# Award Category – Mentoring (to be filled by Teachers)

Award Eligibility - For Teachers only

Duly completed form to be submitted by August 20, 2021.

Email *

v_sowmya@cb.amrita.edu

Name of the teacher, age, institution name, course name *

Dr.V.Sowmya, Assistant Professor (Selection Grade), Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India.

2. Provide names of your 10 best mentees with their email id and mobile numbers? *

1. BhanuKumar.M - bhanu0893@gmail.com - 8374468348 2. VamsiKrishna - vamsik1995@gmail.com - 8448867184 3. M.GaneshKumar - m.ganeshdeveloper@gmail.com - 8015391529 4. Pydimarry Harsha Vardhan Pavan Kumar - cb.en.p2cen20020@cb.students.amrita.edu – 9912552285 5. AkulaHarika - akulaharika12@gmail.com – 7036854780 6. Yamini Madan – maddy1996@gmail.com – 7604880469 7. Sanjana.K – sanjanakala1996@gmail.com – 7012319593 8. Vyshnav. M.T – vyshnav94.mec@gmail.com – 7736339524 9. Ramji.B – ramji.b28@gmail.com – 9629040643 10. Harini. N. - harininarasimhan123@gmail.com - 9176002639

3. Pls categorize the mentees as per the below mentoring activities and upload documents in support of above. Professional Guidance / Career Advancement / Course Specific Work / Laboratory Specific / All – round Developments *

Career_Proof_1 - ...
Career_Proof_2 - ...
Course_Feedbac...
Course_Feedbac...
Course_Feedbac...
Professional_Gui...
Professional_Gui...
Professional_Pro...
Professional_Gui...
Research_Offer_...

Indian Institute of Technology Palakkad
भारतीय प्रौद्योगिकी संस्थान पालक्काड
Under Ministry of Human Resource Development, Govt. of India
मानव संसाधन विकास मंत्रालय के अधीन, भारत सरकार

IIT PALAKKAD

**Ref: IITPKD/2020/025/EE/MRP/001**

26 August 2020

To,
Ms. Harini Narasimhan
202/6, Golden Opulence,
Poonamallee Bypass Road,
Thirumal Nagar, Thiruvallur,
Tamil Nadu PIN: 600056

Dear Ms. Harini Narasimhan,

**Subject: Letter of Appointment as Senior Research Fellow**

On behalf of the Indian Institute of Technology Palakkad, we are pleased to offer you an appointment as a **Senior Research Fellow** in our institute in the discipline of Electrical Engineering.

As a Senior Research Fellow, you will work on the project titled "An automated lung ultrasound workflow for diagnostic assistance in COVID-19 and beyond" under the mentorship of Dr. Mahesh R. Panicker, Assistant Professor, Electrical Engineering, who would be your Reporting Faculty. We believe that this would provide you opportunities for scholastic, research and professional development.

Please take note of the following details of this offer of appointment.

1) The position is temporary and the appointment will be on contract for a period of one year starting from the date of joining the institute.
2) The monthly remuneration for this position will be Rs. 35,000/- (Thirty Five Thousand rupees only). Hostel accommodation may be provided at the institute based on availability on a chargeable basis. In case of non- availability of institute hostel, admissible HRA at 8% will be provided as per Government of India rules.
3) During your employment at IIT Palakkad, your services and conduct will be governed by the administrative orders and rules in force from time to time at the institute.

Please confirm your acceptance of this offer by replying no later than 02 September 2020.
We expect you to join at the earliest and let us know your joining date after consulting with your reporting officer and mentor, Dr. Mahesh R. Panicker. You are advised to contact him to finalize your joining date.

Yours sincerely

Prof. Vinod A. Prasad
Dean - Industry Collaboration & Sponsored Research (ICSR)
Indian Institute of Technology Palakkad

DEAN (Industry Collaboration and Sponsored Research)
Indian Institute of Technology Palakkad
Ahalia Integrated Campus,
Kozhipara, Palakkad - 678 557

# Convolution Neural Network and Generative Model based Classification in Medical Imaging

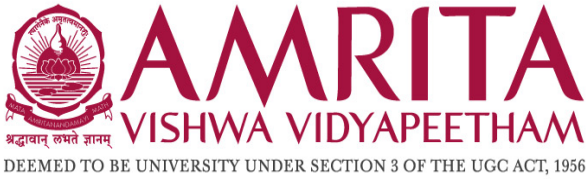
A THESIS

*Submitted by*

**Ramji Balasubramanian**
**(CB.EN.P2CEN18011)**

*in partial fulfillment for the award of the degree of*

**MASTER OF TECHNOLOGY**
**IN**
**COMPUTATIONAL ENGINEERING AND NETWORKING**




**Center for Computational Engineering and Networking**

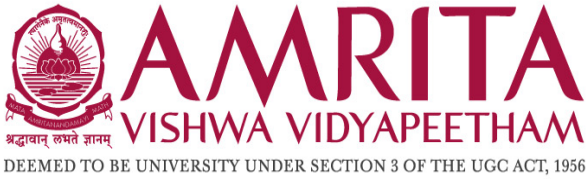
**AMRITA SCHOOL OF ENGINEERING**

**AMRITA VISHWA VIDYAPEETHAM**

COIMBATORE - 641 112 (INDIA)

**June - 2020**

# AMRITA SCHOOL OF ENGINEERING
# AMRITA VISHWA VIDYAPEETHAM
COIMBATORE - 641 112




# BONAFIDE CERTIFICATE

This is to certify that the thesis entitled **"Convolution Neural Network and Generative Model based Classification in Medical Imaging"** submitted by **Ramji Balasubramanian (Register Number- CB.EN.P2CEN18011),** for the award of the **Degree of Master of Technology** in the **"COMPUTATIONAL ENGINEERING AND NETWORKING"** is a bonafide record of the work carried out by him under our guidance and supervision at Amrita School of Engineering, Coimbatore.

Dr. Sowmya V
Project Guide
Assistant Professor (Sr.Gr.), CEN

Dr. E. A. Gopalakrishnan
Project Co-Guide
Assistant Professor (SG), CEN

Dr. K. P. Soman
Project Co-Guide
Head, CEN

Mr. Vijay Krishna Menon
Project Co-Guide
Assistant Professor (Sr.Gr.), CEN

Mr. Sajith Variyar V. V.
Project Co-Guide
Assistant Professor, CEN

*Submitted for the university examination held on* 20/06/2020

**INTERNAL EXAMINER** **EXTERNAL EXAMINER**

## AMRITA SCHOOL OF ENGINEERING

# AMRITA VISHWA VIDYAPEETHAM

COIMBATORE - 641 112

## DECLARATION

I, **Ramji Balasubramanian (CB.EN.P2CEN18011),** hereby declare that this thesis entitled **"Convolution Neural Network and Generative Model based Classification in Medical Imaging "**, is the record of the original work done by me under the guidance of **Dr.Sowmya V**, Assistant Professor, **Dr.E.A.Gopalakrishnan**, Assistant Professor, **Mr.Vijay Krishna Menon**, Assistant Professor,**Mr.Sajith Variyar V.V**, Assistant Professor and **Dr.K.P.Soman**, Professor and Head, Centre for Computational Engineering and Networking, Amrita School of Engineering, Coimbatore. To the best of my knowledge this work has not formed the basis for the award of any degree/diploma/ associateship/fellowship/or a similar award to any candidate in any University.

**Place:** Coimbatore

**Date:** 12/6/2020

**Signature of the Student**

COUNTERSIGNED

14/6/20

Dr. K.P.Soman
Professor and Head
Center for Computational Engineering and Networking

# Contents

# Acknowledgement

First of all, I thank Almighty for the immeasurable blessings for smooth completion of my project. I am immensely pleased to express my sincere obligation to my project guide **Dr. Sowmya V**, Assistant Professor (Sr.Gr.) and my project co-guides Dr. **E. A. Gopalakrishnan**, Assistant Professor (SG), **Mr. Vijay Krishna Menon**, Assistant Professor (Sr.Gr.), **Mr. Sajith Variyar V. V**, Assistant Professor and **Dr. K.P.Soman**, Professor and Head, Centre for Computational Engineering and Networking, Amrita School of Engineering, Coimbatore, for their valuable guidance, dedication and encouragement for the successful completion of this project.

I am grateful to the entire staff of CEN, and research scholars of the Department for their timely cooperation. I avail this opportunity to thank my friends for their whole-hearted support in the difficult stages of the project.

I am greatly indebted to my loving family members for being the motivating forces behind the completion of this dissertation. Also I am grateful for their invaluable help, moral support and encouragement throughout the course of study.

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| SCG | Single Cell Gel |
| MGE | MicroGel Electrophoresis |
| HOG | Histogram of Gradient |
| SVM | Support Vector Machine |
| ANN | Artificial Neural Networks |
| CNN | Convolutional Neural Network |
| DR | Diabetic Retinopathy |
| GAN | Generative Adversarial Network |
| DCGAN | Deep Convolutional Generative Adversarial Network |
| G | Generator |
| D | Discriminator |

# Abstract

The deep neural network is an emerging machine learning method that has proven its potential for different classification tasks. Notably, the Convolutional Neural Network(CNN) dominates with the best results on varying image classification tasks. CNN plays an important role to extract features from Medical images, which inturn helps in medical field to detect the diseases faster. However, medical image datasets are hard to collect because it needs a lot of professional expertise to label them. In this work, we took two different applications in medical field. One is to classify the damaged comets from undamaged comets using Comet assay images and another one is to generate retinal synthetic images using generative adversarial network (GAN) for highly imbalanced Diabetic Retinopathy(DR) disease grading. In first application, the human effort on analysing the effect of genetoxicity and DNA damage detection from microscopy examination is addressed using CNN based image processing approach. The proposed CNN based model were compared with existing models and proved to achieve a better classifcation f1-score of 0.9045. An accurate automatic classification model requires sufficient data for training. In second application analysis of classification results is carried out by augmenting highly imbalanced class in the EYEPACS dataset using generated synthetic images. We generated highly diverse images for imbalanced cases without any constraints. The generated synthetic images do not influence other class images and have also improved the classification results obtained by the model, over that which was trained without synthetic generation. The results obtained before and after augmentation by the proposed generative based model is compared over various model attributes.

# Chapter 1

# Introduction

In recent days, Deep Learning (DL) is very popular and is used for solving research problems in wide range of fields like biomedical[1], cyber security[2], autonomous vehicles[3], etc. Deep learning models are replacing traditional machine learning based models due the fact that, DL models have the ability to automatically extract the useful features from the input, while the traditional machine learning models require manual feature engineering. One of the most popular Deep learning (DL) architectures is Convolutional Neural Network (CNN). CNN is very popular in the computer vision field as it uses convolution filters in order to extract the various location invariant features. Since CNN model works well with the images, various CNN models and its variants are used in biomedical fields like x-ray reconstruction [4], mammogram detection [5], liver lesion classification[6], Brain MRI scan segmentation [7] etc.

The medical images are hard to collect, as the collecting and labeling of medical data confronted with both data privacy concerns and the requirement for time-consuming expert explanations. In the two general resolving directions, one is to collect more data, such as crowd sourcing[8] or digging into the existing clinical reports[9]. An-

other way is to increase the performance of a small dataset, which is very important because the knowledge achieved from the research can migrate to the research on big datasets. CNN-based methods have various strategies to increase the performance of image classification on small datasets using data augmentation techniques. The current work focuses on various applications of medical image classification using CNN and generative model.

## 1.1 Litereature Review

There is a wide variety of medical imaging modalities used for the purpose of clinical prognosis and diagnosis and in most cases the images look similar. This problem is solved by deep learning, where the network architecture allows learning difficult information. Hand crafted features work when expert knowledge about the field is available and generally make some strict assumptions. These assumptions may not be useful for certain tasks such as medical images. Therefore, with the hand-crafted features in some applications, it is difficult to differentiate between a healthy and non-healthy image.Applications of CNNs in medical image analysis can be traced to the 1990s, when they were used for computer-aided detection of microcalcifications in digital mammography [10],[11] and computer-aided detection of lung nodules in CT datasets[12]. With revival of CNNs owing to the development of powerful GPU computing, the medical imaging literature has witnessed a new generation of computer-aided detection systems that show superior performance.For instance, in [13], the fully connected layers of a pre-trained CNN were replaced with a new logistic layer, and then the labeled data

were used to train only the appended layer while keeping the rest of the network the same. This treatment yielded promising results for classification of unregistered multi-view mammogram. Chen et al.[14] suggested the use of a fine-tuned pre-trained CNN for localizing standard planes in ultrasound images. With the rapid increase in data and computation resources, CNN-based methods have significantly improved DR classification performance. Pratt proposed a 13- layer CNN for screening DR[15]. Recently, some researchers have also exploited GANs to synthesize retinal fundus images. Using GANs for medical image synthesis[16] could potentially address the shortage of large and diverse annotated databases. Costa et al.[17] first adopted a U-Net architecture to transfer vessel segmentation masks to fundus images using a vanilla GAN architecture. However, the generated samples have block defects and do not have controllable grading information.

