National College of Ireland

Analysis of Microscopic Blood Images in Sickle Cell Classification Using Deep Learning Algorithm

MSc Research Project
MSc in Data Analytics

Ifeoma Oduntan
Student ID: x18191274

School of Computing
National College of Ireland

Supervisor: Jorge Basilio

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Ifeoma Oduntan
Student ID:	X18191274
Programme:	MSc in Data Analytics **Year:** 2022
Module:	Research Project
Supervisor:	Jorge Basilio
Submission Due Date:	15/08/2022
Project Title:	Analysis of Blood Image in Sickle Cell Classification Using Deep Learning Algorithm
Word Count:	**9843 Page Count 28**

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: ……

Date: ……

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Analysis of Microscopic Blood Images in Sickle Cell Classification Using Deep Learning Algorithm

Ifeoma Oduntan
X18191274

Abstract

Abstract. Sickle cell disease is the most common form of hereditary blood disorder that is associated with hemoglobin abnormality due to mutation in β-globin genes known as hemoglobin S. It is estimated that 20 million people around the world live with the disease and a total of 176,000 deaths were recorded in 2013. The traditional method of diagnosing it is through conventional analysis of peripheral blood smears under the microscope by a pathologist which is laborious, time consuming and can lead to delays and misdiagnosis. Currently, the conventional machine learning technique still depends on the expertise knowledge of medical practitioners to select the features, and this can affect the classifier's accuracy due to the subjective nature of the process. Developing an automatic way of diagnosing this disease through classification of the red blood cells as early as possible is a challenge due to lack of data in the medical field. This research aims to apply a deep learning technique that implements a novel Deep Convolutional Generative Adversarial Networks (DCGANs) for image synthesis to overcome small dataset issue for efficient classification and diagnosis of sickle cell disease. The augmented erythrocytesIDB1 dataset is used as an input to DCGANs to generate more images which can be used to train six deep transfer learning image classification models namely DenseNet121, ResNet50, InceptionV3, VGG16, VGG19, and MobileNet based on three types of red blood cells namely circular (normal), elongated (sickle cells), and other abnormality. The performance of the models is compared on the original images, GAN generated images/original images, and the traditional augmented images/original images to see the effect of each dataset on each model and find out if GAN generated images are realistic and can be an alternative source for augmenting data for classification in situation where the data size is very small, especially in the medical field and also identify the optimal classification model. The results are presented based on weighted metrics of accuracy, precision, recall, and F1-score and it showed that model performance on GAN generated images improved between 4.5% to 136% in all the models and MobileNet model achieved the highest accuracy and recall of 99.70%.

Keywords - *Microscopic blood images, Generative Adversarial Networks, Data Augmentation, Red blood cells, Sickle Cell Disease, Transfer learning, Deep learning*.

1 Introduction

Sickle cell disease (SCD) is a hereditary blood disorder that affects hemoglobin molecules in red blood cells (RBCs). This abnormality occurs because of mutation in β-globin genes known as hemoglobin S (Li et al., 2022). This molecule delivers oxygen to the cells throughout the body and its mutation can cause polymerization of the hemoglobin molecules in RBCs. Its occurrence affects the shape and elasticity of the cells and causes fragility of RBCs. It is difficult for these cells to move through small blood vessels, and this can lead to multiple organ failures, pain crisis and even death if undiagnosed

The most available study of 2013[1] reported that about 3.2 million people live with SCD, and this number has grown to 20 million people around the world (ibid.). However, reports have shown that a total of 43 million people have sickle cell trait, and a total of 176,000 deaths were recorded in 2013 with most of the patients from African continent. Considering the seriousness of this condition with more than 300,000 babies born every year with SCD. Half of these births occurred in three countries (Nigeria, India, and Democratic Republic of Congo). It should be noted that many of the countries with high poverty levels have not benefited from the advances made in sickle cell treatment[2] . Due to lack of cure, early diagnosis is very crucial in the treatment and management of the disease.

1.1 **Research Motivation and Background**

SCD patients are faced with risk of life-threatening complications in the form of organ damage and stroke over time and this can lead to reduced life expectancy. Reason being that sickled red blood cells which are stiff, and rigid can get stuck in small blood vessels. Figures 1a and 1b below show the images of normal red blood cells and sickled red blood cells, and an abnormally shaped cells in sickle cell disease.

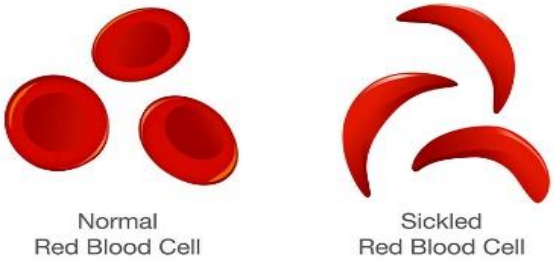


Figure 1a

Figure1b

Figures 1a and 1b: Images of normal and sickled red blood cells and an abnormally shaped cells in Sickle Cell Disease

Figure 1b is the image of normal red blood cells which can easily pass through the blood vessels, but the sickled RBCs get stuck in the blood vessel, and this deprives the organs and tissues of oxygen rich blood. Signs and symptoms of SCD starts in early childhood and is characterised by low counts of RBCs and even though the symptoms can be very severe, there

[1] https://linkinghub.elsevier.com/retrieve/pii/S0140673614616822

[2] How Common Is Sickle Cell Disease? (sickle-cell.com)

are limited options for treatment like bone marrow transplant and is only possible if a suitable donor can be found which can reduce the complications if diagnosed early (Li et al., 2022).

Approaches to SCD diagnosis have evolved from manual classification of peripheral blood smears by a pathologist to the conventional machine learning technique. The manual process is laborious and time consuming. Moreover, the lack of an expertise medical practitioner can lead to delays and inaccurate results. This is because of the heterogenous nature of the red blood cells which can make manual diagnosis very complicated. Also, the conventional machine learning technique still depends on the expertise knowledge of medical practitioners to select the features, which can affect the classifier's accuracy due to the subjective nature of the process. However, classification of SCD is not without challenges (Alzubaidi et al., 2020). These challenges are encountered in the area of data acquisition, lack of nuclei, blurry boundary, overlapped cells, and complex nature of RBCs shapes present in SCD. Moreover, there is still no known efficient technique that can analyse, detect, and classify sickle cell disease in RBCs. There is need to take advantage of breakthroughs in deep learning algorithms to find an automated method of detecting, and classifying this condition as early as possible, so as to reduce the time consuming and labour-intensive tasks performed by these medical practitioners, reduce cost to patients and prevent early death. But lack of large dataset is a big problem.

1.2 Research Question

How well can deep learning techniques be employed in image generation for data augmentation to overcome small size dataset challenges in the classification of sickle cell disease for early detection and diagnosis to give SCD patients good quality of life, prevent early death, and save the medical practitioners time to concentrate on treating their patients?

1.3 Research Objectives and Contribution

For the research question to be addressed, this research project will need to critically analyse the literature review to give insight into the current state of the art and help identify research gaps.

Table 1: Research Objectives

Objective	Description	Evaluation Method
1	Review current state of the art in Sickle Cell Disease based on classification and data augmentation techniques (2012 – 2022)	Literature Review
2	• To apply augmentation techniques using novel GAN technique and other traditional data augmentation methods like rotation, vertical and horizontal flipping, zoom, width and height shift, and shear to generate images.	• Inception Score (IS) • Visual quality of generated images
3	Proposed Techniques to be Implemented in the classification of Sickle Cell Disease are: • To implement DenseNet121 • To implement Inception V3 • To Implement MobileNet • To implement VGG16 • To implement VGG19 • To implement ResNet50	• Confusion Matrix • Weighted Accuracy • Weighted Precision • Weighted Recall • Weighted F1-score • Loss
4	To comparatively analyse the performance metrics of all models based on: • Original images • Traditional augmented images and original images • GAN generated images/Original Images	• Confusion Matrix • Weighted Accuracy • Weighted Precision • Weighted Recall • Weighted F1-score
5	To compare the proposed models with the state - of - the - arts	

Table 1 is the research objectives that will help in addressing the research question.

This research project is aimed at using Deep Convolutional Generative Adversarial Networks (DCGANs) to generate new images for data augmentation and apply six deep transfer learning techniques to classify sickle cell disease based on original images, augmented and GAN generated images. The major contribution of this research is the application of a novel Generative Adversarial Networks (GANs) model for synthesizing new data to augment the dataset so as to overcome small data size problem and compare the results to find out if GAN generated images are realistic and can be an alternative source of augmenting data for classification in situation where the data size is very small, especially in the medical field. This will build on the work done by (Alzubaidi et al., 2020) in which they used transfer learning techniques that were trained on data that is related to their dataset, instead of the state-of-the-art models that were trained on ImageNet dataset to address the issue of lack of training data. Six deep learning classification algorithms will be used to model the images from the erythrocytesIDB1 dataset and the GAN generated images for comparison. A minor contribution is the detailed review of the state-of-the-art techniques for detection and classification of sickle cell disease.

This research project also discusses different techniques used in traditional augmentation and GAN image synthesis, and classification of biomedical and microscopic images (2012 - 2022) in section 2. Section 3 describes the research methodology used in this project. Section 4 discusses the design components for the analysis of microscopic blood images in sickle cell classification. Section 5 discusses the implementation of this research. Section 6 presents and discusses the research results and evaluations. Section 7 concludes the research and discussions on future work.

2 Related Work on Classification Analysis of Sickle Cell Disease (2012 – 2022)

2.1 Introduction

So much research have been carried out on classification analysis of sickle cell disease and different techniques have been proposed based on machine learning and recently on deep learning of microscopic images because of breakthrough in the field of computer vision. These methods have mainly centered on deep learning algorithms used in data augmentation, (Szegedy et al., 2015), classification problems (Huang et al., 2017) and data segmentation problems (Ronneberger, Fischer, and Brox, 2015). However, the literature review will concentrate on data augmentation and classification which is the task of this project. The literature review is divided into different subsections of challenges in classification analysis of biomedical images, data augmentation techniques like traditional augmentation and GAN techniques, critical review of deep learning techniques used in the classification analysis of biomedical images, and conclusion and identified gaps.

2.2 Challenges in Classification Analysis of Sickle Cell Disease

Classification analysis of SCD is not without challenges and some of these are in data acquisition which led to the proposal of GANs by (Goodfellow et al., 2014) in image generation so as to aid data augmentation. Other challenges are in blurry boundary due to the influence of imaging procedure, overlapping of cells, complex nature of the RBCs shapes present in SCD, low intensity contrast between RBCs region and the background, lack of nucleus in RBCs meaning that nuclei location marker technique cannot be used, presence of artefacts as a result

of dirt on the imaging light path and shading (Alzubaidi et al., 2020). Some of these challenges will be reviewed in this literature.

2.3 Literature Review on Data Augmentation Techniques in Sickle Cell Classification and Biomedical Images

Lack of data in the medical and health sector is a big challenge because of the release of patient's information without their consent which can cause ethical issues and the issue of lack of enough data annotators, which requires experienced and skilled medical professionals to label these data correctly. Also, the subjective nature of these annotators can lead to different interpretations by different medical professionals. Because of these challenges it is difficult to obtain large datasets in the medical field unlike other sectors (Calimeri et al., 2017).

2.3.1 Traditional Data Augmentation Techniques

Traditional data augmentation is a technique of applying some alterations to existing data in order to increase its diversity and enlarge dataset without collecting new data (Zhong et al., 2020). It helps in preventing deep neural networks from learning irrelevant features that prevents it from generalizing, and this results in better model performance. To effectively train the classifier to address the issues of insufficient data problem (Ding et al., 2019) implemented several methods of augmentation to generate additional training samples from existing data. Their report showed that AlexNet accurately classified the three diffraction images. While (Naruenatthanaset et al., 2020) applied random flips and rotation to augment their imbalance datasets as a result of the classes' sensitivity to size and color. Their report showed that EfficientNet model with augmentation has the best accuracy as to classification without augmentation.

