism that improves image synthesis steadily. Foreground-Generation method is being used in it, which means the model will generate the synthetic lesions in selected areas

8

of normal/actual images to produce positive case images. A progressive fusing mechanism also makes sure that the synthetic lesion's continuity on the background to preserve the textures.

The other and more significant challenge in training deep learning models for medical images is the desired images are either very less to train the model on or they are extremely unbalanced as most cases would fall in normal/negative class.

A study, proposed in April and Published in May of 2020, merely a couple of months after covid was declared a worldwide pandemic and with an obvious heavy shortage of training images for positive cases, AC-GAN has been put in use for Synthesizing both Covid CXR and normal CXR images to train a classification model for covid detection (Waheed et al., 2020) . On other hand, instead of Image Translation (AC-GAN), (Srivastav et al., 2021) has achieved significant improvement in pneumonia detection by augmenting positive images using image generative GAN model – DC-GAN. However, both studies were not focusing on the "Class Imbalance" issue which is very common across the medical domain.

While most research related to data augmentation using GAN variants were focused to overcome the scarcity of the data itself, there were some researches focused on the challenge of data being extremely biased towards certain class(es) and the rest classes would rarely occur. In (Qasim et al., 2020) researchers talk about the class imbalance issue in the BraTS and ISIC datasets. To achieve the image segmentation task, unlike the traditional GAN where two components, Generator and Discriminator would compete, they introduced a SPADE based GAN with third component called "Segmentor" (Figure 2.2: Basic representation of Red-GAN.) which is fixed and pretrained on the same dataset to obtain the synthetic image segments on the fly.
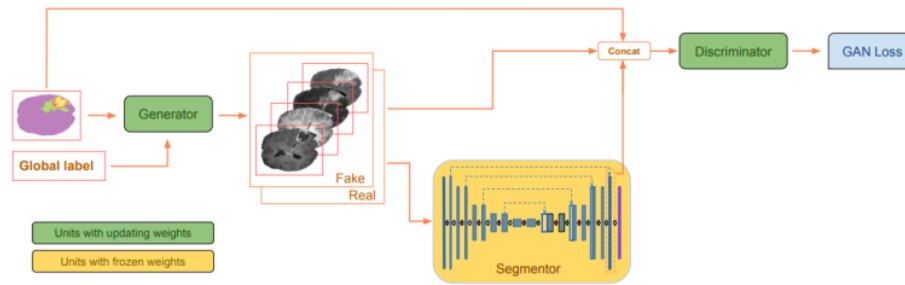

Figure 2.2: Basic representation of Red-GAN. Here we can observe the third component pre-trained "Segmentor" being introduced (Qasim et al., 2020).

9

A study (Frid-Adar et al., 2018) explored two very basic variants of GANs and those were DCGAN and ACGAN. Unlike DCGAN, ACGAN is a conditional GAN and as external conditional information, ACGAN provides class information in the GAN network. Trained on liver CT images for lesion segmentation, their study not only demonstrates the performance improvement but also compares the performance difference of classification when the model is trained on traditional image augmentation and GAN Based augmentation. Cycle-based GAN combined with YOLO (you only look once) architecture (Hammami et al., 2020) was used for generating synthetic MRI images to be used to train a multi-organ detector model. Instead of one set of generators and discriminator, Cycle GAN is made of two sets and works as bidirectional image translation. The output of the Cycle GAN is then fed into YOLO for detection.

To obtain a reliable GAN based image synthesis on skin lesion images, a study, (Bissoto et al., 2021) reviewed 18 prominent research that claimed of gaining significant improvement in the model for classification or segmentation tasks that were trained on GAN based synthetic images. Further, their study has validated how different real:synthetic image ratio leads to a different outcome. Researchers tried four different GAN variants: SPADE, pix2pixHD, PGAN, and StyleGAN to generate synthetic images and trained classification model Inception v4 with the generated training dataset using various real:synthetic image ratios. Researchers then went ahead and compared two basic techniques of utilizing the synthetic images in the classification model, Augmentation and Anonymization. However, in any terms, they could not achieve as good results as it was claimed in the referred papers.