# Chapter 2

# Comet Assay Damage Detection using CNN based Deep Learning Approach

## 2.1 Introduction

### 2.1.1 Overview of Comet Assay

The comet assay, also called the single cell gel assay (SCG) and microgel electrophoresis (MGE) was first introduced by Ostling and Johanson in 1984 as a microelectrophoretic technique for the direct visualization of DNA damage in individual cells. A small number of irradiated cells suspended in a thin agarose gel on a microscope slide were lysed, electrophoresed, and stained with a fluorescent DNA binding dye. The electric current pulled the charged DNA from the nucleus such that relaxed and broken DNA fragments migrated further. The resulting images as shown in Figure 1.1, which were subsequently named for their appearance as 'comets', were measured to determine the extent of DNA damage/ DNA breakage.

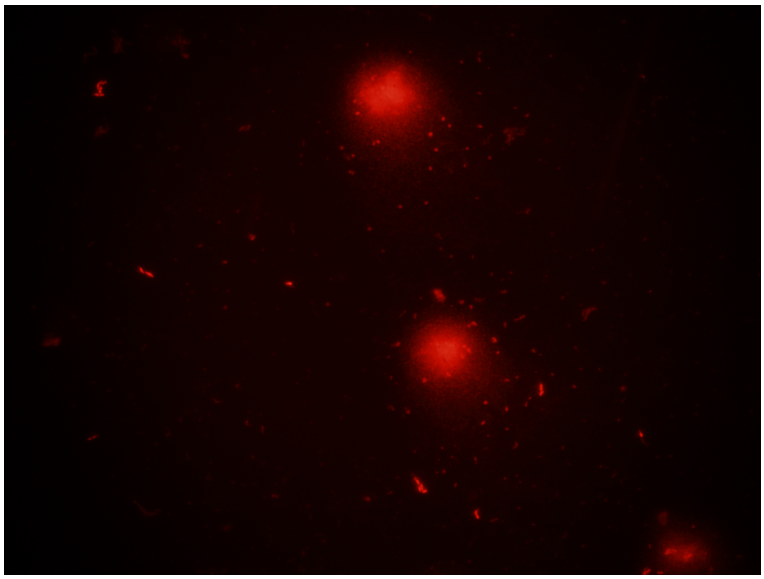


Figure 2.1: Comet Assay

### 2.1.2 Related Works

The Single cell gel electrophoresis is a method for analysing the effect of genotoxicity by DNA breakage analysis [18]. The effect of genotoxicity is measured based on comparing the damage of DNA treated with different concentrations of chemicals with the damage of DNA treated with Control(water) and Positive Control(Hydrogenperoxide). It is known that the water does not cause any damage to the DNA but Hydrogenperoxide causes higher damage[19]. The damage analysis can be performed either by visual scoring technique or image analysis techniques such as measuring parameters like tail length, tail moments of a comet [20].The visual scoring needs experts advice and also time consuming. Therefore, various software tools were developed to quantify the DNA damage of comets and to classify them based on quantification [21][22][23]. Most of the open-source software do not classify comets based on its level of damage except

hi-comet[23].

Hi-comet[23] extracts histogram of gradient (HOG) features from comet images along with box-ratio and classify comets using support vector machine(SVM), Neural Networks, AdaBoost and classification and regression tree. The SVM with linear kernel achieved best with 0.904% accuracy. In [24], the damage classes of comets are normal,abnormal and fail that are determined by measuring ratio,roundness and peak height as parameters. The classification algorithm uses decision tree algorithm which is later tested on 300 comets has achieved 86.8% accuracy.

In [25], a new approach to automate the visual scoring technique using digital image processing with machine learning is proposed. Fourteen different features like area,perimeter,min radius were extracted and classified the features into five types based on degree of damage.This algorithm uses Artificial Neural Networks with two hidden layers with relu activation function. The proposed ANN model was able to achieve 90% for some types whereas it could not meet the reliable accuracy for all types due to less number of samples.

A transfer learning based Convolutional neural network model to classify the levels of damage in comet images for tiny dataset were proposed in [26]. The transfer learning model VGG16 gives 70.5% accuracy, CNN-SVM and ordinary CNN gives 62.5% and 63.5% accuracy respectively. This approach includes few preprocessing techniques like image size normalization, noise data removal and image segmentation. A threshold value at HSV color space is used to segment the comets.Further the segmented comets were used to train the CNN based model.

### 2.1.3 Motivation

- The existing open-source software fails to classify the comets based on its level of damage except Hi- Comet.
- The existing work [23],[24],[25] and [26] needs human intervention either to extract features or setting up a threshold in segmenting the comets.

### 2.1.4 Contribution

In this work, we propose an algorithm to classify the comet assay images as damaged or undamaged using CNN based deep learning model. To do so, we experimented with empirically chose the model which gives a promising f1-score. This classification method can be used as an initial step by researchers to decide on further study on the effect of genotoxicity.

## 2.2 Convolutional Neural Network

Convolutional Neural Networks are very similar to ordinary Neural Networks from the previous chapter: they are made up of neurons that have learnable weights and biases. Each neuron receives some inputs, performs a dot product and optionally follows it with a non-linearity. The whole network still expresses a single differentiable score function: from the raw image pixels on one end to class scores at the other. And they still have a loss function (e.g. Softmax) on the last (fully-connected) layer.

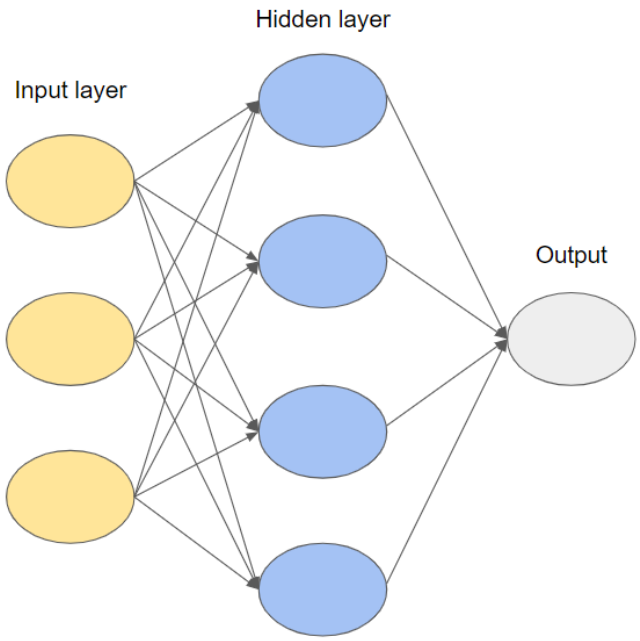



Figure 2.2: Neural Network

### 2.2.1 Neural Network

Neural Networks as shown in Figure 1.2 receive an input (a single vector), and transform it through a series of hidden layers. Each hidden layer is made up of a set of neurons, where each neuron is fully connected to all neurons in the previous layer, and where neurons in a single layer function completely independently and do not share any connections. The last fully-connected layer is called the "output layer" and in classification settings it represents the class scores.

### 2.2.2 Convolutional layer

A convolutional layer contains a set of filters whose parameters need to be learned. The height and weight of the filters are smaller than those of the input volume. Each filter

is convolved with the input volume to compute an activation map made of neurons. The output volume of the convolutional layer is obtained by stacking the activation maps of all filters along the depth dimension. Since the width and height of each filter is designed to be smaller than the input, each neuron in the activation map is only connected to a small local region of the input volume. In addition, as the activation map is obtained by performing convolution between the filter and the input, the filter parameters are shared for all local positions. The weight sharing reduces the number of parameters for efficiency of expression, efficiency of learning, and good generalization.

### 2.2.3 Leaky ReLU

Leaky Rectified Linear activation is first introduced in acoustic model[27]. Mathematically, we have

$$y_i = \begin{cases} x_i & x_i \geq 0 \\ \frac{x_i}{a_i} & x_i < 0 \end{cases}$$

where $a_i$ is a fixed parameter in range $(1, +\infty)$. In paper, the authors suggest to set $a_i$ to a large number like 100.

### 2.2.4 Dense Layer/Fully Connected Layer

Neurons in a fully connected layer have full connections to all activations in the previous layer. Their activations can hence be computed with a matrix multiplication followed by a bias offset.

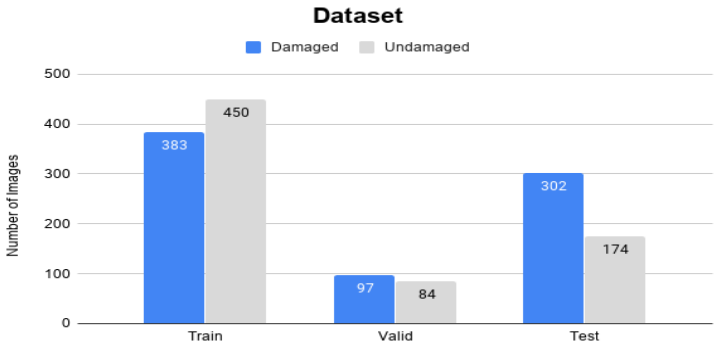



Figure 2.3: Dataset Description

## 2.3 Methodology

The Comet Assay images are annotated as damaged or undamaged using ImageJ software. The annotated images are classified with various existing works along with the proposed CNN model and the results are compared to decide on better solution.

### 2.3.1 Dataset Description

The total of 3000 comet assay images treated with Control(water), Positive Control(hydrogen peroxide) and three different concentrations of diethyl phthalate (1ngml,10mgml and 100ngml) were collected and annotated with the help of experts for damaged and undamaged comets using ImageJ software. From the annotated images 383 damaged comets were taken from Positive control and 450 undamaged comets from control are used for training and the remaining damaged and undamaged comet images of cells treated with control and positive control were used as validation data. The test data includes comets images cropped from different concentrations of diethyl phthalate. The data set is splitted as shown in Figure 1.3.

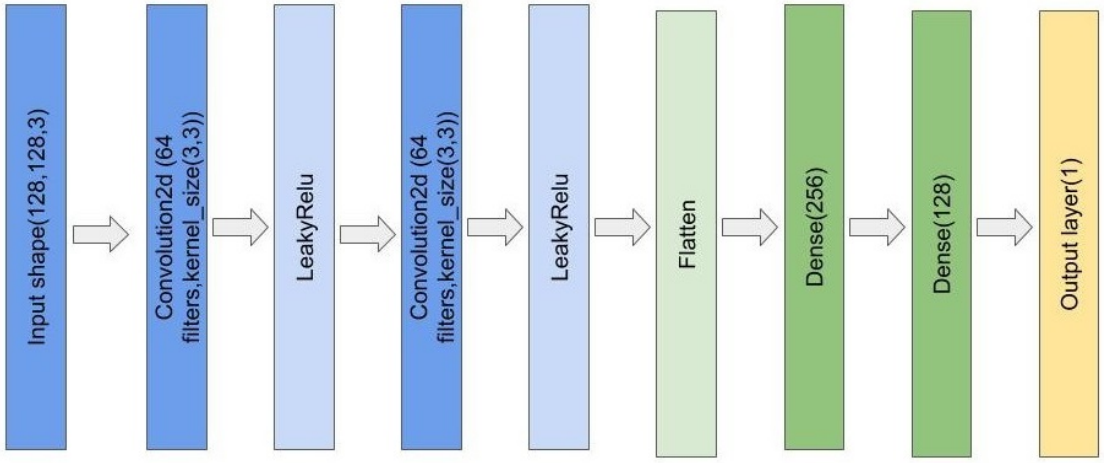



Figure 2.4: Proposed Architecture

### 2.3.2 Classification

The comets are classified as damaged or undamaged comet. In this module,we propose an algorithm to classify the comet images using CNN based deep learning model as shown in Figure 1.4 by training only the DNA comets of cells treated with control and positive control. Since cells treated with control and positive control is required for damage analysis of genotoxins, training the model with them makes it a generalized approach. The proposed CNN model includes two convolutional layers and two dense layers with LeakyRelu activation between the layers. The choice of proposed model is based on empirical tuning of hyperparameters. The ground truth of class labels are provided by the domain experts. The performance of the various experimented models were compared based on f1-score and the model with highest f1-score has been chosen as the proposed model to classify the comets.