2.3.2 Generative Adversarial Networks (GANS)

Paper by (Ronneberger, Fischer, and Brox, 2015) opined that there is need for large amount of annotated training samples in order to achieve successful training of deep networks. Generative adversarial networks (GANs) by (Goodfellow et al., 2014) is a deep learning model that consist of Generator (G) and Discriminator (D) which plays a minimax two player game in which the generator G generates fake images from latent vector which is similar to original images and the discriminator (D) tries to distinguish between the fake and real images. A lot of research have been carried out in GANs leading to many different versions and more research have been carried out on how to improve their training stability in order to solve the gradient vanishing problem and mode collapse. This has been resolved by (Goodfellow et al., 2014) in which the objective function was modified. However, (Metz et al., 2017) proposed unrolled GANs which successfully resolved the issue of model collapse, while (Salimans et al., 2016) developed Inception score for evaluating GAN generated images so as to remove human annotators. Paper by (Radford & Metz, 2016) used DCGANs on three datasets. They only applied scaling of the data to the range of the tanh activation function without any other pre-processing technique. They used mini-batch size of 128 with leakyRelu of slope 0.2 and adam optimizer. However, they used a learning rate of 0.0002 as they found the 0.001 rate was very high. They also noted that momentum of 0.9 value led to oscillation and instability, and they reduced it to 0.5 which helped in stabilizing the training. They concluded that DCGAN performed very well when compared to others with a test error of 2.98% and 1.48% respectively at 50,000 samples and 10 million samples. However, (Bailo, Ham, and Min Shin, 2019) applied conditional generative

adversarial networks (cGANs) in the synthesis of microscopic new images for the augmentation of the size of their small dataset. They implemented this using image-to-image translation technique, and this resulted in a marginal improvement of their result.

Paper by (Bowles et al., 2018) investigated the application of GAN in modelling the underlying distribution of training data for augmentation purposes. Their work compared the results of GANs with the rotation augmentation using Progressive Growing of GANS (PGGAN), because of its training stability on large image sizes and robustness to hyperparameter selection. They explained that populating training data with realistic synthesized data can reduce overfitting significantly and also improve the generalization ability of the algorithm thereby boosting the overall classification accuracy. They opined that there is clear evidence that improvement in synthesized data depends on the amount of available data which is likely due to little data to train the GAN properly. They concluded that even though GANs does not have the ability to extrapolate unlike the traditional augmentation, it can still provide an effective way of filling in the gaps in discrete training data distribution thereby augmenting the sources of variance that are not able to augment in other ways. However, they pointed out that GANs will not be able to extend data distribution beyond the extremes of the training data and noted that improvements achieved by using both GAN and traditional augmentation techniques were consistently more compared to the sum of the improvements achieved by using both methods separately. Paper by (Wieczorek et al., 2021) used DCGAN to generate images in the classification of holed drilled in laminated chipboard. They increased their data size by traditional data augmentation before implementing GAN to generate images for each class and they opined that the generated images were good quality to be used in training the model.

In (Bang and Shim), they proposed Representative feature based GANs to resolve the training instability by extracting representative features from a pre-trained autoencoder. This is transferred to the discriminator thereby implicitly enforcing the discriminator to be updated by effectively considering both reverse and forward KL divergence.

2.4 Critical Review of Deep learning Techniques Used in the Classification Analysis of Biomedical Images

Deep convolutional neural networks have been widely applied in medicine for diagnosis of disease, abnormality detection, organ segmentation and classification in many fields including biomedical images (Zhou et al., 2022). Work by (Xu et al., 2017) used deep Convolutional Neural Networks (dCNN) in the classification of SCD. Their study was based on a very small dataset, and this shows that the model is not robust. Whereas (Tengshe et al., 2021) implemented CNN in sickle cell detection and their work achieved an accuracy of 95%. Work by (Elsalamony, 2016) implemented neural network in classifying healthy and unhealthy red blood cells. Work by (Gual-Arnau, Herold-García, and Simó, 2015) achieved an accuracy of 96.10% on erythrocytesIDB1 dataset and paper by (Rodrigues, Naldi, and Mari, 2016) achieved an accuracy of 94.59% on the same dataset, but (de Faria, Rodrigues, and Mari, 2018) achieved an accuracy of 93.67% on the same dataset. However, paper by (Alzubaidi et al., 2020) addressed the issue of small data size issue by applying transfer learning technique using data from the same domain area. Their work also applied different augmentation techniques so as to minimize overfitting. They claimed that their lightweight deep learning models outperformed the latest methods with an accuracy of 99.98% on the erythrocytesIDB1 dataset.

Paper by (Cheuque et al., 2022) implemented a two stage multi-level scheme by applying Faster R-CNN to identifying the region of interest in the white blood cells to separate mononuclear cells from polymorphonuclear ones. They then implemented two parallel Convolutional Neural Networks (CNNs) which is based on MobileNet architecture for

classification task. MobileNet architecture is a class of lightweight deep convolutional neural networks that are notably smaller in size and at the same time faster in performance than many other state-of-the-arts models. Due to their small low latency and low power, they can be used in classification and detection tasks (Howard et al., 2017). They achieved an accuracy of 98% using Monte Carlo cross validation method. Paper by (Sandler et al., 2019) proposed MobileNet V2 which is the second version of the MobileNet architecture and has a significantly lower number of parameters compared to MobileNetV1 and this makes it more lightweight. It is best suited for mobile devices as they are faster due to a reduction in size and complexity.

However, (Huang et al., 2017) introduced DenseNet and this outperformed other state of the art algorithms due to its ability in strengthening feature propagation, solve the vanishing gradient problem and reducing the number of parameters. Work by (Sen et al., 2021) implemented five state of the arts pre-trained models and different types of data augmentation techniques. They used small data size, and this could have affected the performance of their models. They reported that Inception V3 model achieved the highest accuracy. They could have evaluated their work using at least F1-score, recall and precision considering that misclassification can lead to early death in SCD. Work by (Yee and Raymond, 2020) implemented Inception V3 in extracting features from the chest X-ray images and then used different machine learning algorithms to classify the pneumonia. They reported that Neural Network outperformed the other machine learning algorithms. Study by (Bressem et al., 2020) compared different deep learning architectures in chest radiographs classification. Their study showed that deeper neural networks do not necessarily achieve better results than shallow networks like AlexNet or VGG-16 and they opined that CNN with fewer layers have low computational requirement and shorter training time. Paper by (Sharma et al., 2022) implemented DenseNet121 to classify different types of white blood cells (WBC). They applied normalization and augmentation techniques with different batch sizes and claimed their model achieved an accuracy of 99% with batch size of 8. Their work concentrated on white blood cells. But (Norouzifard et al., 2018) implemented Inception ResNet-V2 and VGG19 in the classification of glaucoma on optic nerve head (ONH) images to overcome overfitting because of the small size of data. Their report showed that the VGG19 model had an overfitting problem leading to very poor performance. They concluded that VGG19 model was unable to generalize but the Inception ResNet-V2 model achieved a very high performance at epoch 30.

Table 2: Summary of the state-of-the-art models for classification of erythrocytesIDB1 dataset

Authors	Achieved results (Accuracy- %)
Gual-Arnau et al. 2015	96.10
Rodrigues et al., 2016	93.18
Rodrigues et al., 2016	93.07
Rodrigues et al., 2016	94.59
de Faria et al., 2018	92.52
de Faria et al., 2018	93.67
Alzubaidi et al., 2020 (Scenario 4, Model 2)	99.54
Alzubaidi et al., 2020 (Scenario 4, Model 2 + SVM)	99.98

Table 2 is the summary of the state-of-the-art models for classification of erythrocytesIDB1 dataset, and it shows that most of the deep learning algorithms achieved very good accuracy with (Alzubaidi et al., 2020) achieving the highest result on the same dataset.

2.5 Conclusion and Identified Gaps

The interest and seriousness of SCD can be seen from the amount of literature in this research area. Although some of the methods in the literature achieved good results based on small sized datasets, the advancement in technology has given way to application of new techniques of deep learning which can employ transfer learning and Generative Adversarial Network (GAN) to overcome small size dataset issues for efficient diagnosis of SCD. This is a significant gap, and this research project is aimed at addressing this gap by using the GAN model to synthesize data in conjunction with different augmentation techniques like rotation and flipping applied in (Alzubaidi et al., 2020). In conclusion, literature review in the classification of biomedical images as a whole and microscopic blood cell images for SCD classification will be beneficial in image synthesis using GAN and other traditional data augmentation techniques. This will help to overcome the challenges posed by lack of dataset in the medical field.

3 Research Methodology

3.1 Introduction

The ability to detect and classify the red blood cells in Sickle Cell diagnosis can be complicated by the heterogenous nature of the red blood cells. This is due to the different shapes, sizes, location, and edge. Finding an automatic way to classify and diagnose the disease is very crucial in the treatment, management, and prevention of early death in SCD patients. The methods and specifications are discussed in detail below.

3.2 Modified CRISP-DM Methodology Design

Of all the different methods that can be applied in research in the field of Computer Science (CRISP_DM, KDD and SEMA), adapted Cross Industry Standard Process for Data Mining (CRISP-DM)[3] methodology was chosen as shown in figure 2 below. This is because of its flexibility which allows for modification of the approach due to the nature of the project.

[3] https://www.ibm.com/docs/en/spss-modeler/saas?topic=understanding-determining-business-objectives