One common trend that has been noticed in (Bissoto et al., 2021) and (Qasim et al., 2020) is that both were not able to perform well for the skin lesion dataset, while Red-GAN could perform reasonably okay for the brain tumor dataset. The concluded reason for these GANs' inability on performing better was, that "skin lesion images have a more visual appearance in comparison with brain tumor MRI images (or other radiographic images), thus image segmentation and mask to image mapping become more difficult in comparison with MRI images". And this opens a large gap for GAN based image synthesis for camera images and the reason given above, it should not be limited to skin lesion images but other medical images like surgical images or endoscopic images as well.

Other than radiographic images, studies had been carried out on rich in color and texture microscopic images of human protein where DC-GAN has been applied (Verma et al., 2020)

and on dermoscopy skin images (Litjens et al., 2017; Rashid et al., 2019; Qin et al., 2020; Bissoto et al., 2021) where a different variant of GANs has been used for image augmentation. However, none of them focused on handling class imbalance, and only (Bissoto et al., 2021) tried and failed to improve the ultimate classification model. Although modified Style-GAN has provided promising results for skin lesion image generation (Qin et al., 2020)

3. Research Questions

On the bases of reviewing the prominent works of literature so far, the below questions are formulated that the current research will ultimately explore.

- Does class imbalance present in the dataset affect the outcome of the classification of skin lesion images?
- Does GAN based data augmentation help in creating a synthetic dataset for camera/dermoscopic skin lesion images that can improve classification performance?
- Does the skin lesion dataset generated by GAN based data augmentation outperform the dataset generated by traditional image augmentation techniques?
- For the classification of skin lesion images, does the model train on data augmentation perform better than the model train on data anonymization?
- Does "image translation" based GAN perform better than "image generative" GAN?

4. Aim and Objectives

The main aim of this research is to develop a stable GAN model that can generate reliable synthetic medical images. The skin lesion dataset is highly imbalancedd and biased, the goal is to be able to generate synthetic images for a specific class(es) to handle the class imbalance present in the dataset that ultimately results in better trained and reliable classification models.

To achieve the aim following objectives are formulated:

- To load and analyze the dataset to identify and eliminate any error/impurity in the dataset
- To perform the image preprocessing to normalize the images and bring them to a uniform size

11

- To generate GAN models using different techniques to identify the most suitable GAN based on the nature of the given dataset
- To generate classification models being trained on the augmented dataset.
- To evaluate the performance of GAN and classification models

5. Significance of Study

This research is contributing to the synthetic medical camera image generation by using different variants of GAN models to handle the 'class imbalance' problem in dataset and scarcity of training images which leads to poor performance of classification models. Dermoscopic skin lesion images are selected to be used in this research as in this dataset, images are camera-based images and demonstrate extreme class imbalance. Among all types of skin cancers, 'melanoma' is the most lethal one thus it becomes very critical for medical science to have a stable and reliable melanoma detection mechanism as early diagnosis can greatly improve the survival rate of patients.

'melanoma' is one of the classes of skin lesions in the dataset which is being shadowed by the dominating class 'melanocytic nevus (nv)' the classification models benign trained on such biased datasets mostly perform poorly in melanoma detection. This research is aimed to overcome this issue by oversampling the minority class (here 'melanoma') with synthetic images of the melanoma class generated by using GAN.

In addition, a generic GAN model will not only help in balancing the skin lesion images but can also be utilized in generating other camera based medical images like surgical images of rare conditions or endoscopic images of anomalies found. This research will also open gates for further extended research to develop GANs that can be used domain agnostically.

6. Scope of the Study

To keep the research focused and feasible to be completed in given time duration, the scope of the research work has been limited as below:

- This research will explore only two approaches to image augmentation, traditional image transformation, and GAN based image synthesis. Image synthesis using "variational autoencoders (VAEs)" is included in the research

- Only noise-based Image generative GANs will be explored and only DC-GAN and Style-GAN variants will be further implemented for image augmentation. Image translation-based GAN techniques are not included in the research and so does the image segmentation.

- The classification models are only meant to evaluate the dataset balanced by image augmentation techniques and further improvements of the classification models are not in scope.

- Using reinforcement learning to improve the quality and speed of GAN models by introducing periodic feedback mechanisms in GAN architecture is not included in the scope of current research.

7. Research Methodology

In this research, the primary focus is on developing a GAN model that can perform well on colored and textured medical camera images like dermoscopic skin lesion images rather than focusing more on the image classification model. The whole research is divided into three main parts: Image Generation, Image Classification, and Evaluation.