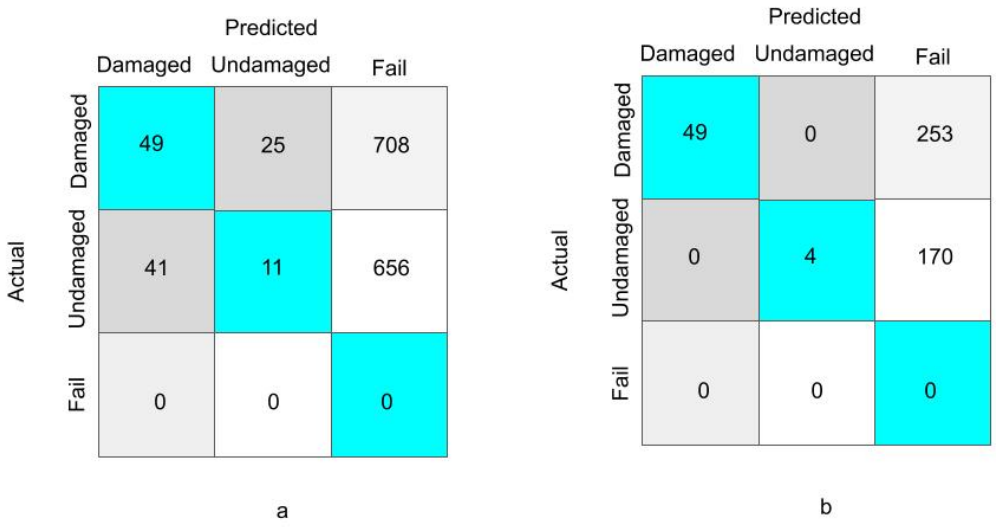



Figure 2.5: Confusion matrix a) whole data b) test data

## 2.4 Experiments and Results

### 2.4.1 Decision Tree Algorithm

As explained in paper [24] ratio,roundness and peak height features were extracted for all the 1490 images. The box ratio of an image is calculated as the ratio of width and height of the image. The correlation coefficient between the comet image and an oval shape image is defined as roundness. Peak Height is defined as the relative intensity of the nucleus over the highest intensity column from the image. Since we do not have any training process for this algorithm, the performance of the model on the complete annotated dataset and test data alone is performed and shown as confusion matrix in Figure 1.5.

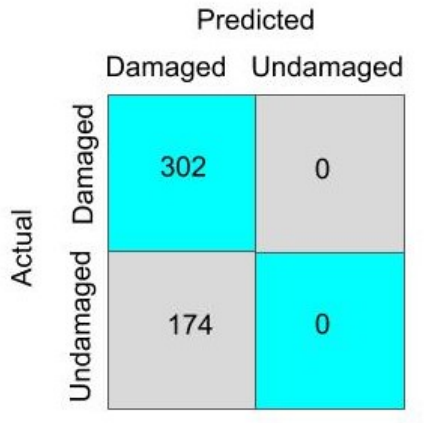



Figure 2.6: Confusion matrix for ANN model

### 2.4.2 Artificial Neural Network

The Artificial neural network proposed in [25] has 2 hidden layers each with 12 and 8 neurons respectively were trained to classify the five different types from the extracted features. The model is tested on comet images and the obtained result is shown in Figure 1.6.

### 2.4.3 Support Vector Machine

The histogram of gradient (HOG) features are used in computer vision object detection task. The HOG features will extract the edge features along with its directions and store as a descriptor. These features are extracted as mentioned in [23] along with its box ratio (width and height ratio) and classified using Support Vector Machine. The confusion matrix for SVM model is shown in Figure 1.7.

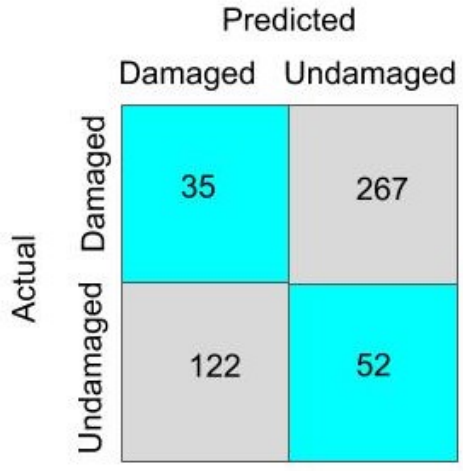



Figure 2.7: Confusion matrix for SVM model

### 2.4.4 Proposed Method

All the existing algorithms mentioned above does not meet the significant result. So, we came up with an algorithm using CNN based deep learning approach to classify the comet images as damaged and undamaged. The preprocessing of the comet images includes resize to (128,128,3), normalization and histogram equalization. The hyper parameters such as number of layers in the network, number of filters in the convolution layers and the number of neurons in the dense layers of the model are empirically tuned with fixed batch size of 32 to get better results and their performances are compared based on f1 score as shown in Table 1.1. The chosen model is then experimented with data augmentation such as horizontal flip and vertical flip which improved the f1 score from 0.8405 to 0.9045.

The performance of the model is evaluated on 302 damaged comets and 174 undam-

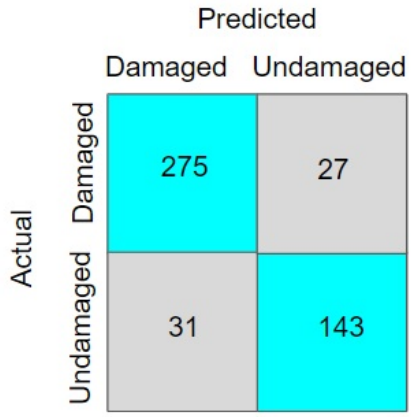



Figure 2.8: Confusion matrix for proposed model

aged comets. The Confusion matrix which summarizes the model is shown in Figure 1.8. Further the precision,recall and f1-score for the proposed model is 0.8986, 0.9106 and 0.9045.

Table 2.1: Models with f1-score

| **Model** | **F1-score on test data** |
|---|---|
| Conv(64) + Dense(256) | 0.8135 |
| Conv(64) + Dense(256) + Dense(128) | 0.8355 |
| **2 * Conv(64) + Dense(256) + Dense(128) (Proposed Model)** | **0.8405** |
| 2 * Conv(64) + Maxpool + Dense(256) + Dense(128) | 0.8402 |
| 2 * Conv(64) + Maxpool + Conv(128) + Dense(256) + Dense(128) | 0.7977 |
| 2 * Conv(64) + Maxpool + Conv(128) + Dense(512) + Dense(256) | 0.8303 |
| 2 * Conv(64) + Maxpool + Conv(128) + Conv(64) + Maxpool + Dense(512) + Dense(256) | 0.8342 |
| 2 * Conv(64) + Maxpool + Conv(128) + Maxpool + 2 * Dense(256) | 0.8217 |
| 2 * Conv(32) + Maxpool + Conv(64) + Maxpool + Dense(256) | 0.8379 |
| 2 * Conv(32) + Maxpool + 2* Conv(64) + Maxpool + Dense(256) | 0.8305 |
| 2 * Conv(32) + Maxpool + 2*Conv(64) + Maxpool + Dense(256) + Dense(64) | 0.8297 |

## 2.5 Acknowledgement

I acknowledge Dr. Bindhu Paul,Assistant Professor, Nano science Department, Amrita Institute of Medical Science, Cochin for the access to the comet assay data set on analysis of compound diethyl phthalate and for help in manual annotation of the data set.

# Chapter 3

# Analysis of Adversarial based Augmentation for Diabetic Retinopathy Disease Grading

## 3.1 Introduction

### 3.1.1 Overview of Diabetic Retinopathy

Diabetic Retinopathy (DR) is a disease which causes blindness to diabetic patients and the number of such cases are increasing dramatically. Study shows early detection and treatment of DR can restore 90% of the lost vision effectively; it is recommended to regularly screen patients. DR is graded into five classes: normal(No DR), mild, moderate, severe and proliferative DR as shown in Figure 2.1[28]. The grading is a time consuming and tedious process. Whereas with the advancement in technology a huge difference can be made with automating the grading process[29][30][15].

### 3.1.2 Related Works

Seoud et al. extracted 35 different features from fundus images and classified them using random forest [29]. The first step towards automatic grading by Seoud et al showed

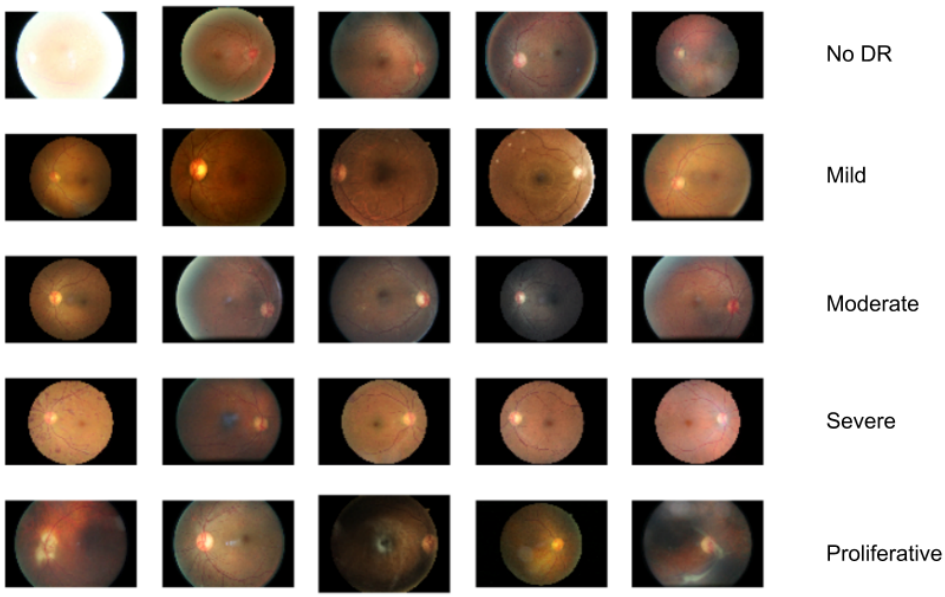



Figure 3.1: Sample Images for DR grading

a comparable result to that achieved by the domain experts. The deep learning era has made the classification even better with deep convolutional layers. Several works like [31],[32] and [33] use a deep Convolutional Neural Network(CNN) for classification and gave state of the art accuracy to the problem. Training CNN networks requires a diverse and huge volume of data. The collection of such data is very hard due to privacy laws and medico legal complications; the data is not available. This might be the reason the EYEPACS dataset[34] is highly imbalanced and biased towards the normal class. Statistically too, observing a disease is analogous to observing a defect or a fault; both are significantly less likely events contributing to imbalanced datasets. In literature, the use of many augmentation techniques in CNN, to overcome the imbalance in the datasets,are detailed in [35],[36],[37] and [38]. We have employed some of these traditional augmentation techniques namely rotation, flipping and cropping.

Though this overcomes the imbalance problem to some extent, the augmented images are not diverse enough limiting the model performance [39]. So, synthetically generated images are used to create diversity within the dataset.. Generative Adversarial Networks (GAN) [40], introduced in the year 2014, is widely being used to generate synthetic images. GANs have the ability to address the shortage of images and have been successfully used in medical image analysis involving chest X-Rays [41], Computerized Tomography scans[42], and Magnetic Resonance images[17]. Recently, Costa et al. [17] generated fundus images using image-to-image translation technique, but the generated images have lesser information about grading. Zhao et al proposed Tubs-GAN [43], to extend style transfer and increase the diversity of generated images. In [44], Zhou et al generated synthetic retinal fundus images by conditioning the structural and lesion masks. Though the generated images showed improvement in accuracy, the conditioned lesion masks make the model more complicated and the generation of images for normal is not necessary since it already has enough data to train the model. Kaplan et al [45], generated a diverse fundus image using an unconditioned DCGAN model, but the model failed to capture the vessel trees of the fundus images clearly.