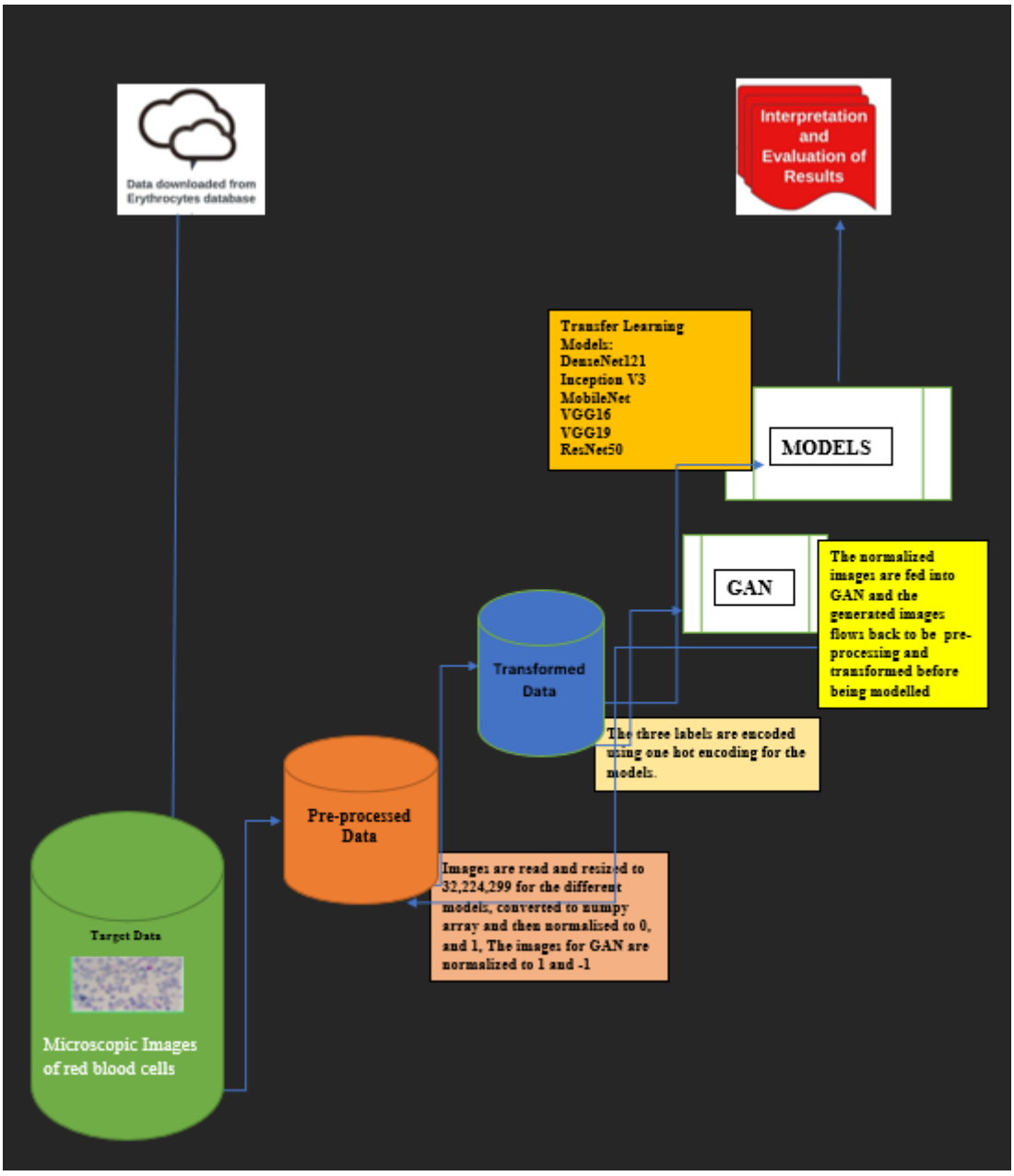


Figure 2: Adapted CRISP-DM methodology process for sickle cell classification

The methodology process entails the following steps:

3.2.1 Data Understanding and Exploratory Data Analysis

There are diverse types of available data around blood images, but the data that will be appropriate for this research is the one that is more tailored to the aim of the research in sickle cell classification. The dataset used is the erythrocytesIDB1 dataset (Gonzalez-Hidalgo *et al.*, 2015) which are images of peripheral blood smear samples from sickle cell patients in the Special Hematology Department of the Santiago de Cuba General Hospital[4]. This is a private

[4] http://erythrocytesidb.uib.es/

dataset which was requested through email from their website and downloaded from the link provided by the data owner from their cloud storage and was then stored in google drive.

Exploratory data analysis shows that the dataset has 196 full-field images and 626 individual cell images of size 80 x 80 pixels in three classes as circular (202), elongated (211), and other (213). The dataset is in Joint Photographic Experts Group (JPEG) format and initial analysis shows that it is a balanced dataset with almost equal classes and there are no damaged images.

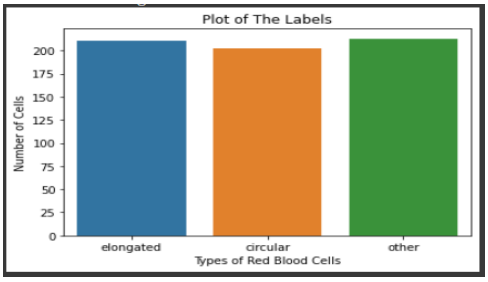


Figure 3a

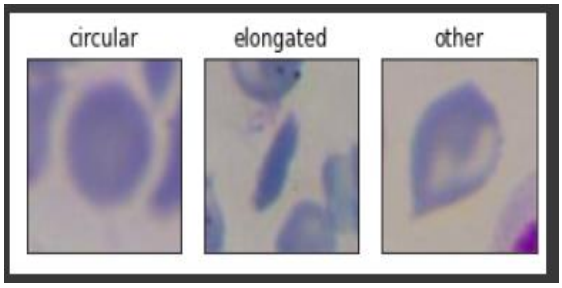


Figure 3b

Figures 3a and 3b: Bar plot of the classes and sample image from each class

Figure 3a is the bar plot of each class and figure 3b is the sample image from each class, which are circular (normal), elongated (sickle cell) and other types.

3.3 Data Pre-processing

The aim of image pre-processing is to improve the quality of the images by standardizing it before being fed into a model to increase its accuracy and at the same time reduce model complexity. In the implementation of sickle cell disease classification, the first process was to mount the drive in Google Colab and install all the necessary packages in Google Colab using Python, TensorFlow and Keras API. The data was read for the initial exploratory data analysis. Some of the images in different classes were visualized. The images were then resized to 244 x 244, and 32 x 32 and 299 x 299 to match the size of the input layer of each of the algorithms, using the open CV library. The images are then transformed into an image array.

Normalization, which is also known as rescaling is the process of projecting pixels (intensity) of an image data to a predefined range. The images were normalized to between 0 and 1 by dividing by 255, as the range of pixels in an image is from 0(black) to 255(white) and between -1 and 1 for GAN training. The aim is to make all the images to equally contribute to the total loss, rather than some images with high pixels contributing more than low pixel images. This will provide a standard learning rate because images with high pixels require a low learning rate and vice versa. The three image classes were encoded using the sklearn preprocessing Label binarizer into one hot encoding for the labels. Two types of augmentations were performed.

The traditional data augmentation method is implemented due to the small size of the dataset. The **erythrocytes** dataset was augmented, and this was achieved by using the traditional augmentation method of rotation at an angle range of 30°. Vertical and horizontal flipping were used as they will not distort the context of the images. Slight width and height shifting was applied in the range of 0.2 respectively, zoom range of 0.2, and shear range of 0.2, were implemented. Each image in each class was augmented 7 times with a total augmented images of 4382 plus the original images bringing the final total to 5001 images and was saved in a folder named **erythrocytes2**

The GAN image augmentation technique: The **erythrocytes2** dataset was used as an input in training GAN model as implemented in (Bowles et al., 2018, Wieczorek et al., 2021) in which they opined that there is clear evidence that improvement in synthesized data depends on the amount of available data, which is why traditional augmentation is applied first before feeding the data into GAN with the original images. The images were processed by resizing to 32 x 32 using the open CV library and were transformed into an image array.

Normalization: The images were normalized by converting to a float and then converted to between -1 and 1 by subtracting 127.5 and divided by 127.5. The resized images were first downsampled in the discriminator as it was trained first. Then the generator was fed with a noise vector which was upsampled to size 32 x 32 and the generated images were then passed to the discriminator to determine whether they are fake or real images. This was monitored at different epochs and the images achieved 100% accuracy at epoch 500 for each of the classes. The images generated by GAN were saved in a folder named **erythrocytes3** which has 2000 images per each class totaling 6626 images inclusive of the original images. This was done so that comparison can be made on the effect of different datasets on the models. The final data for fitting the models was then split into training, validation and test sets using a ratio of 80:10:10. This partitioning of the data will help in modelling to make sure that the model has not seen the test and validation data. However, the data used for GAN was split into 80 :20 ratio. The weights of the models for transfer learning were used to extract features which are the shape, size and edges from the dataset, and the top layers of each model was set to false because they are the classification layers for the pre-trained models which is different from the sickle cell classification task. The top layers were fine-tuned to the task of this project which is 3 classes rather than 1000 classes they were trained on. Table 3 below is the final total number of images in each dataset.

Table 3: Final Total of Images in each Dataset

Datasets	Total Number
Erythrocytes (Original Dataset)	626
Erythrocytes2(Augmented/Original Images)	5001
Erythrocytes3(GAN generated images/Original images)	6626
GanImages (GAN generated Images for Evaluating the quality of the images)	6000

3.4 Model Application

The model application stage of the methodology is particularly important because making the right decision on algorithm selection that is in line with the dataset and the task is very crucial in achieving a better outcome. The following diverse deep learning algorithms which are DenseNet121, ResNet50, InceptionV3, VGG16, VGG19, and MobileNet will be used to fit the three datasets.

3.5 Performance Evaluation

Given that this project is in the medical field, the ideal metrics to evaluate the performance of each model is weighted metrics for accuracy, precision, recall and F1-score.

4 Design Specification

The 2-tier design specification process for the classification of sickle cell disease is shown in figure 4 below and it shows how the data will flow from the client tier as stored data in the cloud into the business tier where it will be pre-processed, then both the augmented images and original images are transformed through normalization before they are fed into GAN to generate new images. The generated images flow back to preprocessing and normalization stages before it is fed into the transfer learning models for classification. Both original images, traditional augmented images and GAN generated images are then preprocessed again before they are fed into the six pretrained models using the TensorFlow and Keras API on Python language in the Google Colab platform. The results will flow back to the client tier where they are visualized, evaluated, and interpreted.

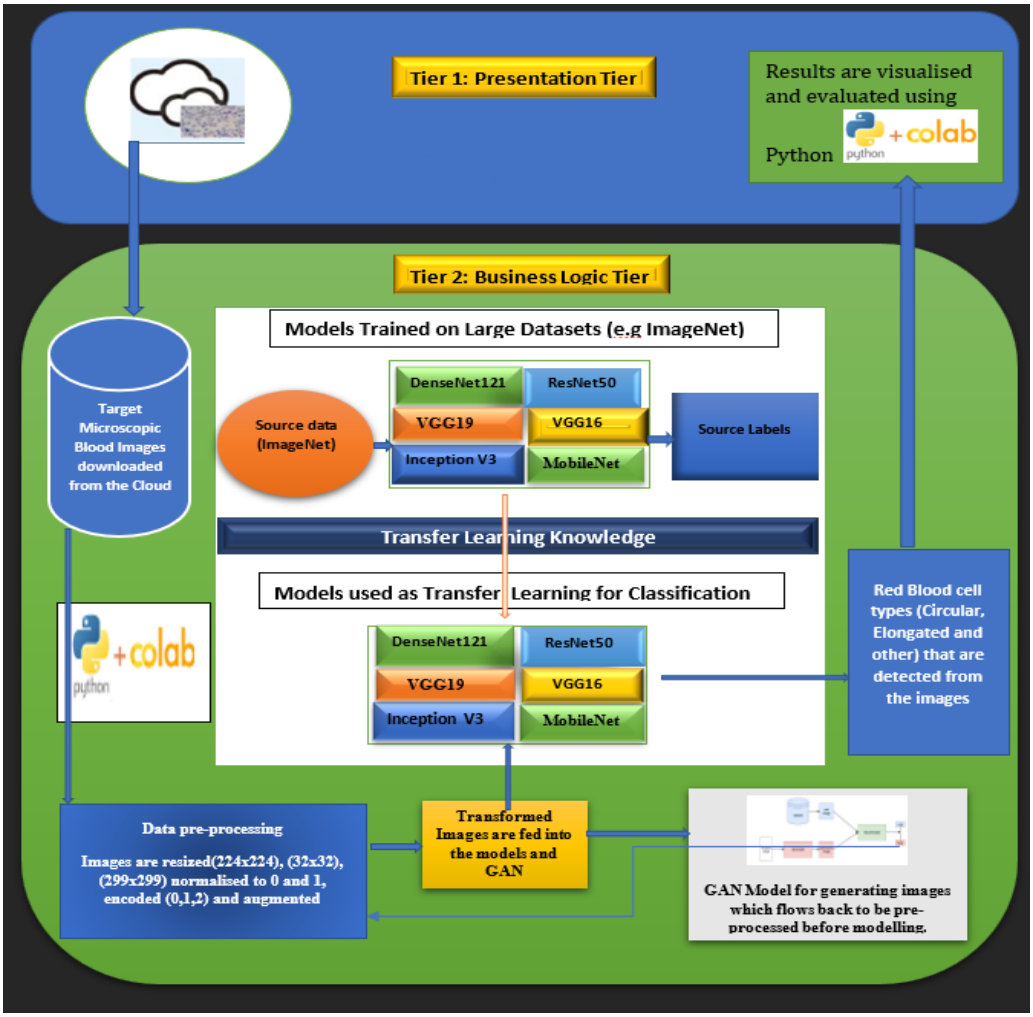


Figure 4: Design process for Sickle Cell Anemia Classification

4.1 Model Specifications

4.1.1 Generative Adversarial Networks (GANS)

The discriminator consists of convolution2D of 64, 128, 128 and 256 with LeakyRelu of alpha 0.2 and adam optimizer with learning rate of 0.0002, dropout rate of 0.4, sigmoid activation

and was compiled with binary cross entropy activation. The generator model used convolution2D transpose of 256, 128, 128, 128 with LeakyRelu of alpha 0.2, tanh activation and adam optimizer with learning rate 0.0002 and beta_1 of 0.5.) the generator has a 4x4 x256 input nodes which is upsampled to 32x32 with Conv2DTranspose and stride of 2 by 2, kernel size of 4 x 4, tanh activation and LeakyRelu of 0.2.