The detailed flow of the entire research has been discussed in this section. The flowchart in the Figure 7.1: Methodology flow, shows the sequential order of different steps being performed to complete the objectives and achieve the main aim.
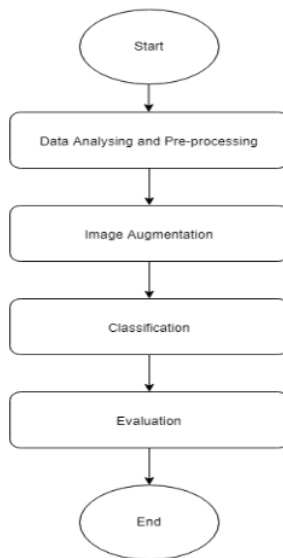
Figure 7.1: Methodology flow

13

7.1 Dataset analysis and pre-processing

A GAN model to be able to generate medical camera images, (International Skin Imaging Collaboration. SIIM-ISIC 2020 Challenge Dataset., 2020) is being used.

ISIC 2020 dataset contains:

1. 33,126 JPEG and DICOM images
2. Metadata containing information (patient ID, lesion ID, gender, age, and general anatomic site) for all 33,126 images
3. Duplicate images list
4. Ground truth of all 33,125 images

ISIC 2020 dataset is well organized and clean. However, a few basic steps will be performed as data pre-processing

1. EDA on the metadata of the images and ground truth information
2. Dropping the images
   a. Which were associated with dropped entries of metadata.
   b. Keeping the class ratio constant, reducing the dataset size to make further development feasible yet realistic.
3. Resizing the images to a uniform size
4. Normalizing the image pixel intensity values between (0,1)

7.2 Image Augmentation
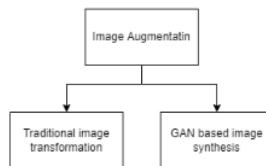
Figure 7.2: Image augmentation techniques

On the Assumption that present class imbalance in ISIC dataset will impact the classification model trained on this dataset and will be highly biased towards majority classes, image augmentation becomes critically important and thus it is the primary focus of this research. (Bissoto et al., 2021; Guan et al., 2022) extensively talks about different generative data augmentation techniques that include both image-to-image translation and noise-based image generation. However, fundamentally speaking two main ways of augmenting the images

14

(shown in Figure 7.2: Image augmentation techniques) will be explored in this research, Traditional image transformation and GAN based image synthesis (Dumagpi and Jeong, 2021)

7.2.1   Traditional image transformation

Although less sophisticated, image transformation techniques like rotating, zooming, cropping, etc. have been used to upsample the images for any particular class(es). And in many studies (Verma et al., 2020; Waheed et al., 2020; Dumagpi and Jeong, 2021) , image transformation has either been used with image synthesis or compared with image synthesis concerning the effectiveness.

Given the nature of the images and the factors responsible for classification, a few techniques of transformation like thresholding, erosion, dilation, opening, closing, etc. cannot be used to augment new images as they might alter the color, contrast, texture of the image. Whereas linear transformation techniques like resizing/scaling, cropping, zooming in/out, rotating, and flipping can be safely used.

In the context of traditional image transformation techniques, this research will be a comparative study of the effectiveness of classification models trained on the dataset that included image transformation + GAN in data augmentation, only used GAN based synthetic images for data augmentation, and standalone usage of image transformation for data augmentation.

7.2.2   GAN based image augmentation

Mainly classified into two types, image to image translation model and noise-based image generation model, many variants of GAN based models are discussed (Singh and Raza, 2020; Bissoto et al., 2021).

Inspired by studies (Qin et al., 2020; Verma et al., 2020) with comparatively similar dataset and promising outcome, this research will explore and experiments with two widely accepted GAN variants, DC-GAN and Style-GAN. DC-GAN is a relatively simpler GAN variant with both generator and discriminator comprising of the deep convolutional network. Unlike conditional GANs, DC-GAN doesn't have external conditioning as the input and output layer of the

discriminator network contains a single neuron and thus can't produce probability distribution for the generated image.

GAN can produce realistic images but being stable they cannot achieve high resolution. Style-based GANs can produce higher resolution output images where vanilla GAN might collapse. The low-resolution issue can be resolved by PGGAN too, PGGAN has limitation in effective control over the features of the image and style of the image during the image generation on the other hand Style-based GAN can generate high resolution images with good control over the features and style (Qin et al., 2020).