### 3.1.3 Contribution

In this paper, we aim to generate synthetic images for proliferative class and balance it with the severe class which has the second-lowest number of images. The generated and its original train images are trained, and compared the class-wise accuracy of test data between before and after the synthetic generation. To make sure that the

generated images are from the proliferative class, the average cosine similarity for the real proliferative class was compared with the generated images. The diversity of images generated among the classes is also analysed. The proposed model serves as a base version on improving class accuracy using generated images. Here, we aim to improve the classification results of specifically the proliferative class (as it contains the least number of data samples) by augmenting with proliferative generated images. The novel contribution in this paper includes the following.

- Unconditional GAN model to generate DR images, to reduce the complexity of image generation.
- Diversity of intra class image generation,specifically for proliferative cases.

## 3.2 Genarative Adversarial Network

In 2014, Goodfellow et al.[40] proposed an adversarial network framework as an alternative generative model estimation process for deep generative networks, called GAN. The main principle of a GAN is that there are two neural networks called the generator (G) and the discriminator (D) that compete to maximize their gains. The main goal for both networks is to improve their capabilities to generate and discriminate the data. In this context, G draws samples from a random noise distribution and D discriminates whether the samples are drawn from G or from real data (the training data).

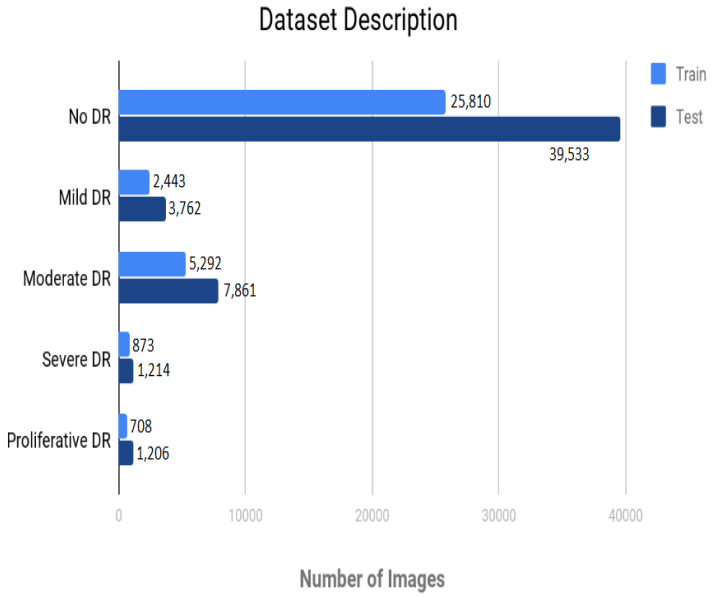



Figure 3.2: EYEPACS Dataset used for the Proposed Adversarial based Data Augmentation for DR Grading.

## 3.3 Methodology

The images collected from the Kaggle EYEPACS dataset [34] are processed before feeding into a GAN model. The GAN model generates a new set of synthetic images for proliferative class and the best batch is selected based on the cosine distance measure between the original images and the generated images. The generated images with original images are used to train the classification model.

### 3.3.1 Dataset Description

The open-source Kaggle EYEPACS dataset[34] includes high-resolution retinal fundus images collected under different imaging conditions. The dataset is split into train and test images, containing 35,126 and 53,576 retinal images respectively. The five classes

of DR with the number of images are represented in Figure 2.2 and it is evident that the proliferative class is highly imbalanced with 708 images in the train set.

### 3.3.2 Preprocessing techniques

Among the whole dataset, images are collected and preprocessed in such a way that extra black regions are removed. Furthermore, the images are resized to 128 x 128 x 3 and also normalized with the maximum value (255). These preprocessed images are used for generation and classification models.

### 3.3.3 Retinal Synthetic Image Generation

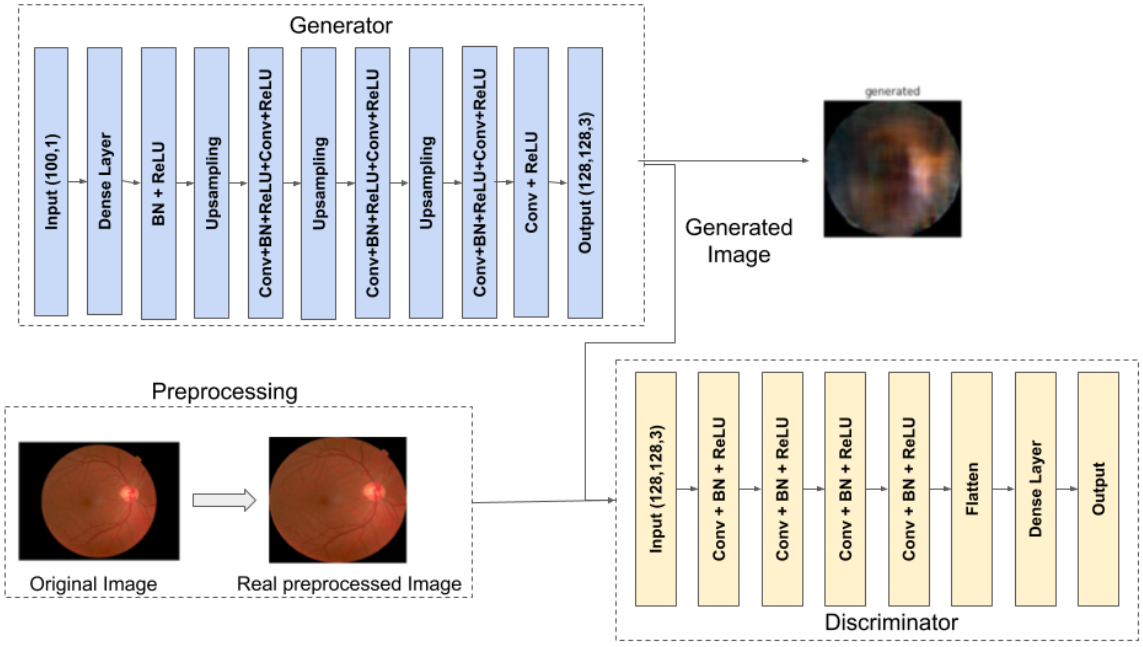



Figure 3.3: Proposed DCGAN Model for the Image Generation used for DR Grading

The preprocessed proliferative images are fed into the discriminator of the Deep Convolution Generative Adversarial Network (DCGAN) model inspired from [45] with

few changes in convolutional layers to get the required output shape as shown in Figure 2.3 is trained for 500 epochs. The Synthetic retinal images (165 images) are generated by the generator of the DCGAN model for every 10 epochs. The reason behind the number of images in each generated batch is 165 because this is the difference between the number of images in proliferative class (which has the least data samples (708)) and the Severe DR class, which has the second least data samples (873 images). The diversity of the generated images are analyzed using cosine distance. For 50 different generated image batches (each batch with 165 images), the difference between average cosine distance for raw proliferative images (708 retinal images) and the mixture of raw and generated proliferative images (708 + 165 generated images) are calculated. The batch with the minimum difference is chosen for data augmentation.

### 3.3.4 Classifier

The classifier model [35] shown in Figure 2.4, is selected with cross-entropy as a loss function, as the model resulted in the benchmark kappa score for the original EYEPACS dataset. The experimentation includes training the chosen classifier model with and without synthetically generated images. The generated proliferative images are chosen based on the best cosine similarity index. Both the above-trained models i.e., trained with and without augmentation are validated using the test data.

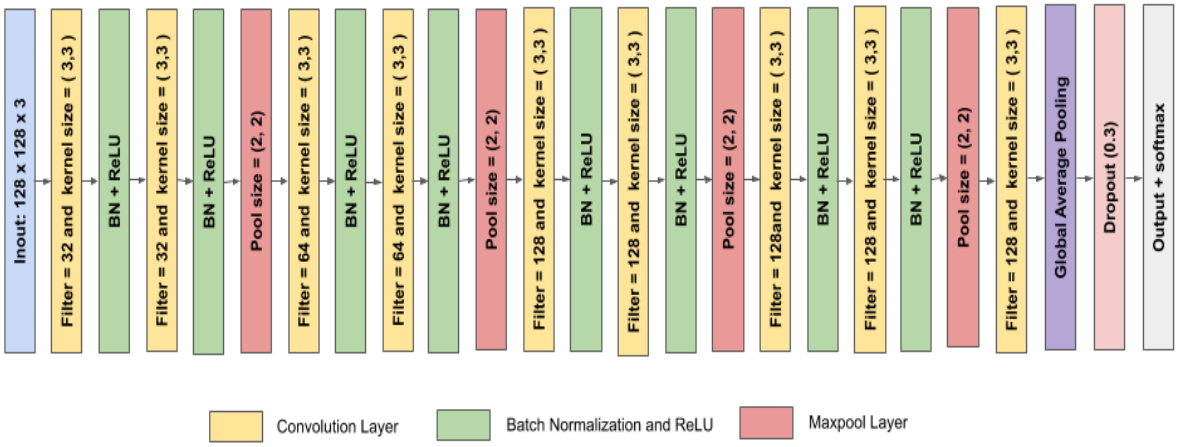



Figure 3.4: Classifier Model for validating the DR grading with and without data augmentation using generated images

## 3.4 Experiements and Results

The synthetic proliferative image generation is a crucial step in this work. The DCGAN model is trained for 500 epochs with learning rate 1e-4 and different experiments are performed by varying the batch size (shown in Table 2.1). The randomly generated latent samples of vector size 100 x 1, are fed into the generator to generate 165 random synthetic images. The cosine similarity measure is used to find the best batch size generated by each model. Each DCGAN model with different batch sizes, generates 50 different batches (each with 165 synthetic images), as the generation is performed for every 10 epochs.

Since InceptionV3 is one of the predominant models used to attain the benchmark result in DR classification[30], the feature vector of size 2048 x 1 is extracted from the same. This model is pre-trained with Imagenet images and the final global average pooling is chosen to extract the feature vector. The cosine distance between the ex-

tracted feature vectors for 708 proliferative images is calculated (708 x 708) and the average cosine distance ($A_{original}$) is computed as given below.

$$Average\ Cosine\ distance = \frac{1}{N - N_d} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}$$

*where*,

$a_{ij}$ - *Cosine distance matrix of size (n x n)*

$N$ - *Number of elements in* $a_{ij}$

$N_d$ - *Number of diagonal elements in* $a_{ij}$

Similarly, the average cosine distance ($A_{augmented}$) is calculated for shuffled original proliferative images and synthetic images. The average cosine distance ($A_{original}$) is 0.244492. From all the experimented models, the model with batch size 4 has produced the $A_{augmented}$ closer to the obtained $A_{original}$ as shown in Figure 2.5. The minimum difference between $A_{original}$ and $A_{augmented}$ is the best batch among the generated batches for each model. The same was repeated for different models i.e., the model with the different batch size, and the computed cosine similarity index are tabulated in Table 2.1.

From Table 2.1, it is clear that the synthetic images generated by the model with batch size 4 are more similar to the original proliferative images. Further evaluation on choosing the right batch size is carried by comparing the classification results (given in Table 2.2) obtained by the generated images from all three models.