4.1.2 DenseNet121

This model has 120 convolutional layers and 4 average pooling layers. In this study, DenseNet121 model with growth rate of k = 32 is used due to limited sample, since more complex networks requires more samples to be trained.

4.1.3 Inception V3

The Inception V3 model which has a 48 layer deep convolutional neural network is also used in image classification and recognition tasks. It uses several techniques of label smoothing, auxiliary classifier to propagate label information lower down the network. It uses factorized 7x7 convolutions and batch normalization for optimizing the network and this network is deeper than model V1 and V2.

4.1.4 MobileNet

The model has 30 layers made up of 27 convolutional layers with 1 average pool layer and 1 fully connected layer and SoftMax. The MobileNet architecture has a significantly lower number of parameters compared to other models. It is best suited for mobile devices as they are faster due to a reduction in size and complexity.

4.1.5 VGG16

Visual Geometry Group (VGG16) has 16 layers made up of 13 convolutional layers with stride size of 1, and 2 x 2 pooling layers with stride of 2 and 3 fully connected layers.

4.1.6 VGG19

This model is another variant of VGG models and is made up of 19 layers of which 16 are convolutional layers and 3 are fully connected layers with 5 Max pooling layers and one SoftMax.

4.1.7 ResNet50

The ResNet50 has a total of 50 layers that is made up of 5 blocks of 3 convolutional blocks with 3 layers and 3 convolutional layers in each identity block. It has over 23 million trainable parameters and uses skip connections thereby making the network learn identity function that enables it to pass the input through the block without passing through the other weight layers and this solves the gradient vanishing problem.

5 Implementation of Sickle Cell Disease Classification

5.1 Introduction

Modelling is the major step in methodology. Six different classification models (DenseNet121, ResNet50, InceptionV3, VGG16, VGG19, MobileNet) were implemented to see the effect of each model when fitted on the three different datasets which are (1) **erythrocytes** - the original images with augmentation of training dataset only, (2) **erythrocytes2 -** the traditional augmented images, and (3) **erythrocytes3 -** the GAN generated images and the original images. However, different techniques were used in training the models to identify the best hyperparameters. The results were visualised using plot of training and validation accuracy/loss, accurately predicted classes, and confusion matrix.

5.1.1 GAN

Two models were constructed (Discriminator and Generator). The generator takes a latent dimension vector of size 4 x 4 x 256 to upsampled to size 32 by 32 using Conv2DTranspose, while the discriminator takes an image size of 32 by 32 to downsampled to 256x4 x 4 through Conv2D. Only the discriminator was compiled, while the generator was not compiled because it is trained in the combined model (GAN). Sample size of 1200, 1300 and 1300 for each class was used with a batch size of 128, Adam optimizer (0.0002 and beta 0.2) and 500 epochs for each of the classes. Inception Score was implemented in evaluating the quality and diversity of the generated images. The results were plotted, and the images saved in a folder named erythrocytes3 and GANImages.

5.1.2 DenseNet121

The DenseNet121 was implemented using the three datasets with an input size of 224 x 224, two dense layers of 1024 and 512 neurones with Relu activation, Batch Normalization, two dropout rates of 0.5, 20 epochs with batch size of 128 and SoftMax activation function at the final classification layer. The model was fine-tuned by training the last 8 layers for better classification accuracy and was compiled using Adam optimizer at a learning rate of 0.0001/0.0002 and categorical cross entropy as the loss function. However, call-backs were implemented, and validation accuracy was monitored with a patience of 5.

5.1.3 Inception V3

This model was implemented using input shape of 224 x 224, 1024 and 512 dense connections with two dropout rates of 0.25 and global average pooling on the three datasets. The last 15 layers of the model were trained to fine-tune the model as this will help the model learn some of the features. It was compiled using the adam optimizer with a learning rate of 0.0001/0.0002, SoftMax activation and early stopping were used with a patience of 5. Categorical cross entropy was used as the loss function. Model checkpoint was implemented to save the best model. Epoch of 20 was used with a batch size of 128.

5.1.4 MobileNet

This model was implemented using the Keras functional API on the three datasets with an input shape of 224 by 224, two dense connections of 1024, and one 512 neurones with Relu activation, Batch Normalization, two dropout rates of 0.5. The model was fine-tuned by using

the 6th to the last layer of the original model to build a new model with SoftMax activation function. So, the last 23 layers were trained, and the new fine-tuned model was compiled using an adam optimizer and learning rate of 0.0001/0.0002, and 20 epochs with a batch size of 128.

5.1.5 VGG16

This model was implemented using a sequential model on the three datasets with an input shape of 224 x 224, 1024 and 512 dense layers, global average pooling and two dropout rates of 0.25. The last 8 layers were fine-tuned by flattening them to reflect the task of this project which has 3 classes, as opposed to 1000 classes it was trained on. The model was compiled using Adam as the optimizer with a learning rate of 0.0001/0.0002, SoftMax activation with categorical cross-entropy as the loss function because of the 3 classes (multiclass classification). Reduce Learning rate on plateau was applied to monitor validation accuracy with a factor of 0.5. Batch size of 128 and 20 epochs were used.

5.1.6 VGG19

This model was implemented using a sequential model on the three datasets with an input shape of 224 x 224, 1024 and 512 dense layers, global average pooling and two dropout rates of 0.25. The last 8 layers were fine-tuned by flattening them to reflect the task of this project which has 3 classes. The model was compiled using Adam as the optimizer with a learning rate of 0.0001/0.0002, SoftMax function, with categorical cross-entropy as the loss function because of the multiclass classification. Batch size of 128 and 20 epochs were used.

5.1.7 ResNet50

This model was implemented using sequential model on the three datasets with input size of 224 x 224, two dense layers of 1024 and one 512 with two dropout rates of 0.5 and global average pooling with BatchNormalization, SoftMax activation and Relu. The model was compiled using Adam optimizer with a learning rate of 0.0001/0.0002, categorical cross entropy as loss function, and the metric used is accuracy. The last 8 layers of the model were trained to fine tune the model to the task of sickle cell classification. The epoch used was 20 with batch size of 128. Call backs were implemented while monitoring validation accuracy.

5.2 Conclusion

Modelling these deep learning algorithms were time consuming and computationally challenging due to the constraints on usage in Google Colab, as it is timed and can be disconnected at any time especially in the middle of fitting the models. As the data was augmented, this led to crashing of the model fitting due to memory size. Purchasing a monthly subscription from Google Colab was the only way to increase the memory size. Different hyperparameters were used to identify the best performing model. Also, the lack of data availability was challenging in terms of model performance but was rectified by synthesizing images using GAN.

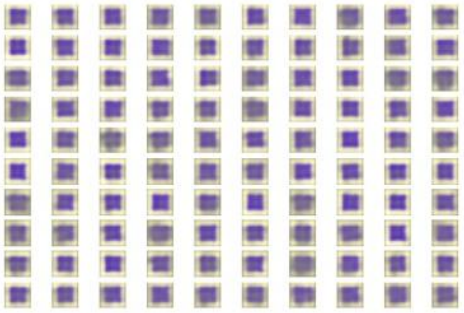
6 Evaluation

Weighted metrics were used to evaluate the performance of the models. This is because there is no true positive and true negative in multiclass classification task unlike binary classification.

The weighted metrics takes into account the total number of samples in each class. This will help in determining which model is the most suitable for this classification task.

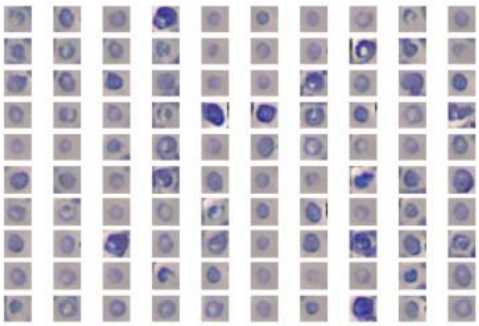
6.1 Experiment 1: GAN

Three different GANs were implemented for the three classes using the **erythrocytes2** dataset and figure 5 below are the progress of image generation at different epochs until the final epochs of 500. The generated images were evaluated and an Inception Score (IS) of 1.82 was reported out of the three classes.

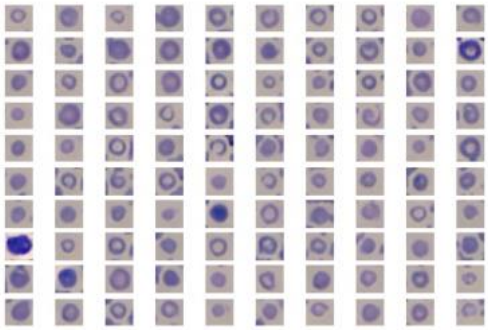


Epoch 10

Epoch 100



Epoch 200



Epoch 500

Figure 5: Example of GAN generated images at different epochs for the Circular cells

6.2 Experiment 2: DenseNet121

The DenseNet121 performed very well on the three datasets, though the model was overfitting on the original dataset. Figures 6a, 6b and 6c below are the plot comparisons of training and validation accuracy/loss, correctly predicted images and confusion matrix.