7.3 Image Classification

The main aim of this research is limited to generating desired GAN model to overcome the class imbalance problem and thus this research doesn't focus on improving the image classification models. These models will be used only for comparing the quality of the training dataset.

This research will use two classification models, basic CNN architecture and VGG Net using transfer learning. The image classification models will be trained on different datasets while keeping constant hyper-parameters, activation functions, and overall architecture. Once trained, these models will be evaluated on the same test dataset using the same evaluation matrices. Dataset generation is already discussed in the above sections.

7.4 Evaluation

Evaluation of this research will be a comparative study of the outcome of different models and experiments. As the development work in this research will be done in two parts, they both will be evaluated separately.

7.4.1   Evaluating GAN models

(Borji, 2019) talks briefly about the different measures to evaluate the performance of GAN models. In their study, they have proposed basic characteristics of a good GAN model evaluation measure

16

- Evaluation measures should favor the GAN model that can generate high-fidelity samples
- It should favor the GAN model that can generate a diverse sample
- It should favor the GAN model with controllable sampling
- It should favor the GAN model with well-defined bounds
- Evaluation measures should be sensitive to image distortion and image transformations.
- The evaluation measure's outcome should be in line with human perceptions.
- It should be less computational complexity.

This study will mainly be dependent on evaluating the classification model to determine whether the dataset generated using GAN is helping the classification process or not. However, there can be independent measures to evaluate GAN performance. Broadly speaking, there are two ways of evaluating the GAN, quantitive measures and qualitative measures.

Inspired by, (Qin et al., 2020) this study will evaluate the GANs based on

- Quantitive Measures
  - Inception Score – IS
    $$IS = \exp(E_x[KL(p(y|x)\|p(y))]) \text{ (Qin et al., 2020)}$$
  - Frechet Inception Distance – FID
    $$FID(r,g) = \|\mu_r - \mu_g\|_2^2 + Tr[\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{\wedge}(1/2)] \text{ (Borji, 2019)}$$
- Manually validating by visualizing the output of GAN.

7.4.2 Evaluating classification models

Given that the dataset is highly imbalanced and biased towards dominating class(es), the High Accuracy value is often misleading. For medical image classification, a high rate of false negatives cannot be accepted, on another hand high rate of false positives will require a continuous cross-checking mechanism, making the final diagnosis more time-consuming and expensive.

With this understanding, this research will evaluate the classification model using the following measures. The confusion matrix will be generated as one class is a positive case and the rest all being negative cases.

- Sensitivity (Recall, True Positive Rate): The number of positive cases that are correctly predicted out of the total positive cases

- o Sensitivity = True Positive / (True Positive + False Negative)
- o The value of sensitivity should be as high as possible
- Specificity: The number of negative cases that are correctly predicted out of the total negative cases
  - o Specificity = True Negativ / (True Negative + False Positive)
  - o The value of specificity should be as high as possible
- Precision(Positive predictive rate): Rate of correctly predicted positive case our of total positive prediction.
  - o Precision = True Positive / (True Positive + False Positive)
  - o The value of precision should be as high as possible
- ROC curve plotting will be used to visualize the performance of the classification.

However, the conclusion of the research will not be comparative but quantitive. Based on the evaluation result, this research will try to propose the answers to the questions mentioned in the 'research question' section.

8. Required Resources
Below listed software and hardware will be required to carry out the research.

8.1 Software requirement
- Operating System: Windows 10 20H2 or above
- Language: Python 3.8
- For on-prem development work
  - o Conda Package (Python packages) Manager: Anaconda Navigator 2.0
  - o Notebook/IDE: Jupyter Notebook
- For online development work
  - o Google collab
- Commonly used python packages for GAN and Classification model development, Data loading, and visualization, and for supporting development tasks.
- Microsoft Office 360 16.0.14

8.2 Hardware requirement
- Processor: Intel® Core ™ i7 – 10510U

18

o   Clock rate: 1.80 GHz

o   Cores: 4

o   Logical Cores: 8

- RAM: 16 GBs

- Storage: (to support development environment, dataset, and development) 100 GBs