The classifier model is trained for 100 epochs with learning rate 1e-4 and batch size

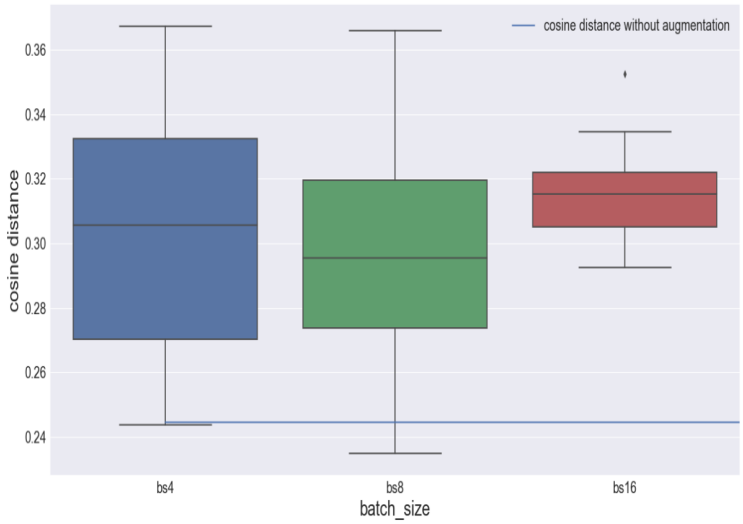



Figure 3.5: Cosine distance between the augmented images using the proposed DCGAN for different batch size and the original images present in the proliferative class

Table 3.1: Average Cosine Distance with a varying batch size of the proposed DCGAN model used for DR image generation

| **Batch Size** | **Average Cosine Distance ($A_{original}$)** | **Average Cosine Distance ($A_{augmented}$)** | **Difference ($A_{original}$ -$A_{augmented}$)** |
|---|---|---|---|
| 4 | 0.244492 | 0.243563 | **0.000929** |
| 8 | 0.244492 | 0.234786 | 0.009706 |
| 16 | 0.244492 | 0.292456 | 0.047964 |

Table 3.2: Classification result for DR grading with and without augmentation

| **Experiments** | **Batch Size While generating GAN** | **Difference ( $A_{original}$ -$A_{augmented}$)** | **Class** | **Score** | | |
|---|---|---|---|---|---|---|
| | | | | **Precision** | **Recall** | **F1-score** |
| Without Augmentation | Nil | Nil | 0 | 0.78 | 0.87 | 0.82 |
| | | | 1 | 0.10 | 0.03 | 0.05 |
| | | | 2 | 0.27 | 0.25 | 0.26 |
| | | | 3 | 0.17 | 0.07 | 0.10 |
| | | | 4 | 0.28 | 0.07 | 0.11 |
| With Augmentation | **4** | **0.000929** | 0 | 0.77 | 0.89 | 0.83 |
| | | | 1 | 0.10 | 0.03 | 0.04 |
| | | | 2 | 0.26 | 0.21 | 0.23 |
| | | | 3 | 0.16 | 0.09 | 0.11 |
| | | | **4** | **0.32** | **0.08** | **0.13** |
| | 8 | 0.009706 | 0 | 0.77 | 0.89 | 0.83 |
| | | | 1 | 0.09 | 0.05 | 0.07 |
| | | | 2 | 0.28 | 0.17 | 0.21 |
| | | | 3 | 0.17 | 0.12 | 0.14 |
| | | | 4 | 0.26 | 0.08 | 0.12 |
| | 16 | 0.047964 | 0 | 0.78 | 0.88 | 0.82 |
| | | | 1 | 0.09 | 0.04 | 0.05 |
| | | | 2 | 0.28 | 0.21 | 0.24 |
| | | | 3 | 0.15 | 0.06 | 0.08 |
| | | | 4 | 0.25 | 0.12 | **0.17** |

15 with the original EYEPACS dataset and tested. The same classifier model is trained by augmenting the existing proliferative images with generated synthetic images and the test results are compared, which is shown in Table 2.2.

From Table 2.2, the precision score of the classifier model for proliferative class (class 4) is higher when the dataset is augmented by the DCGAN model with batch size 4. Whereas, the generated images by DCGAN model with batch size 16 has achieved the best F1 score and recall score. It is important that this augmentation should not impact the class wise accuracies of other classes (0,1,2 and 3), as it is only targeted to improve the performance of the proliferated class. The metrics of class 0 and class 1 have no major difference in all the experiments. The F1 score of class 2 is better for the DCGAN model with batch size 16 whereas for class 3, batch size 4 is better. Though the classifier model with batch size 16 performs better classification, the average cosine distance

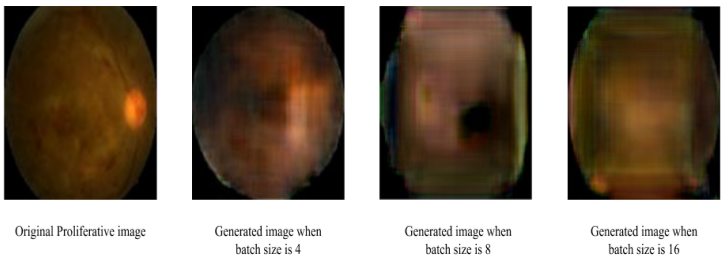



Figure 3.6: Original Proliferative and generated synthetic images when batch size is 4,8 and 16

between the original proliferative images and the augmented proliferative images is higher. This implies that the generated images are not close enough with the original images, when compared with the DCGAN model (with batch size 4), which produces lower average cosine distance and better F1 score. To visually support that DCGAN model with batch size 4, the generated images are shown in Figure 2.6. It is observed from Figure 5 that the images generated by the DCGAN model with batch size 4 are closer to the original image.

Each generated image is compared with all the original proliferative images in the dataset using cosine distance. The minimum cosine distance gives information about the generated image being similar to an original image. It is found that the generated images are similar to different varieties of original images within the same class. Hence, the generated model captured the intraclass variation.

# Chapter 4

# Conclusion and Future Work

In this work, the importance of CNN in medical image classification is proved using two different applications.

- **Comet Assay Damage Detection using CNN based Deep Learning Approach:** The comet classification using CNN based deep learning approach is experimented and achieved a promising result with 0.9045 as f1-score. The work can further be improved on hypertuning the parameters to increase the f1-score. This may help researchers to go for further analysis on the effect of genotoxicity based on number of correctly classified damaged comets. The proposed model is also incorporated in a software tool and validated using different test dataset.

- **Analysis of Adversarial based Augmentation for Diabetic Retinopathy Disease Grading:** Data imbalance is handled by the synthetic generation of images for the classification of Diabetic retinopathy grading from fundus images. Furthermore, a novel method on analysing the efficiency of generated images is also proposed based on the similarity between the original images and augmented

images (original and generated images). This paper limits the improvement of the classification accuracy by balancing only the highly imbalanced class (proliferative) as a base version model. The extension of this base work to all other classes can improve the classification results. The possible future scope of the present work includes the hyper-tuning of the DCGAN model used to generate synthetic images and as well as the classification model.

# Bibliography


[1] MA Anupama, V Sowmya, and KP Soman. "Breast Cancer Classification using Capsule Network with Preprocessed Histology Images". In: *2019 International Conference on Communication and Signal Processing (ICCSP)*. IEEE. 2019, pp. 0143–0147.

[2] R Vinayakumar et al. "Application of Deep Learning Architectures for Cyber Security". In: *Cybersecurity and Secure Information Systems*. Springer, 2019, pp. 125–160.

[3] N Deepika and VV Sajith Variyar. "Obstacle classification and detection for vision based navigation for autonomous driving". In: *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE. 2017, pp. 2092–2097.

[4] Eunhee Kang, Junhong Min, and Jong Chul Ye. "A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction". In: *Medical physics* 44.10 (2017), e360–e375.

[5] Wener Borges Sampaio et al. "Detection of masses in mammogram images using CNN, geostatistic functions and SVM". In: *Computers in Biology and Medicine* 41.8 (2011), pp. 653–664.

[6] Maayan Frid-Adar et al. "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification". In: *Neurocomputing* 321 (2018), pp. 321–331.

[7] Fausto Milletari et al. "Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound". In: *Computer Vision and Image Understanding* 164 (2017), pp. 92–102.

[8] Danail Stoyanov et al. *Intravascular Imaging and Computer Assisted Stenting and Large-scale Annotation of Biomedical Data and Expert Label Synthesis: 7th Joint International Workshop, CVII-STENT 2018 and Third International Workshop, LABELS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings*. Vol. 11043. Springer, 2018.

[9] Xiaosong Wang et al. “Tienet: Text-image embedding network for common thorax disease classification and reporting in chest x-rays”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 9049–9058.

[10] Wei Zhang et al. “Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network”. In: *Medical physics* 21.4 (1994), pp. 517–524.

[11] Heang-Ping Chan et al. “Computer-aided detection of mammographic microcalcifications: Pattern recognition with an artificial neural network”. In: *Medical physics* 22.10 (1995), pp. 1555–1567.

[12] S-CB Lo et al. “Artificial convolution neural network techniques and applications for lung nodule detection”. In: *IEEE transactions on medical imaging* 14.4 (1995), pp. 711–718.

[13] Gustavo Carneiro, Jacinto Nascimento, and Andrew P Bradley. “Unregistered multiview mammogram analysis with pre-trained deep learning models”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2015, pp. 652–660.

[14] Hao Chen et al. “Standard plane localization in fetal ultrasound via domain transferred deep neural networks”. In: *IEEE journal of biomedical and health informatics* 19.5 (2015), pp. 1627–1636.

[15] Harry Pratt et al. “Convolutional neural networks for diabetic retinopathy”. In: *Procedia Computer Science* 90 (2016), pp. 200–205.

[16] Salome Kazeminia et al. “GANs for medical image analysis”. In: *arXiv preprint arXiv:1809.06222* (2018).

[17] Pedro Costa et al. “Towards adversarial retinal image synthesis”. In: *arXiv preprint arXiv:1701.08974* (2017).

[18] Daryl W Fairbairn, Peggy L Olive, and Kim L O’Neill. “The comet assay: a comprehensive review”. In: *Mutation Research/Reviews in Genetic Toxicology* 339.1 (1995), pp. 37–59.

[19] Natacha Driessens et al. “Hydrogen peroxide induces DNA single-and double-strand breaks in thyroid cells and is therefore a potential mutagen for this organ”. In: *Endocrine-related cancer* 16.3 (2009), p. 845.

[20] Andrew R Collins. “The comet assay for DNA damage and repair”. In: *Molecular biotechnology* 26.3 (2004), p. 249.

[21] Krzysztof Końca et al. “A cross-platform public domain PC image-analysis program for the comet assay”. In: *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* 534.1-2 (2003), pp. 15–20.

[22] Benjamin M Gyori et al. “OpenComet: an automated tool for comet assay image analysis”. In: *Redox biology* 2 (2014), pp. 457–465.

[23] Taehoon Lee et al. “HiComet: a high-throughput comet analysis tool for large-scale DNA damage assessment”. In: *BMC bioinformatics* 19.1 (2018), p. 44.

[24] Taehoon Lee et al. “Robust classification of DNA damage patterns in single cell gel electrophoresis”. In: *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2013, pp. 3666–3669.

[25] Mirayr RQ de Souza et al. “A Framework for Automatic Recognition of Cell Damage on Microscopic Images using Artificial Neural Networks”. In: *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. 2018, pp. 636–639.

[26] Edgar Anarossi et al. “Comet assay classification for buccal Mucosa’s DNA damage measurement with super tiny dataset using transfer learning”. In: *Asian Conference on Intelligent Information and Database Systems*. Springer. 2019, pp. 279–289.

[27] Andrew L Maas, Awni Y Hannun, and Andrew Y Ng. “Rectifier nonlinearities improve neural network acoustic models”. In: *Proc. icml*. Vol. 30. 1. 2013, p. 3.

[28] CP Wilkinson et al. “Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales”. In: *Ophthalmology* 110.9 (2003), pp. 1677–1682.

[29] Lama Seoud, Jihed Chelbi, and Farida Cheriet. “Automatic grading of diabetic retinopathy on a public database”. In: (2015).

[30] Varun Gulshan et al. “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs”. In: *Jama* 316.22 (2016), pp. 2402–2410.

[31] Karen Simonyan and Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition”. In: *arXiv preprint arXiv:1409.1556* (2014).

[32] Christian Szegedy et al. “Rethinking the inception architecture for computer vision”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 2818–2826.

[33] Deepika Jaswal, S Vishvanathan, and S Kp. “Image classification using convolutional neural networks”. In: *International Journal of Scientific and Engineering Research* 5.6 (2014), pp. 1661–1668.

[34] ““Kaggle diabetic retinopathy detection competition,” https://www.kaggle.com/c/diabetic-retinopathy-detection.” In: ().

[35] Jordi de La Torre, Domenec Puig, and Aida Valls. “Weighted kappa loss function for multi-class classification of ordinal data in deep learning”. In: *Pattern Recognition Letters* 105 (2018), pp. 144–154.