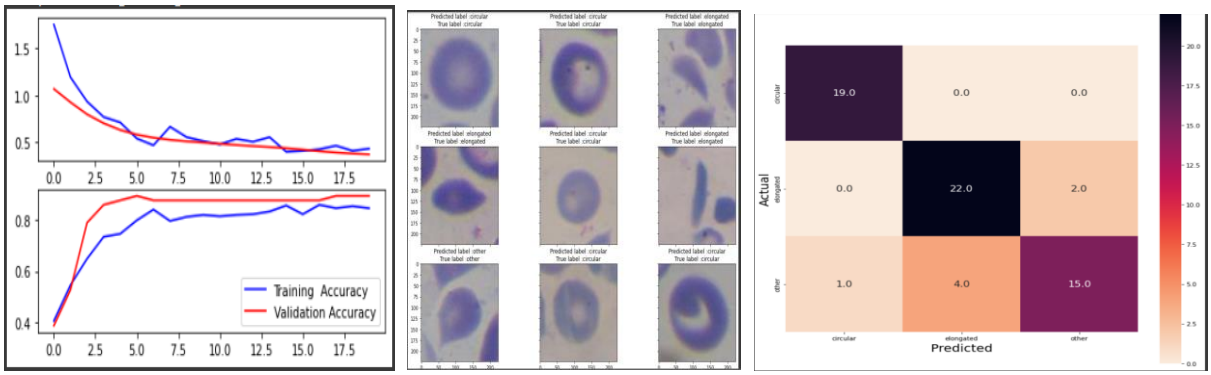


Figure 6a: Results of Modelling the Original dataset (erythrocytes)

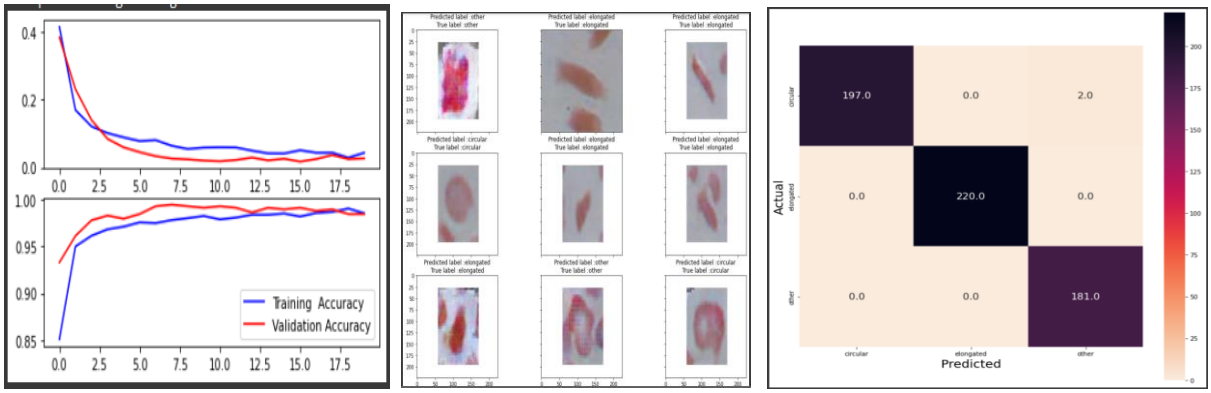


Figure 6b: Result of Modelling GAN Generated Images/Original Images Dataset

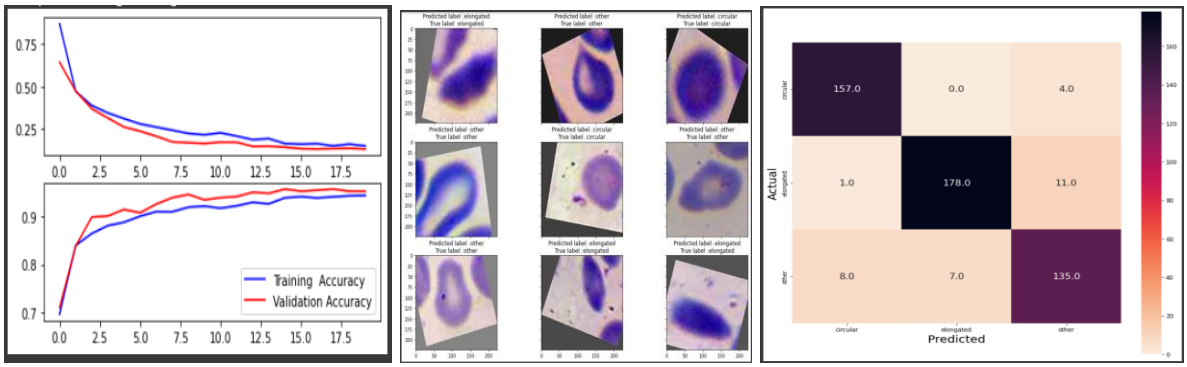


Figure 6c: Result of Modelling Augmented Data Images/Original Images

Figures 6a, 6b and 6c: DenseNet121 Comparison plot of training and validation accuracy and loss, correctly predicted images, and confusion matrix for the three datasets

From figure 6b, it clearly shows that the model performed better with GAN generated images achieving accuracy of 98.90%, with only 2 misclassifications out of 600 images. The model did not misclassify any of the sickle cells while in 6c it achieved an accuracy of 93.80% with 31 misclassifications out of 501 images and 6a achieved an accuracy of 85.70%

Experiment 3: ResNet50

From figure 7b below, it clearly shows that the model performed better with GAN generated images achieving accuracy of 90%, with 3 misclassifications out of 600 images, while in 7c, it achieved an accuracy of 68.70% with 157 misclassifications out of 501 images and 7a achieved an accuracy of 38.10% with 39 misclassifications out of 63 images.

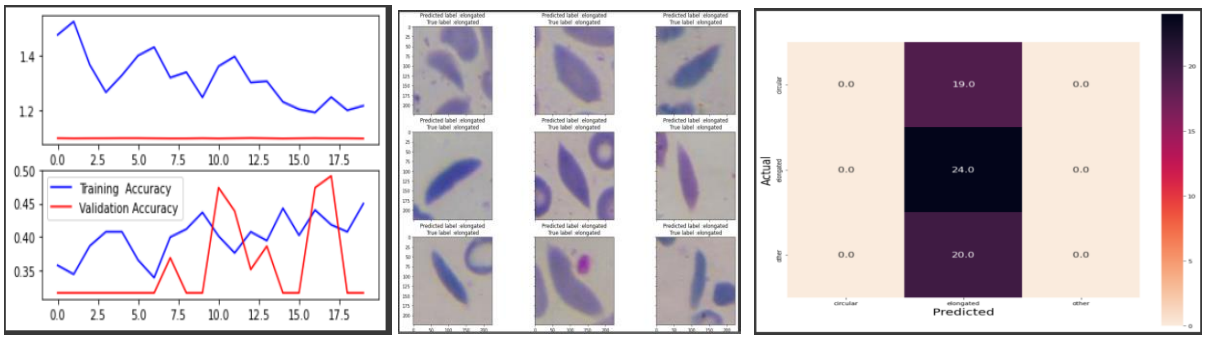


Figure 7a: Result of Modelling the Original dataset (erythrocytes)

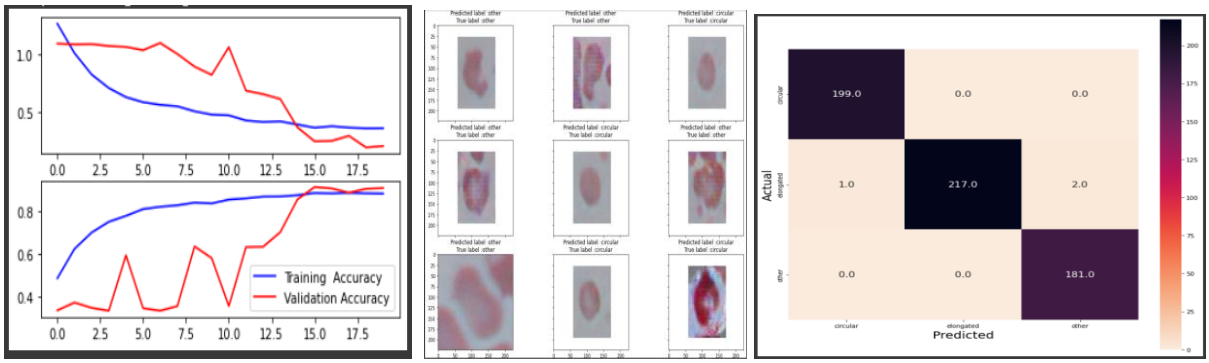


Figure 7b: Result of Modelling GAN Generated Images/Original Images Dataset

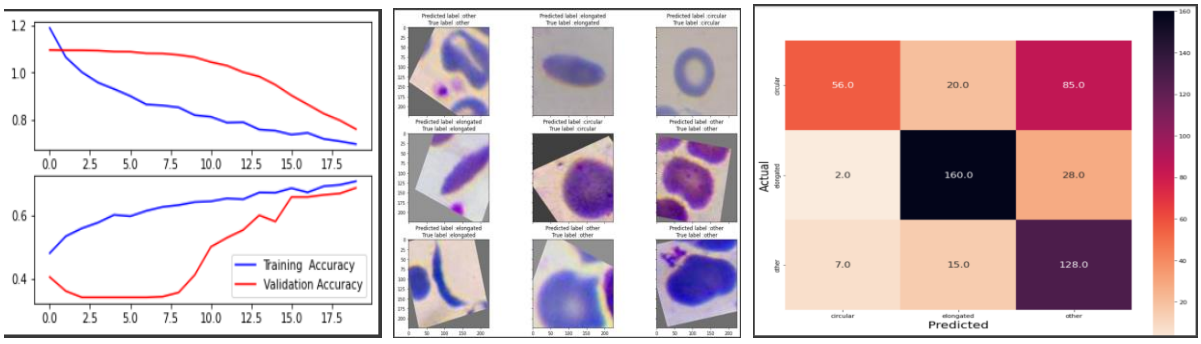


Figure 7c: Result of Modelling Augmented Data Images

Figures 7a, 7b and 7c: Plot of training and validation accuracy and loss, visualised correctly predicted images and plot of the confusion matrix for ResNet50

6.3 Experiment 4: Inception V3

From figure 8b below, it shows that the model performed better with GAN generated images achieving accuracy of 99.50% with no misclassifications out of 600 images, while in 8c, it achieved an accuracy of 97.60% with only 9 misclassifications out of 501 images and 8a achieved an accuracy of 95.20% with 3 misclassifications out of 63 images

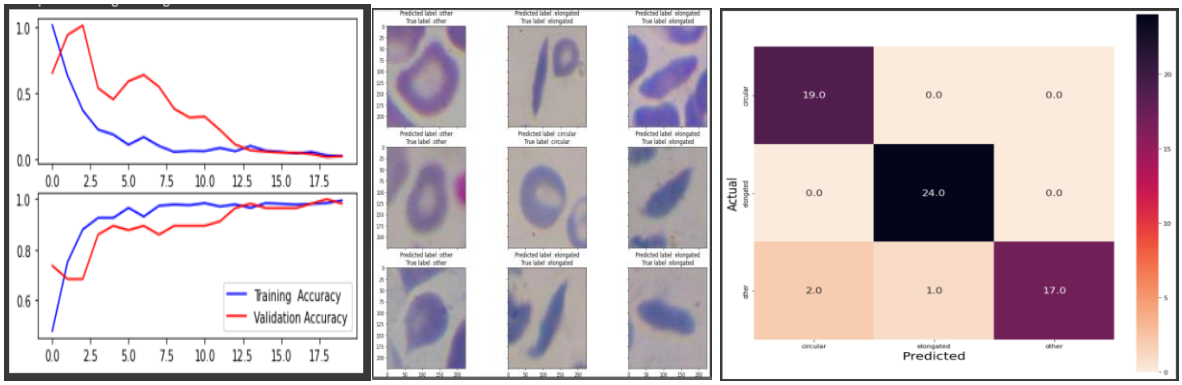


Figure 8a: Result of Modelling the Original dataset (erythrocytes)