9. Research Plan

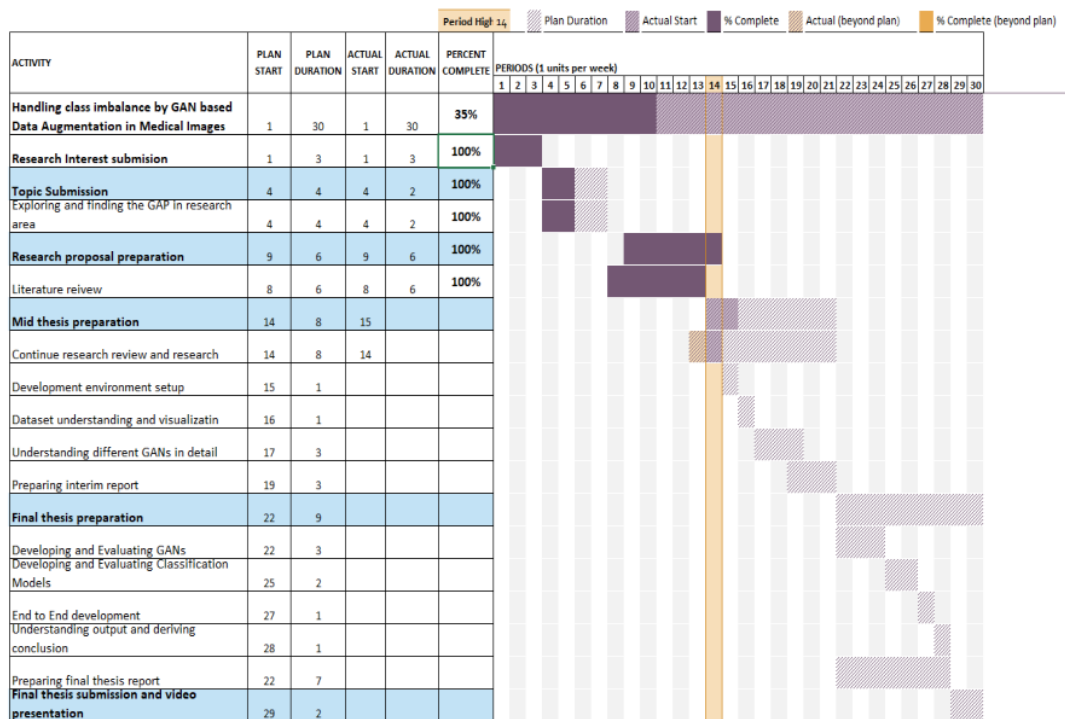Below is shown a brief research plan that this research is following.

ACTIVITY	PLAN START	PLAN DURATION	ACTUAL START	ACTUAL DURATION	PERCENT COMPLETE
Handling class imbalance by GAN based Data Augmentation in Medical Images	1	30	1	30	35%
Research Interest submision	1	3	1	3	100%
Topic Submission	4	4	4	2	100%
Exploring and finding the GAP in research area	4	4	4	2	100%
Research proposal preparation	9	6	9	6	100%
Literature reivew	8	6	8	6	100%
Mid thesis preparation	14	8	15		
Continue research review and research	14	8	14		
Development environment setup	15	1			
Dataset understanding and visualizatin	16	1			
Understanding different GANs in detail	17	3			
Preparing interim report	19	3			
Final thesis preparation	22	9			
Developing and Evaluating GANs	22	3			
Developing and Evaluating Classification Models	25	2			
End to End development	27	1			
Understanding output and deriving conclusion	28	1			
Preparing final thesis report	22	7			
Final thesis submission and video presentation	29	2			

Figure 8.1: A brief research plan

10. Risk and contingency plan

Followings are the potential risk factors that can affect the timeline and outcome of the research work. To maintain the resiliency in the research work, against each risk factor suitable mitigation has been planned.

- Risk: Hardware/Software issues
  - All the documents, references, and development work including the dataset are maintained on and in sync with cloud storage.
  - Windows operating system's image capturing the development environment has been taken as a backup
  - Google collab (or any other online development platform) can be utilized to continue development work in case of on-prem development environment is not available.
- Risk: Time constraint
  - The scope of the study has been planned according to the available time
  - However, scop is designed in a way that some buffer time should be available for additional experiments (e.g., working on classification model improvement). In case of critical time constraints, these extra experiments can be dropped to achieve the main goal by completing all the objective in proper manner.
- Risk: The research is not generating the expected outcome
  - Instead of waiting for the final outcome, be continuously in contact with the thesis supervisor and discuss the periodic progress and outcome and seek his guidance and if needed university professor's guidance and proceed accordingly.