[36] Sheikh Muhammad Saiful Islam, Md Mahedi Hasan, and Sohaib Abdullah. "Deep learning based early detection and grading of diabetic retinopathy using retinal fundus images". In: *arXiv preprint arXiv:1812.10595* (2018).

[37] Xiaogang Li et al. "Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification". In: *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. IEEE. 2017, pp. 1–11.

[38] Jiaxi Gao, Cyril Leung, and Chunyan Miao. "Diabetic retinopathy classification using an efficient convolutional neural network". In: *2019 IEEE International Conference on Agents (ICA)*. IEEE. 2019, pp. 80–85.

[39] Yi Zhou et al. "DR-GAN: Conditional Generative Adversarial Network for Fine-Grained Lesion Synthesis on Diabetic Retinopathy Images". In: *arXiv preprint arXiv:1912.04670* (2019).

[40] Ian Goodfellow et al. "Generative adversarial nets". In: *Advances in neural information processing systems*. 2014, pp. 2672–2680.

[41] Debangshu Bhattacharya et al. "GAN-Based Novel Approach for Data Augmentation with Improved Disease Classification". In: *Advancement of Machine Intelligence in Interactive Medical Image Analysis*. Springer, 2020, pp. 229–239.

[42] Maayan Frid-Adar et al. "Synthetic data augmentation using GAN for improved liver lesion classification". In: *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. IEEE. 2018, pp. 289–293.

[43] He Zhao et al. "Synthesizing retinal and neuronal images with generative adversarial nets". In: *Medical image analysis* 49 (2018), pp. 14–26.

[44] Gao Huang et al. "Densely connected convolutional networks". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 4700–4708.

[45] Sinan Kaplan et al. "Evaluation of Unconditioned Deep Generative Synthesis of Retinal Images". In: *International Conference on Advanced Concepts for Intelligent Vision Systems*. Springer. 2020, pp. 262–273.

# DNA damage analysis tool

- DNA damage analysis includes three different modules,
  - **Module 1:** Detecting the valid comets from comet assay image
  - **Module 2:** Classifying the valid comets as damaged or undamaged
  - **Module 3:** Quantifying the damaged comets
- **Module 1 and 3** were developed by my fellow friends **Mr. Vyshnav M T** and **Ms. Harini N**
- All three modules were incorporated into a single tool and a web based beta version application is created for end users to analyse the DNA damage from comet assay image.

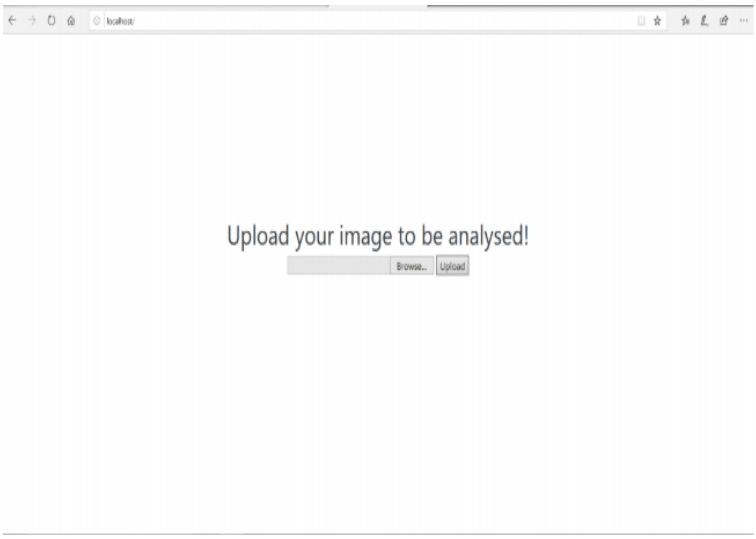



Figure 4.1: Home page

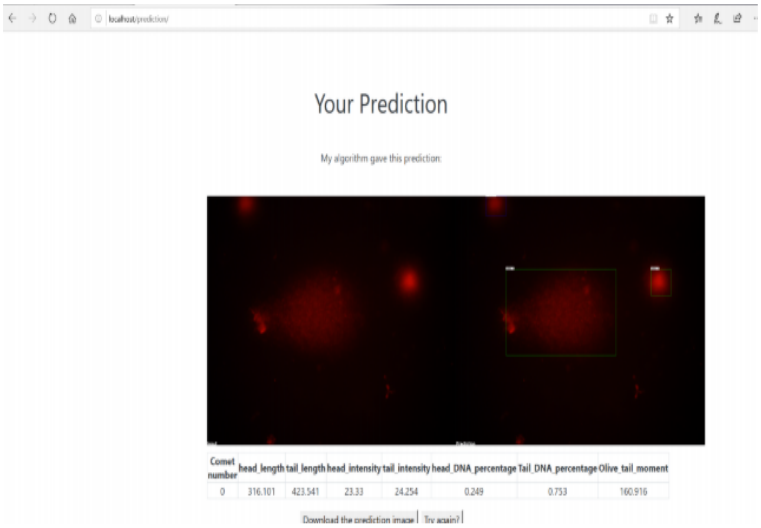



Figure 4.2: Result page

# List of Publications based on this research work

1. Ramji Balasubramanian, Sowmya V, Gopalakrishnan E.A, Vijay Krishna Menon, Sajith Variyar V. V, Soman K.P,"Analysis of Adversarial based Augmentation for Diabetic Retinopathy Disease Grading", INTERNATIONAL CONFERENCE ON COMPUTING, COMMUNICATION AND NETWORKING TECHNOLOGIES (ICCCNT-2020),IEEE (Accepted)

# Explainable AI for Heart Rate Variability in ECG Signal

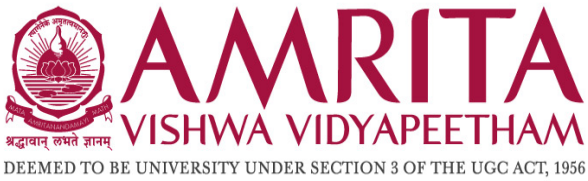
A THESIS

*Submitted by*

**Sanjana K**
**(CB.EN.P2CEN18014)**

*in partial fulfillment for the award of the degree of*

**MASTER OF TECHNOLOGY**
**IN**
**COMPUTATIONAL ENGINEERING AND NETWORKING/REMOTE SENSING AND WIRELESS SENSOR NETWORKS**




**Center for Computational Engineering and Networking**

**AMRITA SCHOOL OF ENGINEERING**

**AMRITA VISHWA VIDYAPEETHAM**

COIMBATORE - 641 112 (INDIA)

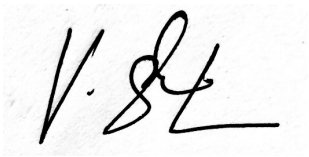
**June - 2020**

# AMRITA SCHOOL OF ENGINEERING
# AMRITA VISHWA VIDYAPEETHAM
COIMBATORE - 641 112

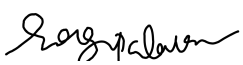
AMRITA VISHWA VIDYAPEETHAM
श्रद्धावान् लभते ज्ञानम्
DEEMED TO BE UNIVERSITY UNDER SECTION 3 OF THE UGC ACT, 1956

# BONAFIDE CERTIFICATE

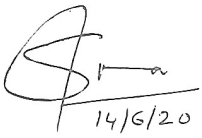
This is to certify that the thesis entitled **"Explainable AI For Heart Rate Variability in ECG Signal"** submitted by **Sanjana K (Register Number – CB.EN.P2CEN18014)**, for the award of the **Degree of Master of Technology** in the **" COMPUTATIONAL ENGINEERING AND NETWORKING/REMOTE SENSING AND WIRELESS SENSOR NETWORKS"** is a bonafide record of the work carried out by her under our guidance and supervision at Amrita School of Engineering, Coimbatore.



**Name of the Guide**
Dr.Sowmya.V
Assistant Professor



**Name of the Co-Guide**
Dr.GopalaKrishnan.E.A
Assistant Professor




**Name of the Co-Guide**
Dr.Soman.K.P
Professor and Head

*Submitted for the university examination held on 20 June 2020*

**INTERNAL EXAMINER** **EXTERNAL EXAMINER**

## AMRITA SCHOOL OF ENGINEERING

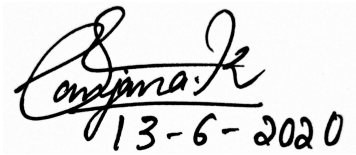
# AMRITA VISHWA VIDYAPEETHAM

COIMBATORE - 641 112

## DECLARATION

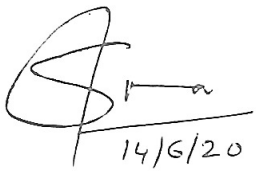
I, **Sanjana K (CB.EN.P2CEN18014),** hereby declare that this thesis entitled **"Explainable AI for Heart Rate Variability in ECG Signal"**, is the record of the original work done by me under the guidance of **Dr.Sowmya.V**, Assistant professor, **Dr.Gopalakrishnan.E.A**, Assitant professor and **Dr.Soman.K.P**, Professor and Head, Centre for Computational Engineering and Networking, Amrita School of Engineering, Coimbatore. To the best of my knowledge this work has not formed the basis for the award of any degree/diploma/ associateship/fellowship/or a similar award to any candidate in any University.

**Place:Amrita** **Signature of the Student**




**Date:13-6-2020**

COUNTERSIGNED




Dr. K.P.Soman
Professor and Head
Center for Computational Engineering and Networking

# Contents

# Acknowledgement

First of all, I thank Almighty for the immeasurable blessings for smooth completion of my project. I am immensely pleased to express my sincere obligation to my project guide Dr.Sowmya.V Assistant Professor, Computational Engineer- ing and Networking, Amrita Vishwa Vidyapeetham, Coimbatore, for her valuable guidance, dedication and encouragement for the successful completion of this project. It is my genuine pleasure to thank my project co-guide Dr.Gopalakrishnan.E.A Assistant Professor, Computational Engineering and Networking, Amrita Vishwa Vidyapeetham, Coimbatore, for his continuous motivation and scholarly advice for the successful completion of this project.

I thank profusely all the research scholars,Computational Engineering and Networking,Amrita Vishwa Vidyapeetham, Coimbatore, for their kind help throughout my project work. I am grateful to our Head, Dr. K.P.Soman and the Project Co-ordinator, Dr. E.A.Gopalakrishnan and the entire staff of CEN for their timely cooperation.

I avail this opportunity to thank my friends for their whole-hearted support during the project. I am greatly indebted to my loving parents for their invaluable help, moral support and encouragement throughout the course of study. This accomplishment would not have been possible without them.

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| AF | Atrial Fibrillation |
| AUC | Area Under Curve |
| CNN | Convolutional Neural Network |
| ECG | Electrocardiogram |
| GRU | Gated Recurrent Unit |
| LSTM | Long Short Term Memory |
| ML | Machine Learning |
| ROC | Receiver Operating Characteristic Curve |
| RSCNN | Residual Skip Convolutional Neural Network |
| RNN | Residual Neural Network |

# Abstract

ECG signal is one of the most reliable methods to analyze the cardiovascular system. The main challenge lies in the analysis and categorization of different cardiac diseases. Cardiac diseases can generally be classified into tachycardia and bradycardia. In literature, there are different deep learning architectures to detect various types of tachycardia diseases such as atrial fibrillation, ventricular fibrillation, and sinus tachycardia. Even though all types of tachycardia diseases have fast beat rhythm as the common characteristic feature, existing deep learning architectures are trained with the corresponding disease-specific features. Hence, the objective of the paper is to explore the features learned by the model from the ECG signal for the detection of different tachycardia diseases. For the detection of the different types of tachycardia diseases, we use a transfer learning approach in which, the model is trained with one of the tachycardia diseases and tested with all other tachycardia diseases such as ventricular fibrillation and sinus tachycardia. The experiments are conducted using standard deep learning architectures such as recurrent neural networks, long short-term memory, gated recurrent unit, convolution neural network, and residual skip convolution neural network. The results are analyzed using standard metrics called accuracy and sensitivity. The most important factor that limits detection is the availability of instances of the diseased condition. In the case of a low resource ECG signal data, deep learning cannot be implemented. We address the problem of cardiac disease detection with less number of samples from an ECG signal using a Chebfun. The Chebfun helps to retrieve the signal information by signal approximation without making use of conventional handcrafted features. With this feature of Chebfun, the study aims to give an interpretation of the features learned by the model for cardiac disease prediction. The performance of the Chebyshev features for cardiovascular disease detection is evaluated using the ICBHI

2019 scientific challenge cardiovascular disease dataset and atrial fibrillation dataset. The input to the classifier was the Chebfun coefficient which is extracted from the raw signal. The evaluation metrics such as accuracy, precision, recall, and specificity are used for measuring the performance of the proposed Chebfun features.