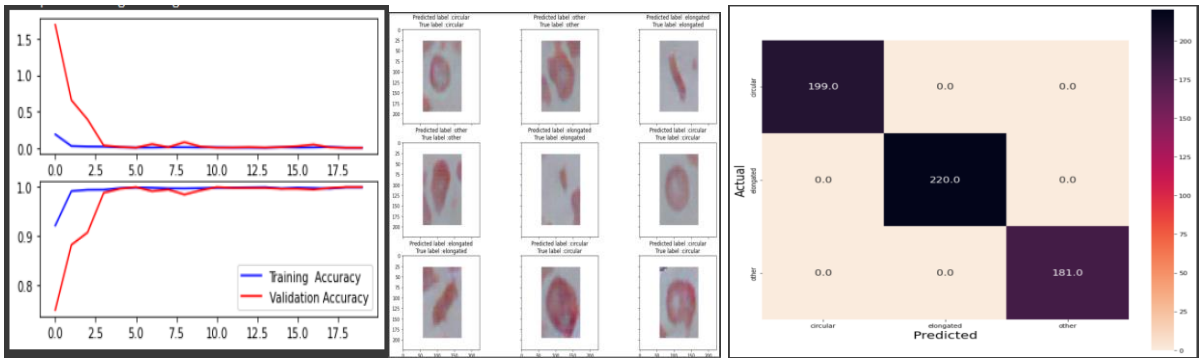


Figure 8b: Result of Modelling GAN Generated Images/Original Images Dataset

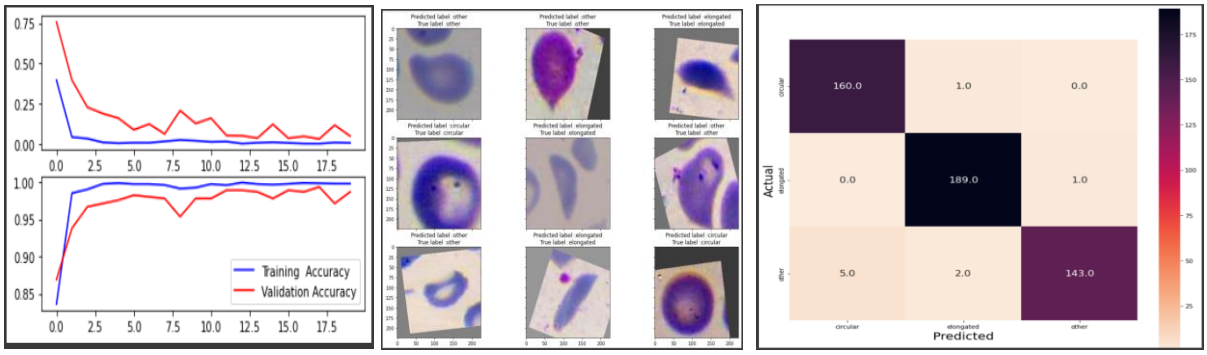


Figure 8c: Result of Modelling Augmented Data Images

Figures 8a, 8b and 8c: plot of training and validation accuracy and loss, correctly predicted images, and confusion matrix for Inception V3

6.4 Experiment 5: VGG16

Figure 9b below shows that the model performed better with GAN generated images achieving accuracy of 96.20%, with 14 misclassifications out of 600 images, while in 9c it achieved an accuracy of 84.44% with 78 misclassifications out of 501 images and 9a achieved an accuracy of 49.20% with 32 misclassifications out of 63 images

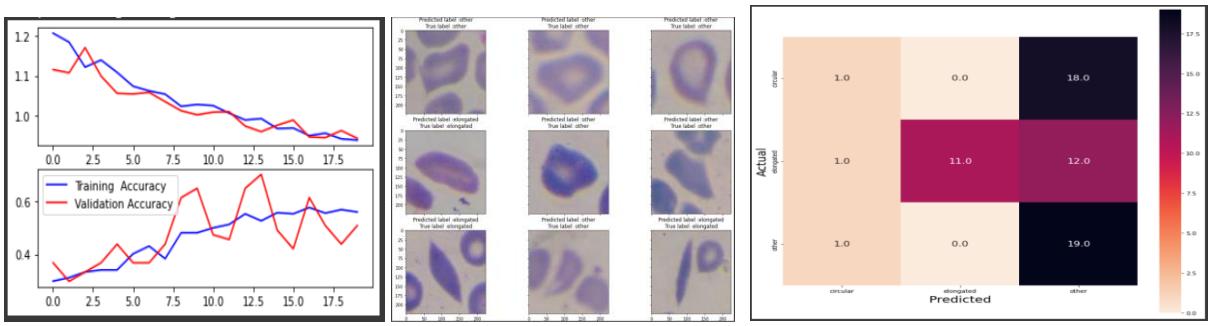


Figure 9a: Result of Modelling the Original dataset (erythrocytes)

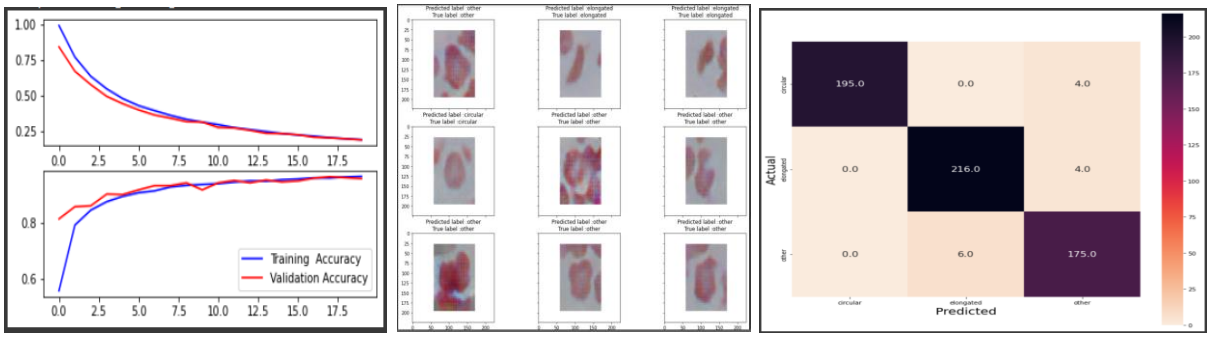


Figure 9b: Result of Modelling GAN Generated Images/Original Images Dataset

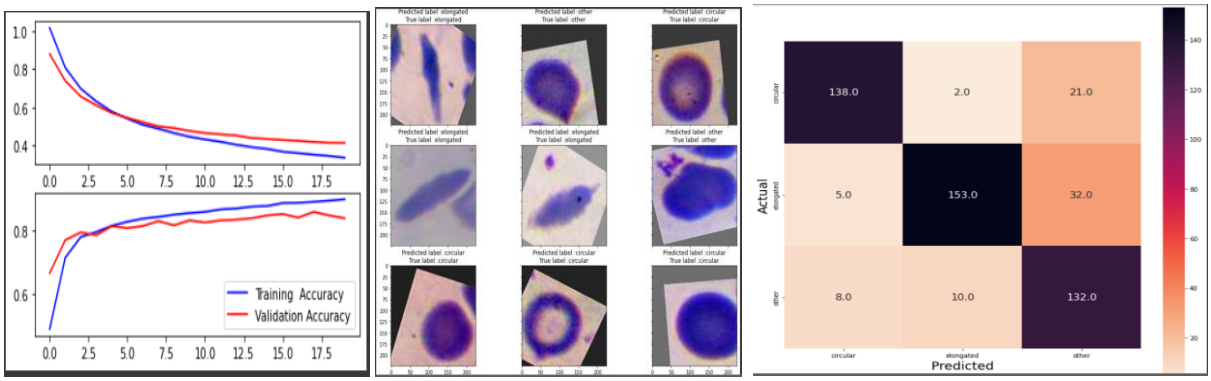


Figure 9c: Result of Modelling Augmented Data Images

Figures 9a, 9b and 9c: Plot of training and validation accuracy and loss, correctly predicted images, and confusion matrix for VGG16

6.5 Experiment 6: VGG19

Figure 10b below clearly shows that the model performed better with GAN generated images achieving accuracy of 97.30%, with only 1 misclassifications out of 600 images, while in 10c it achieved an accuracy of 92.20% with 39 misclassifications out of 501 images and 10a achieved an accuracy of 55.60% with 28 misclassifications out of 63 images.

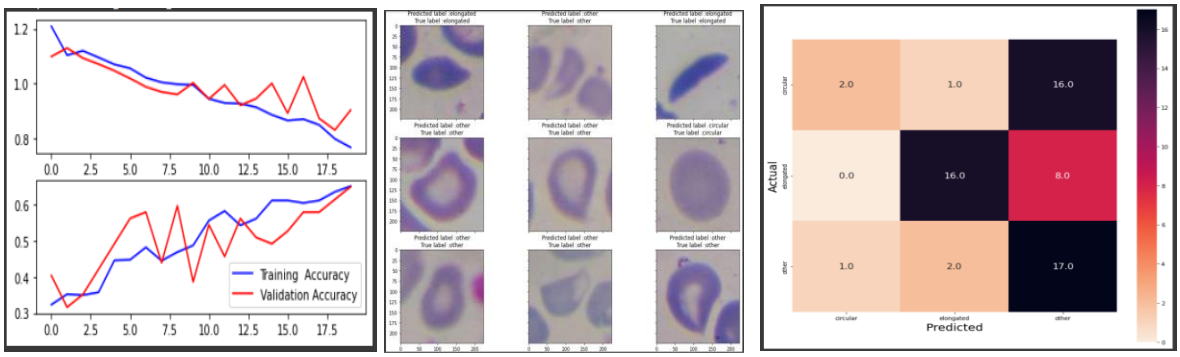


Figure 10a: Result of Modelling the Original dataset (erythrocytes)

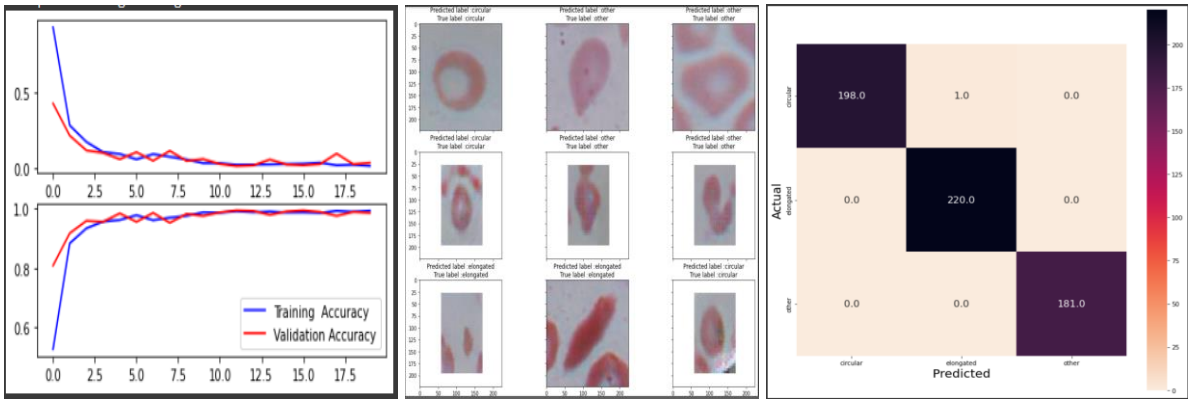


Figure 10b:Result of Modelling GAN Generated Images/Original Images Dataset

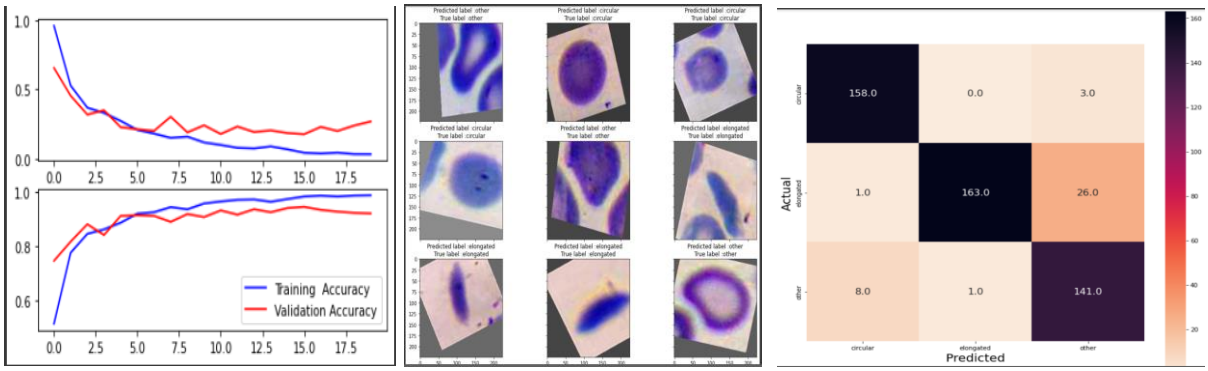


Figure 10c: Result of Modelling Augmented Data Images

Figures 10a, 10b and 10c: Plot of training and validation accuracy and loss, correctly predicted images, and confusion matrix for VGG19

6.6 Experiment 7: MobileNet

Figure 11b below clearly shows that the model performed better with Gan generated images achieving accuracy of 99.70%, with no misclassifications out of 600 images, while in 11c, it achieved an accuracy of 97.60% with 12 misclassifications out of 501 images and 11a achieved an accuracy of 61.90% with 24 misclassifications out of 63 images.