References

Anon (2020) *International Skin Imaging Collaboration. SIIM-ISIC 2020 Challenge Dataset. Creative Commons Attribution-Non Commercial 4.0 International License*.

Bissoto, A., Valle, E. and Avila, S., (2021) GAN-Based Data Augmentation and Anonymization for Skin-Lesion Analysis: A Critical Review. [online] Available at: http://arxiv.org/abs/2104.10603.

Borji, A., (2019) Pros and cons of GAN evaluation measures. *Computer Vision and Image Understanding*, 179, pp.41–65.

Dumagpi, J.K. and Jeong, Y.J., (2021) Evaluating gan-based image augmentation for threat detection in large-scale xray security images. *Applied Sciences (Switzerland)*, 111, pp.1–21.

Dumagpi, J.K., Jung, W.Y. and Jeong, Y.J., (2020) A new GAN-based anomaly detection (GBAD) approach for multi-threat object classification on large-scale x-ray security images. *IEICE Transactions on Information and Systems*, E103D2, pp.454–458.

Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J. and Greenspan, H., (2018) GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*, 321, pp.321–331.

Fu, Y., Li, X. and Ye, Y., (2020) A multi-task learning model with adversarial data augmentation for classification of fine-grained images. *Neurocomputing*, 377, pp.122–129.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., (n.d.) *Generative Adversarial Nets*. [online] Available at: http://www.github.com/goodfeli/adversarial.

Guan, Q., Chen, Y., Wei, Z., Heidari, A.A., Hu, H., Yang, X.H., Zheng, J., Zhou, Q., Chen, H. and Chen, F., (2022) Medical image augmentation for lesion detection using a texture-constrained multichannel progressive GAN. *Computers in Biology and Medicine*, 145.

Hammami, M., Friboulet, D. and Kechichian, R., (2020) *CYCLE GAN-BASED DATA AUGMENTATION FOR MULTI-ORGAN DETECTION IN CT IMAGES VIA YOLO. 2020 IEEE International Conference on Image Processing (ICIP)*.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., van Ginneken, B. and Sánchez, C.I., (2017) *A survey on deep learning in medical image analysis. Medical Image Analysis*, .

M E Vestergaard, S W Menzies, P Macaskill and P E Holt, (2008) Dermoscopy compared with naked eye examination for the diagnosis of primary melanoma: a meta-analysis of studies performed in a clinical setting.

Qasim, A.B., Ezhov, I., Shit, S., Schoppe, O., Paetzold, J.C., Sekuboyina, A., Kofler, F., Lipkova, J., Li, H. and Menze, B., (2020) Red-GAN: Attacking class imbalance via conditioned generation. Yet another perspective on medical image synthesis for skin lesion dermoscopy and brain tumor MRI. [online] Available at: http://arxiv.org/abs/2004.10734.

Qin, Z., Liu, Z., Zhu, P. and Xue, Y., (2020) A GAN-based image synthesis method for skin lesion classification. *Computer Methods and Programs in Biomedicine*, 195.

Rashid, H., Tanveer, M.A. and Aqeel Khan, H., (2019) Skin Lesion Classification Using GAN based Data Augmentation. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, 2019, pp.916–919.

Singh, N.K. and Raza, K., (2020) *Medical Image Generation using Generative Adversarial Networks*.

Srivastav, D., Bajpai, A. and Srivastava, P., (2021) Improved classification for pneumonia detection using transfer learning with GAN based synthetic image augmentation. In: *Proceedings of the Confluence 2021: 11th International Conference on Cloud Computing, Data Science and Engineering*. Institute of Electrical and Electronics Engineers Inc., pp.433–437.

Verma, R., Mehrotra, R., Rane, C., Tiwari, R. and Agariya, A.K., (2020) Synthetic image augmentation with generative adversarial network for enhanced performance in protein classification. *Biomedical Engineering Letters*, 103, pp.443–452.

Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F. and Pinheiro, P.R., (2020) CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection. *IEEE Access*, 8, pp.91916–91923.

Research Proposal-1.docx