# Chapter 1

# Introduction

Cardiovascular disease (CVD) is one that affects the heart and blood vessels. The CVDs include coronary heart disease, rheumatic heart disease, etc. [1]. Risk of the CVDs increase due to blood clots that are caused by the buildup of fat deposits in the coronary arteries. According to the study conducted by WHO, an estimated 17.9 million people died due to CVDs in 2016 i.e 31% of all deaths worldwide [2]. The CVD in a broader sense can be categorized into electrical disorder, circulatory disorder, and structural disorder [3]. The electrical disorder is caused due to the malfunction of the electrical system that synchronizes the heartbeat (e.g. arrhythmia). The circulatory disorder is caused due to the high blood pressure and block in the coronary artery (e.g. stroke or heart attack). The structural disorder is caused due to the damage in the heart muscle or heart valves (e.g. cardiomyopathy).

Most of the people might have experienced irregular heart rhythms at some point in their life. Arrhythmia is developed when there is an abnormality in electrical impulse formation or transformation or abnormality in both [7]. Some of the arrhythmias are a threat to life [3]. When the heart beats are slower than the normal beats (<50bpm)

it is called bradycardia or bradyarrhythmia. In such cases, the blood pressure cannot be controlled and the patient will faint which leads to death. Similarly, when the heart beats faster than the normal beats (>100bpm) it is called tachycardia or tachyarrhythmia. This may lead to pass out and sudden death [4]. As arrhythmias are one of the main causes of mortality, detection of the arrhythmias at the early stage has acquired great importance in recent years. Tachycardia and bradycardia can be classified into different types based on their origin. The different types of tachycardia include ventricular fibrillation (VF), long QT syndrome, premature ventricular contractions, atrial flutter (AFL), supraventricular tachycardia, atrial fibrillation (AFib), sinus tachycardia (ST) and Wolff-Parkinson-White syndrome [8]. The different types of bradycardia include sinus bradycardia (SB), sinus pause or sinus arrest, sick sinus syndrome, etc [9, 10].

As the risk of heart disease is high, the detection of disease must be accurate. There are different techniques that prevailed in the detection of coronary heart diseases. Some of the techniques are an electrocardiogram (ECG), Holter monitoring, echocardiogram, stress test, cardiac catheterization, cardiac computerized tomography (CT) scan, and cardiac magnetic resonance imaging (MRI) [11]. Among the above-mentioned techniques, ECG based analysis is the most commonly used practice to diagnose cardiac disease. An ECG signal is a record of electrical communication of the heart. ECG signal monitoring is a non-invasive technique. ECG signals are recorded by placing small electrodes in the legs, arms, and chest. Cardiac disease is detected through the analyses of variation in the morphology of the ECG signal. The characteristic feature

of a normal ECG signal during one cardiac cycle is the P-wave followed by the QRS complex continued by a T-wave [12, 13]. The sample of the normal ECG signal is shown in Fig 1.1 The intervals between the waves P-QRS-T varies when the person is affected

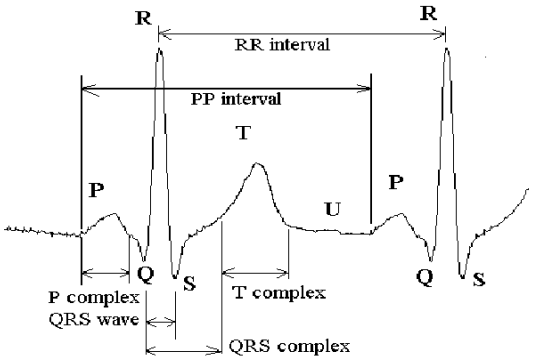



Figure 1.1: Normal ECG Signal [14]

by the disease. The variation of the ECG signal based on the characteristic shape and interval helps the experts in disease diagnosis. However, analysis of the ECG is a complex procedure because the experts should consider various factors such as age, gender, previous health condition, etc. Along with this, the number of patients a doctor would see during a day is also very high and so, it is also prone to error. To make ease of this task, an automatic expert system to diagnose the cardiac disease is preferable.

Automation in the expert system aims to make an intelligent system that can automatically detect disease. Advancement in the field of artificial intelligence made automation in expert systems possible [15]. Conventional methods require feature extraction that is specific for the disease from the raw signal. The model should be fed with optimal data. The model trained with less amount of data shows lesser performances due to overfitting. A deep learning-based model in contradiction to the machine

learning model can learn the required feature by itself [5]. In the case of cardiac disease, each of the tachycardia and bradycardia disease contains features, which have disease-specific variability. Fast beating rhythm is common for all the tachycardia diseases. In case of the atrial fibrillation (AF) and ventricular fibrillation (VF), the distinct feature that makes them distinct from another tachycardia disease i.e. sinus tachycardia (ST) dataset is the absence of the P-wave. The VF has the presence of the fibrillatory waves in the QRS baseline [8, 9, 10]. The deep learning model is expected to study different disease-specific features for the detection of different diseases. Explainable AI is the field that has gained more popularity recently. Explainable AI helps to interpret information learned by deep learning or machine learning models. It tries to interpret the reason for the decision made in the black box of neurons. This interpretability helps to improve performance in various fields of artificial intelligence. In the disease classification problem, we may not exactly know what the model learns. In the case of tachycardia disease, it is expected to learn the fast beat rhythm. The different types of tachycardia diseases such as AF and VF which do not contain specific P-wave segments are different from ST which have distinct P-wave segments. Other features in the ECG signal that make AF different from the other tachycardia dataset is the presence of the fibrillatory waves in the baseline of the QRS complex. VF also has the fibrillatory waves in the baseline of ECG. But more commonly, the ECG of VF is irregular and all segments (P, QRS, and ST) are distorted. One distinct feature used for the identification of ST is the presence of positive upright P-wave before the QRS complex. Other segments such as QRS and ST have normal morphology [30, 31, 32]. The patterns of cardiac diseases

are drastically changing nowadays. ECG is the most reliable method for the detection of these variations in the diseases. The distinct features of the cardiovascular diseases are captured in the QRS complex, P-wave, and T-wave components of the ECG signal. The ECG is the record of the electrical activity of the heart and thus ECG could capture all anomalies that can happen to the cardiac activities. Thus leaves the trace of the abnormal activities in its ECG structure.In most medical scenarios the availability of samples of the diseased condition is less thus we will not get enough data samples for the proper training of the model. In such cases in which there is less availability of data samples, there is a requirement of capturing the proper disease-specific features that are best suitable for the detection of the diseases. The proposed method aims to interpret the features learned by the deep learning model for the detection of different cardiac diseases. In the case of the ECG dataset with low resources, the study aims to find a suitable feature for the better detection of cardiac disease and to interpret the features learned by the machine learning model.

## 1.1 Literature Survey

Arrhythmia is the most dangerous cardiac disease which can be life-threatening because of its abnormal heart rate. Various studies have been conducted in this field for the detection of different types of arrhythmia, which is generally classified as tachycardia and bradycardia. Rajendra et.al [6] proposed an automatic system to classify different segments of an ECG signal. The proposed method used a convolution neural network which classified the data into four classes namely atrial fibrillation, ventricular

fibrillation, atrial flutter and normal. The model was able to achieve an accuracy of 92.50%, sensitivity of 98.09% and specificity of 93.13%. Wang et.al [16] performed a novel short time multi-fractional approach to classify atrial fibrillation, ventricular fibrillation, and ventricular tachycardia. With a fuzzy Kohonen classifier, the proposed method achieved an accuracy higher than 97%. Martis et.al [17] used the discrete cosine transform together with independent component analysis (ICA) as a dimensionality reduction approach. K-nearest neighbor algorithm based classifier has been used to classify diseases such as atrial fibrillation and atrial flutter from normal ECG beats. The method acquired an accuracy of 99.45%. A higher-order spectra method was proposed by Martis et.al [18] for rectifying the problem due to high nonlinearity in the ECG signal and compared two higher-order methods for classification of the 3 diseases namely atrial fibrillation, atrial flutter and normal. This method obtained an accuracy of 97.65% and a predictive value of 99.53%. Khadra et.al [19] used higher-order bi-spectral analysis for classification of arrhythmias such as Atrial fibrillation (AF), Ventricular fibrillation (VF) and Ventricular tachycardia (VT) with respect to normal (NR) ECG. Sensitivity values of 91.7%, 81.8%, 83.3%, and 100 % were obtained for VF, VT, AF, and NR respectively. Li et.al [20] used a support vector machine-based method for the classification of VF and VT. The proposed method achieved an accuracy of 96.3%.

Sujadevi et.al [26] proposed recurrent neural network-based atrial fibrillation detection. The work made use of architectures such as a recurrent neural network (RNN), long short term memory (LSTM) and gated recurrent unit (GRU) for the real-time

detection of AF which gained accuracies of 95%, 100%, and 100% respectively. Kiranyaz et.al [27] proposed a 1-dimensional convolution neural network (CNN) based adaptive method for individual specific ECG signal classification. The method was able to show reliable performance in the classification of ventricular ectopic beats and supraventricular ectopic beats. Kachuee et.al [23] used deep learning architecture for the classification of five different classes of arrhythmia and the approach gained an accuracy of 93.4%. Further authors have used the transfer learning approach because of the less availability of data. Transferred knowledge from the classification of arrhythmia is used to classify ECG signals with and without myocardial infarction with an accuracy of 95.9%. Gopika et.al [28] further showed an improved accuracy from 95.9% to 99% using the features proposed by the Kachuee et.al [23]. From the literature, it is evident that there are various approaches used for the efficient classification of different types of tachycardia disease. The different types of tachycardia diseases are atrial fibrillation, ventricular fibrillation, and sinus tachycardia (AF, VF, and ST) which have the fast beat rhythm as a common feature. In the previous works, even though AF, VF, and ST have a common feature, the models are trained with disease-specific ECG signals for detection of the above mentioned different types of tachycardia diseases. The previous studies lack the interpretation of the features learned by the model. The proposed work aims to interpret the features learned by the model for cardiac disease detection.

Various studies have been conducted in the field of cardiac disease detection, particularly in the detection of atrial fibrillation. The 2017 PhysioNet CinC challenge was conducted to classify the atrial fibrillation (AF) data from normal and other noisy

data. Many teams took part in the challenge in which, four teams had shown equal and better score of 0.83 than other teams. Teijeiro et.al [38] used abductive interpretation to derive the handcrafted features and the sequence information from the ECG signal. The features are fed into the embedded recurrent neural network for the classification. Datta et.al [39] proposed a two-layer cascaded binary classifier for the classification of normal, AF and other rhythms. The real classification happens in the second layer of binary classifier before that the records are intermediately classified to normal and others, and AF and noisy in the second layer. Zabihi et.al [40] used a random forest classifier to classify 150 features chosen from 491 handcrafted features into normal, AF and other rhythms. Hong et.al [41] combined the features extracted using the deep neural networks and extracted features from significant waves for the classification of normal, AF, noisy, and other signals. They used an ensemble classifier for the classification. Kropf et.al [44] proposed a method combining the machine learning and conventional signal analysis approach. They analyzed the different architectures and shown that a gradient boosted tree model is better than the random forest for AF classification. Rizwan et.al [45] used a machine-learning algorithm to classify the feature extracted samples from the ECG signal. Together with the dimensionality reduction technique and sparse coding they used an ensemble decision tree as the classifier. Kamaleswaran et.al [42] proposed a 13 layer deep convolutional neural network-based approach for the classification of the normal, AF, noisy, and other signals. The method also includes the effect of the influence of hyperparameters on model performance. Plesinger et.al [43] proposed an ensemble of convolution neural network and bagged tree for the classifica-

tion of Holter ECG signal into four classes (Normal, AF, other arrhythmia and other noisy samples).