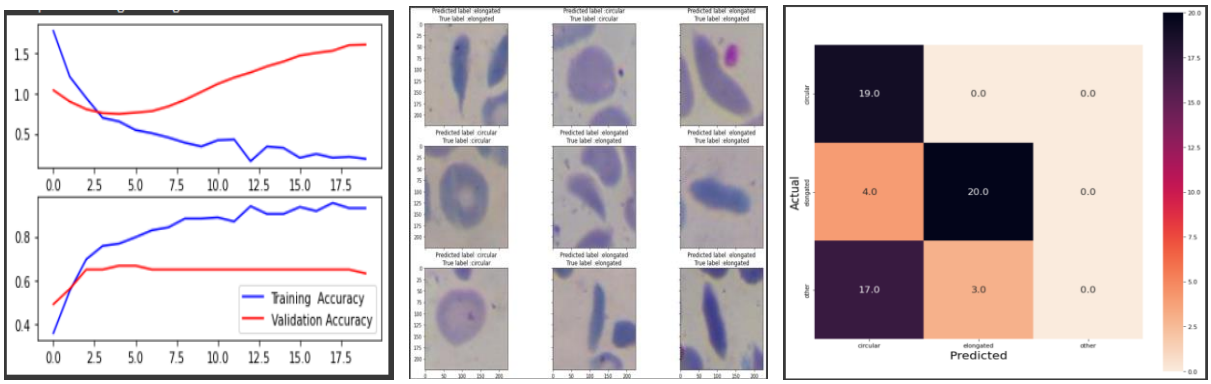


Figure 11a: Result of Modelling the Original dataset (erythrocytes)

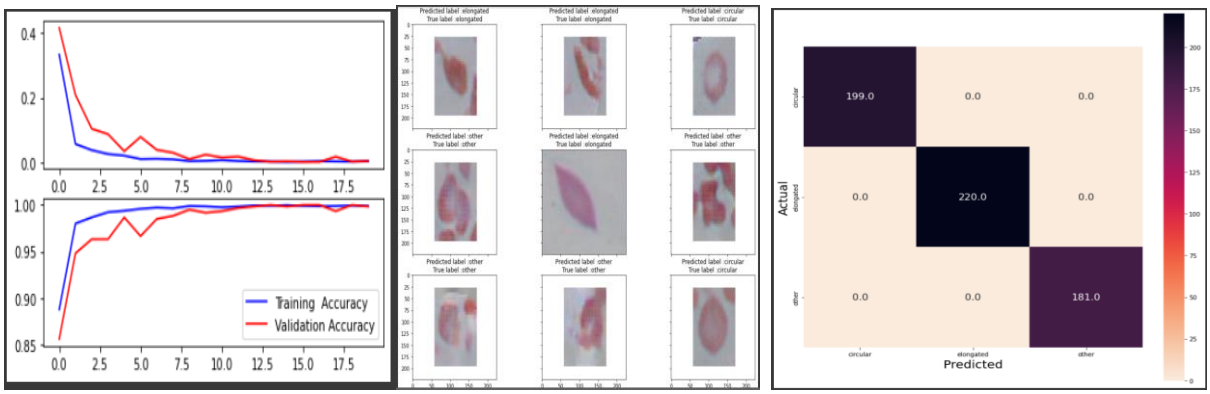


Figure11b: Result of Modelling GAN Generated Images/Original Images Dataset

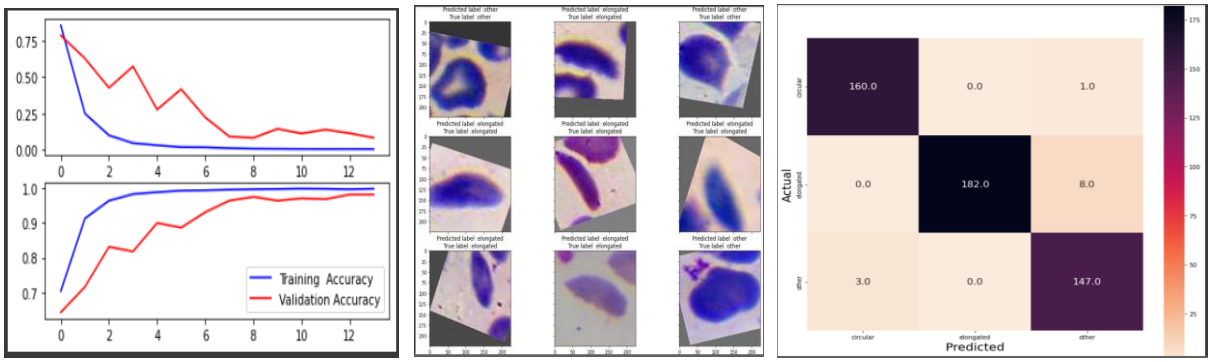


Figure 11c: Result of Modelling Augmented Data Images

Figures 11a, 11b and 11c: Plot of training and validation accuracy and loss, correctly predicted images, and confusion matrix for MobileNet

6.7 Evaluation of the Results

Table 3 below shows that two models performed considerably well in the modelling of the original dataset with Inception V3 outperforming all the models. However, all the models performed extremely well when modelled with GAN generated images and this boosted their performances between 4.5% to 136%. The biggest effect was on ResNet50 at 136%. On the traditional augmented/Original images dataset, it had almost the same impact with incremental range of between 3.2% and 80%.

Table 4: Comparison of the Performance Metrics on the Three Datasets

	Original Images				GAN Generated Images/Original Images				Original/Augmented Images			
Models	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
DenseNet121	85.70%	86.00%	84.00%	85.70%	98.90%	98.90%	99.90%	98.90%	93.80%	93.80%	93.80%	93.80%
Inception V3	95.20%	95.60%	95.20%	95.10%	99.50%	99.50%	99.60%	99.60%	98.20%	98.20%	98.20%	98.20%
MobileNet	61.90%	47.50%	61.90%	51.80%	99.70%	99.70%	99.70%	99.70%	97.60%	97.60%	97.60%	97.60%
VGG16	49.20%	60.50%	49.20%	44.20%	96.20%	96.20%	96.30%	96.20%	84.44%	85.90%	84.44%	84.70%
VGG19	55.60%	65.30%	55.60%	51.50%	97.30%	97.30%	97.30%	97.30%	92.20%	92.90%	92.20%	92.30%
ResNet50	38.10%	14.50%	38.10%	21.00%	90.00%	90.40%	90.00%	90.10%	68.70%	74.70%	68.70%	67.00%

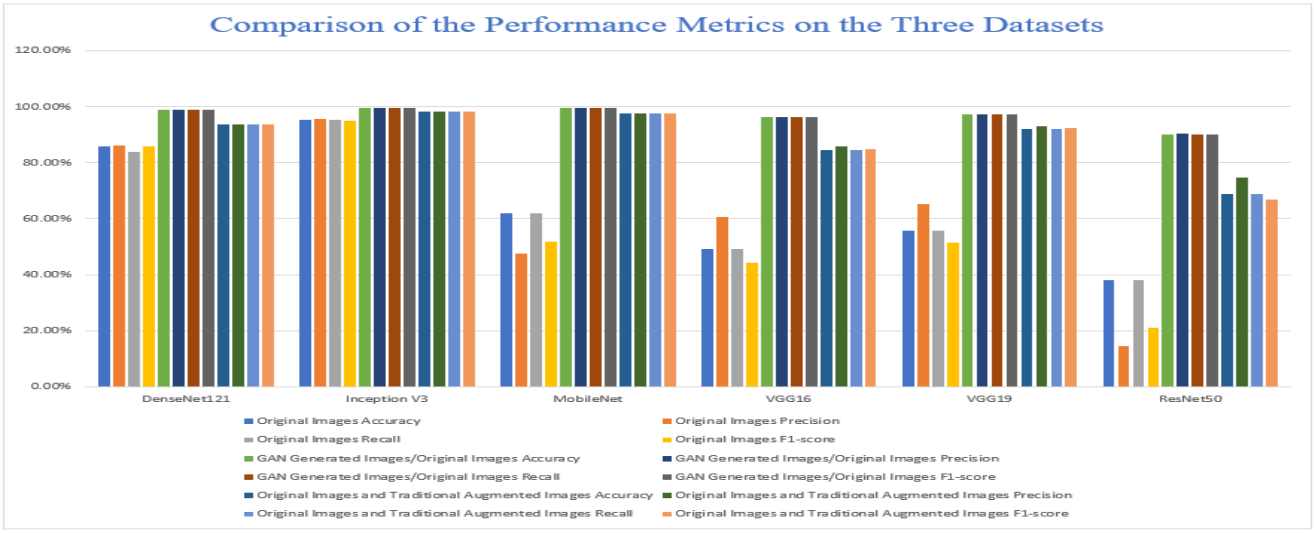


Figure 12: Bar Chart Comparison of the Performance Metrics on the three datasets

Figure 12 is the bar chart comparison of the performance metrics and it clearly showed that modelling with GAN generated images really boosted the performances of all the models with a large margin in all the performance metrics.

6.8 Discussion

Looking at the presented results from figures 6 to 12, and table 3, they clearly showed that Inception V3 and DenseNet121 models have shown very strong performance in all the performance metrics across the three datasets. The result on the original dataset produced low modelling performance for most of the models except for Inception V3 and DenseNet121. However, the improvement in the performance metric in all the models when modelled with GAN generated images/original images was phenomenal in the range of between 4.5% to 136% increase in accuracy. The biggest improvement was in ResNet50 model at 136%, followed by VGG16 at 95%, VGG19 at 75% and MobileNet at 61%. However, improvement in accuracy was also observed when modelled using augmented images/ original images in the range of 3.2% to 80%. This has shown that using GAN to generate images is a much better alternative than augmenting the images only. It shows that even though GANs does not have the ability to extrapolate unlike the traditional augmentation, it can still provide an effective way of filling in the gaps in discrete training data distribution thereby augmenting the sources of variance

that are not able to augment in other ways (Bowles et al., 2018). However, Inception V3 achieved very high-performance metrics in the three datasets showing that it is very good at detecting sickle cell disease no matter what the sample size is. Choosing the best model will depend on the business objective of the task which is to detect SCD at an early stage to prevent early death and plan management of the disease to give patients good quality of life. Although the MobileNet dropped in accuracy on the original dataset, it still achieved best performance across all metrics in GAN generated images and second-best in traditional augmented images. However, the Inception V3 and MobileNet models achieved the highest accuracy of over 99%. But looking at the accuracies alone will not help to determine the best performing model considering that this is a medical data, and any misdiagnosis can be detrimental, especially in this project where the business objective is to detect the disease very early to prevent early death and to plan management of the disease to help give patients decent quality of life. So, the emphasis is on correct classification of elongated (sickle cell) red blood cells which can be very detrimental to a sickle cell patient if misclassified and undiagnosed and the model that does not misclassify the disease will be a better model. MobileNet achieved the highest recall score of 99.70% followed by Inception V3 at 99.50% which is the sensitivity or the positive rate.

Also, consideration is given to lightweight models that has high accuracy like the MobileNet considering that there is great accessibility to mobile phone especially in less developed countries. Any model that can be deployed on a mobile device is a welcome addition and this is considered in choosing the best model. MobileNet still achieved a very high classification accuracy and recall at 99.70% and it correctly predicted sickle cell without any misclassification, and it is a lightweight model. Finally, the experiment showed that models achieved highest accuracy on GAN generated Images with great percentage margins and shows that GAN can generate good quality images and the generated images achieved an Inception Score of 1.823 out of 3 for this dataset though this can be improved on by cropping the images to contain only one type of images for the class and also increase the range of some of the traditional augmentation techniques like increasing the width and height shift and the shear range can help the models. It also demonstrates that the issue of small data size can be overcome through image synthesis using GAN which boosted the classification accuracy of all the models. MobileNet correctly classified all the sickle cell without any misclassification. This means that it can efficiently detect all the disease without any misclassification as recall and precision are 99.70%, although recall was used for evaluation as correct prediction of sickle cell is very crucial.

7 Conclusion and Future Work

The aim of this research project is to use the novel Deep Convolutional Generative Adversarial Network (DCGAN) to synthesize more images to augment the dataset for the efficient classification of sickle cell disease in early detection of the disease to overcome the challenges of small size datasets, prevent early date, give patients good quality of life and compare the results to find out if GAN generated images are realistic and can be used as an alternative source of data augmentation for classification in situation where the data size is very small, especially in the medical field. GAN and six predictive models were implemented and evaluated. This objective was achieved by using GAN to generate plausible images that looked like they were from the original dataset, and it boosted the modelling performance of all the predictive models (Bowles *et al.*, 2018, Ding et al., 2019) from between 4.5% to 136% when compared to the performance on the original images and GAN generated images. All the models also achieved performance increase between 3.2% to 80% when compared on the original images and traditional augmented images. This has addressed the research question and objectives 2, 3, and 4 that it is advantageous to generate images using GAN to overcome

the problem of small size datasets and overfitting of the models. The achieved results showed that this will help solve the problem of lack of big size datasets especially in the medical field, as the models performed better on GAN generated images than the original dataset and the traditional augmented dataset. This project achieved best and highest accuracy and recall of 99.70% with MobileNet model, while paper by (Alzubaidi et al., 2020) reported highest accuracy of 99.98%. The achieved result showed that MobileNet can effectively and efficiently predict and diagnose sickle cell in microscopic blood images without any misclassification as no misclassification of the sickle cell was obtained . This achieved result sits second behind (Alzubaidi et al., 2020) in the state-of-the-art comparison table. This has addressed research objective 5.

On the future work, the segmentation of the cells using U-Net and transfer learning models as an encoder back bones like Inception V3 will be an interesting aspect of this work as this will give a complete automatic classification thereby solving the challenges posed by overlapped cells.

Acknowledgements

I wish to thank God for the life of my younger brother who is a Sickle Cell patient for inspiring this research project, for his bravery and determination to fight for his life every day and I hope that one day he will be free from SCD based on research in this field. I am very grateful and thankful to my supervisor Mr. Jorge Basilio for all his advice, tips, and guidance throughout this research project and very thankful to my children for their amazing support throughout this period.

References

Alzubaidi, L. *et al.* (2020) 'Deep Learning Models for Classification of Red Blood Cells in Microscopy Images to Aid in Sickle Cell Anemia Diagnosis', *Electronics*, 9(3), p. 427. Available at: https://doi.org/10.3390/electronics9030427.

Bailo, O., Ham, D. and Shin, Y.M. (2019) 'Red Blood Cell Image Generation for Data Augmentation Using Conditional Generative Adversarial Networks', in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1039–1048. Available at: https://doi.org/10.1109/CVPRW.2019.00136.

Bang, D. and Shim, H. (2018) 'Improved Training of Generative Adversarial Networks Using Representative Features'. arXiv. Available at: http://arxiv.org/abs/1801.09195 (Accessed: 15 May 2022).

Bowles, C. *et al.* (2018) 'GAN Augmentation: Augmenting Training Data using Generative Adversarial Networks'. Available at: https://doi.org/10.48550/ARXIV.1810.10863.

Bressem, K.K. *et al.* (2020) 'Comparing different deep learning architectures for classification of chest radiographs', *Scientific Reports*, 10(1), p. 13590. Available at: https://doi.org/10.1038/s41598-020-70479-z.

Calimeri, F. *et al.* (2017) 'Biomedical Data Augmentation Using Generative Adversarial Neural Networks', in A. Lintas et al. (eds) *Artificial Neural Networks and Machine Learning*

– *ICANN 2017*. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 626–634. Available at: https://doi.org/10.1007/978-3-319-68612-7_71.

Chen, M. *et al.* (2021) 'Deep Feature Learning for Medical Image Analysis with Convolutional Autoencoder Neural Network', *IEEE Transactions on Big Data*, *Big Data, IEEE Transactions on*, *IEEE Trans. Big Data*, 7(4), pp. 750–758. Available at: https://doi.org/10.1109/TBDATA.2017.2717439.

Cheuque, C. *et al.* (2022) 'An Efficient Multi-Level Convolutional Neural Network Approach for White Blood Cells Classification', *Diagnostics*, 12(2), p. 248. Available at: https://doi.org/10.3390/diagnostics12020248.

de Faria, L.C., Rodrigues, L.F. and Mari, J.F., 2018, November. Cell classification using handcrafted features and bag of visual words. In *Anais do XIV Workshop de Visão Computacional*.

Ding, J. *et al.* (2019) 'A Case Study of the Augmentation and Evaluation of Training Data for Deep Learning', *Journal of Data and Information Quality (JDIQ)*, 11(4), pp. 1–22. Available at: https://doi.org/10.1145/3317573.

Elsalamony, H.A. (2016) 'Healthy and unhealthy red blood cell detection in human blood smears using neural networks', *Micron*, 83, pp. 32–41. Available at: https://doi.org/10.1016/j.micron.2016.01.008.

Gonzalez-Hidalgo, M. *et al.* (2015) 'Red Blood Cell Cluster Separation From Digital Images for Use in Sickle Cell Disease', *IEEE Journal of Biomedical and Health Informatics*, 19(4), pp. 1514–1525. Available at: https://doi.org/10.1109/JBHI.2014.2356402.

Gual-Arnau, X., Herold-García, S. and Simó, A. (2015) 'Erythrocyte shape classification using integral-geometry-based methods', *Medical & Biological Engineering & Computing*, 53(7), pp. 623–633. Available at: https://doi.org/10.1007/s11517-015-1267-x.

Howard, A.G. *et al.* (2017) 'MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications'. arXiv. Available at: http://arxiv.org/abs/1704.04861 (Accessed: 28 June 2022).

Huang, G. *et al.* (2017) 'Densely Connected Convolutional Networks', in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI: IEEE, pp. 2261–2269. Available at: https://doi.org/10.1109/CVPR.2017.243.

Li, Y. *et al.* (2022) 'Machine learning-based approaches for identifying human blood cells harboring CRISPR-mediated fetal chromatin domain ablations', *Scientific Reports*, 12(1), p. 1481. Available at: https://doi.org/10.1038/s41598-022-05575-3.

Metz, L. *et al.* (2017) 'Unrolled Generative Adversarial Networks'. arXiv. Available at: http://arxiv.org/abs/1611.02163 (Accessed: 10 August 2022).

Norouzifard, M. *et al.* (2018) 'Automated Glaucoma Diagnosis Using Deep and Transfer Learning: Proposal of a System for Clinical Testing', in *2018 International Conference on Image and Vision Computing New Zealand (IVCNZ)*. *2018 International Conference on Image*

and Vision Computing New Zealand (IVCNZ), Auckland, New Zealand: IEEE, pp. 1–6. Available at: https://doi.org/10.1109/IVCNZ.2018.8634671.

Radford, A., Metz, L. and Chintala, S. (2016) 'Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks'. arXiv. Available at: http://arxiv.org/abs/1511.06434 (Accessed: 21 May 2022).

Rodrigues, L.F., Naldi, M.C. and Mari, J.F., 2016, November. Morphological analysis and classification of erythrocytes in microscopy images. In *Proceedings of the 2016 Workshop de Visão Computacional, Campo Grande, Brazil* (pp. 9-11).

Ronneberger, O., Fischer, P. and Brox, T. (2015) 'U-Net: Convolutional Networks for Biomedical Image Segmentation', *arXiv:1505.04597 [cs]* [Preprint]. Available at: http://arxiv.org/abs/1505.04597 (Accessed: 4 February 2022).

Salimans, T. *et al.* (2016) 'Improved Techniques for Training GANs'. arXiv. Available at: http://arxiv.org/abs/1606.03498 (Accessed: 13 August 2022).

Sandler, M. *et al.* (2019) 'MobileNetV2: Inverted Residuals and Linear Bottlenecks'. arXiv. Available at: http://arxiv.org/abs/1801.04381 (Accessed: 30 June 2022).
Sen, B. *et al.* (2021) '"Deep Learning based diagnosis of sickle cell anemia in human RBC"', in *2021 2nd International Conference on Intelligent Engineering and Management (ICIEM)*. *2021 2nd International Conference on Intelligent Engineering and Management (ICIEM)*, pp. 526–529. Available at: https://doi.org/10.1109/ICIEM51511.2021.9445293.

Sharma, S. *et al.* (2022) 'Deep Learning Model for the Automatic Classification of White Blood Cells', *Computational Intelligence and Neuroscience*. Edited by S.K. Gupta, 2022, pp. 1–13. Available at: https://doi.org/10.1155/2022/7384131.

Szegedy, C. *et al.* (2015) 'Going deeper with convolutions', in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA: IEEE, pp. 1–9. Available at: https://doi.org/10.1109/CVPR.2015.7298594.

Tengshe, R. *et al.* (2021) 'Sickle cell anemia detection using convolutional neural network', in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kharagpur, India: IEEE, pp. 1–6. Available at: https://doi.org/10.1109/ICCCNT51525.2021.9580165.

Wang, J. *et al.* (2021) 'A Review of Deep Learning on Medical Image Analysis', *Mobile Networks and Applications*, 26(1), p. 351. Available at: https://doi.org/10.1007/s11036-020-01672-7.

Wieczorek, G. *et al.* (2021) 'Multiclass Image Classification Using GANs and CNN Based on Holes Drilled in Laminated Chipboard', *Sensors*, 21(23), p. 8077. Available at: https://doi.org/10.3390/s21238077.

Xu, M. *et al.* (2017) 'A deep convolutional neural network for classification of red blood cells in sickle cell anemia', *PLOS Computational Biology*. Edited by Q. Nie, 13(10), p. e1005746. Available at: https://doi.org/10.1371/journal.pcbi.1005746.

Yee, S.L.K. and Raymond, W.J.K. (2020) 'Pneumonia Diagnosis Using Chest X-ray Images and Machine Learning', in *Proceedings of the 2020 10th International Conference on Biomedical Engineering and Technology*. *ICBET 2020: 2020 10th International Conference on Biomedical Engineering and Technology*, Tokyo Japan: ACM, pp. 101–105. Available at: https://doi.org/10.1145/3397391.3397412.

Zhong, Z. *et al.* (2020) 'Cancer image classification based on DenseNet model', *Journal of Physics: Conference Series*, 1651(1), p. 012143. Available at: https://doi.org/10.1088/1742-6596/1651/1/012143.

Zhou, T. *et al.* (2022) 'Dense Convolutional Network and Its Application in Medical Image Analysis', *BioMed Research International*. Edited by C. Li, 2022, pp. 1–22. Available at: https://doi.org/10.1155/2022/2384830.

Zhu, Y. and Newsam, S. (2017) 'DenseNet for dense flow', in *2017 IEEE International Conference on Image Processing (ICIP)*. *2017 IEEE International Conference on Image Processing (ICIP)*, Beijing: IEEE, pp. 790–794. Available at: https://doi.org/10.1109/ICIP.2017.8296389.