A deep residual skip convolution neural network (RSCNN) was proposed by Kachuee et.al [35] for the classification of arrhythmia and used transfer learning to extract information learned by the model to classify the myocardial infarction. Gopika et.al [46] have shown performance improvement for the above-given method by retraining the network. Arrhythmia is a disease characterized by its irregular heartbeat rhythm. Atrial fibrillation is one of the types of arrhythmia. Through this knowledge, Gopika et.al [37] used the RSCNN for the classification of AF data into normal and abnormal using the feature extracted samples extracted from the ECG signal proposed by Andreotti et.al [33]. However, it resulted in minimal performance in the case of AF classification. Gopika et.al [37] used other deep learning models such as long short term memory (LSTM), gated recurrent unit (GRU), recurrent neural network (RNN). All the above-mentioned architectures are trained for 1000 epochs and RSCNN is trained for 75 epochs. In the existing work [37], all the other deep learning models have shown better performance when trained with a higher number of epochs but the RSCNN model achieved only minimal performance even though it has comparatively less number of learnable parameter. Thus the proposed work also aims to improve the performance of the AF classification using RSCNN by hyperparameter tuning.

Various conventional feature extraction is available for retrieving the features from the ECG. At the initial stage, Jiapu Pan And Willis J. Tompkins [51], proposed the real-time QRS detection algorithm in which they acquired 99.3% accurate QRS com-

plex detection for 24 hours MIT/BIH arrhythmia database. Desai et.al [52] proposed a discrete wavelet transform (DWT) for denoising the signal and independent component analysis (ICA) as a dimensionality reduction technique using support vector machine (SVM) classifier for arrhythmia classification. In addition to this, they used the Pan-Tompkins algorithm for R peak detection. The SVM quadratic kernel provides the highest classification accuracy of 98.4% and the highest kappa coefficient of 0.9520. Kalidas et.al [53] proposed real-time beat detection using stationary wavelet transform along with adaptive thresholding the method was evaluated on MIT–BIH arrhythmia database, QT Database and American Heart Association (AHA) database. The algorithm acquired sensitivity (SE) of 99.88% and a positive predictive value (PPV) of 99.84% on the MIT-BIH arrhythmia database. The algorithm also gained a sensitivity of 99.80% and PPV of 99.91% on the AHA database and sensitivity of 99.97% and PPV of 99.90% on the QT database. Kiranyaz et.al [54] proposed an adaptive implementation of 1-D convolutional neural networks (CNNs) for the classification of arrhythmia. The input of the model is the beat segmented ECG signals using R peak detection. The method acquired less computational complexity, high speed, and it combined the two major processes of feature extraction and classification in machine learning. They utilized the MIT-BIH Arrhythmia dataset on this approach. Xia et.al [55] Proposed a 2D CNN for the detection of Atrial fibrillation. They used the short-term Fourier transform (STFT) and stationary wavelet transform (SWT) for converting the 1D features to 2D features. The MIT-BIH atrial fibrillation (MIT-BIH AFIB) data set was used for the performance evaluation of this approach. The proposed method does not

require the detection of P or R peaks or the extraction of handcrafted manual features for cardiac disease detection. 2D CNNs with the input based on STFT has gained a sensitivity of 98.34%, a specificity of 98.24%, and accuracy of 98.29% and DCNNs with the input based on SWT has acquired sensitivity of 98.79%, a specificity of 97.87%, and accuracy of 98.63%. In all these scenarios the cardiac disease is detected using the model trained with more numbers of samples. In most conventional feature extraction techniques, they capture the peak information. It cannot be generalized that the cardiac abnormalities are captured only at the peaks. The cardio disease can change the structural morphology of the ECG signal. But in most cases, the conventional feature extraction technique fails to capture the structural variation that is continuous in nature and that can contribute to the identification of disease. To tackle the problem of lesser samples for cardiac disease detection and to interpret the features learned by the model, we propose the Chebyshev based feature extraction technique.

## 1.2 Problem statement

In the literature, there are various methods used for cardiac disease detection. In most cases, interpretation for the detection by the respective models is also missing. Hence the present work establishes the concept of explainable AI. The objective of the present work is to explore and analyze the features that the model has learned for the detection of tachycardia diseases. As an inflow work the current work also aims to improve the accuracy of the atrial fibrillation classification using the RSCNN model by hyperparameter tuning. As it shows minimum performance than any other

architecture in the classification of atrial fibrillation [5]. The deep learning models are not suggestable in cases in which the number of ECG samples are less. Thus the proposed work aims to reckon a suitable feature for the conventional machine learning algorithm and to interpret the feature learned by the model for the detection using Chebyshev features.

## 1.3 Objectives

- To interpret the features learned by the model for the detection of tachycardia diseases such as atrial fibrillation, ventricular fibrillation and sinus tachycardia from the ECG signal.
- To find the best suitable model for cardiac disease detection among the deep learning models such as RNN, LSTM, GRU, CNN and RSCNN.
- To improve the performance of RSCNN for the detection of atrial fibrillation.
- To check the possibilities of using the Chebyshev features for cardiac disease detection instead of the conventional features in case of less availability of ECG samples.
- To analyze the variation of features captured by the model in case conventional features and the Chebyshev features using less number of ECG sample.
- To find the best suitable model for cardiac disease detection in machine learning models such as SVM, logistic regression , decision tree and adaBoost.

The chapter is organized as follows: Section 2 describes the background information of the network architecture Section 3 describes the proposed work, and the thesis is concluded in section 4.

# Chapter 2

# Background

## 2.1 Recurrent Neural Network

Recurrent neural network is very popular neural network because of it's ability to store the previous state. It is also known as RNNs. It's widely used in different fields such as time series modelling, machine translation, speech recognition, generating sequences. RNNs work similar to the feed forward networks with an additional capacity of remembering the things they have learnt from the prior input for predicting the output. RNNs have no limitation in the length of input and output sequence.

Fig 2.1 shows the RNN module and its expansion of RNN in time sequential order. Here the X is the input, Z is the output, h is the hidden state and $W_{xh}$ is the weight at the inputs, $W_{hh}$ is the shared weights at the output.This shows that the weights are shared across time steps. RNN learns the time patterns and computes the next time stamp from the previous and present time step. The loop in the Fig 2.1 allows the information to pass from t to t+1. The RNN has a disadvantage of vanishing gradient problems in case of long time step dependency. This vanishing gradient problem is

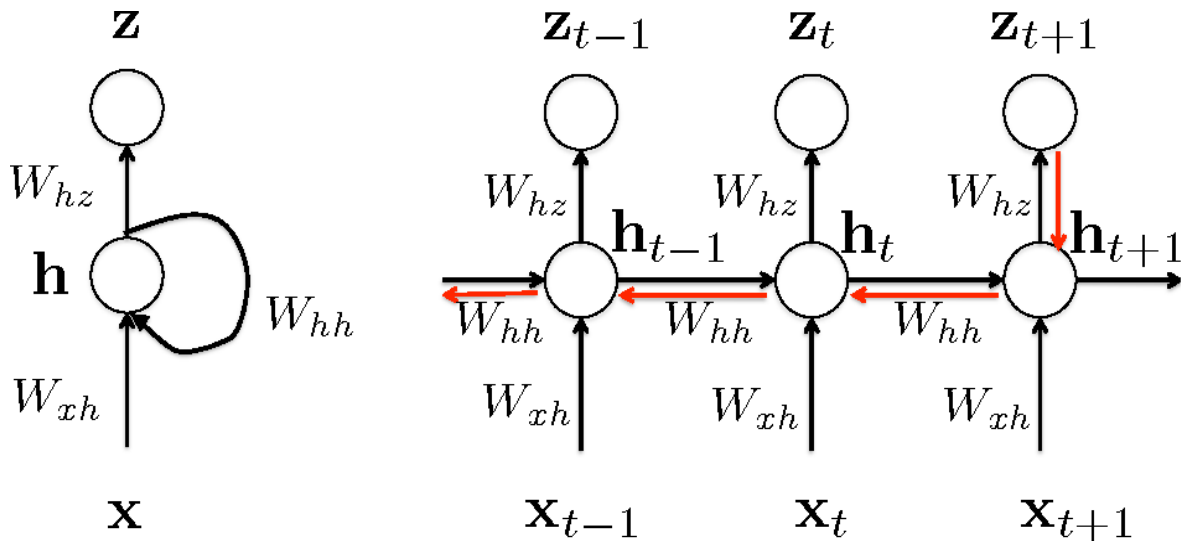



Figure 2.1: The left shows the RNN module and right shows the expansion of RNN in time sequential order [66].

addressed in Long short term memory (LSTM).

## 2.2 Long Short Term Memory

The correlation between the input $X_t$, previous hidden state ht-1information and current hidden state $h_t$ are maintained by using a simple tanh function. Thus creates a vanishing gradient problem. The LSTM maintains this correlation by using a memory unit. The previous information is stored in the memory block.

Fig 2.2 shows the unit of the LSTM. Along with memory blocks LSTM contains 4 gates. An input and output gate which regulates the information to be passed across the time step. A forget gate which nullifies unnecessary past value. The input modulation gate regulates the information to be scaled in the input gate.

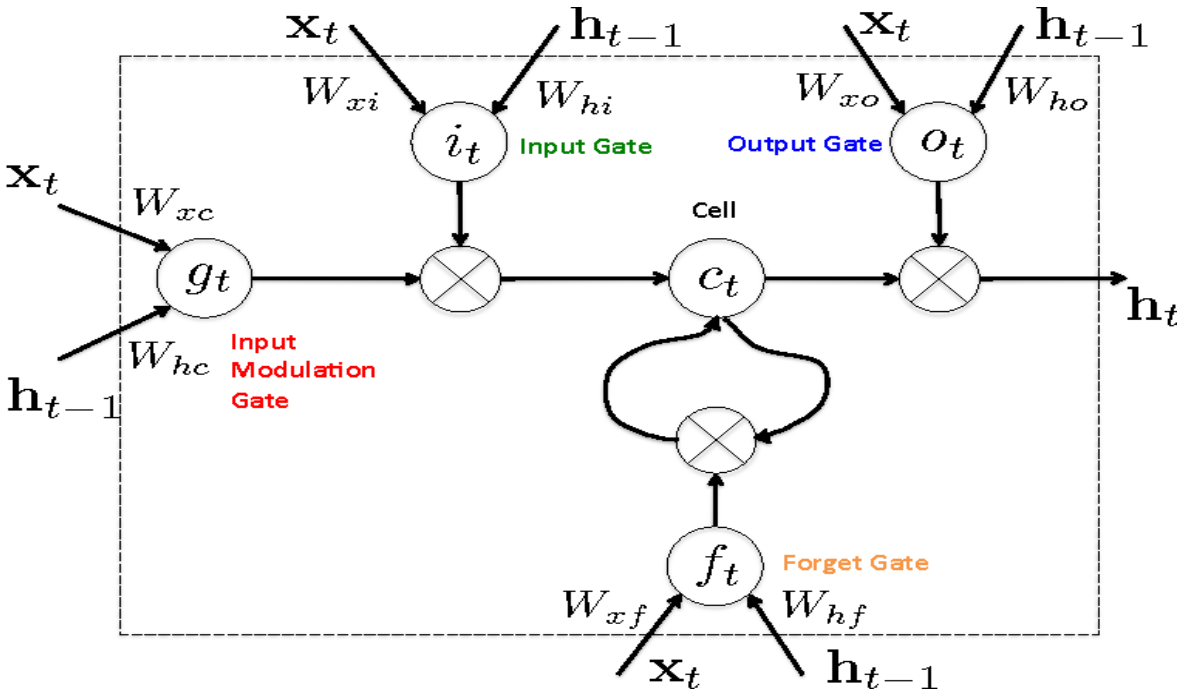



Figure 2.2: An unit structure of LSTM, including 4 gates: input modulation gate, input gate, forget gate and output gate.[66]

## 2.3 Gated Recurrent Unit

Gated Recurrent Unit(GRU) is another version of RNNs which is built to decrease the computational complexity of LSTM. Fig 2.3 shows a unit of GRU. It mainly contains 4 gates: the update gate, the reset gate, the current memory unit, and the final memory unit.The update gate is responsible for updating the weights and eliminating the vanishing gradient problem. As the model can learn on its own, it will continue to update information to be passed to the future. The reset gate acts in an opposing way by deciding how much of the past information should be forgotten, given the current state.

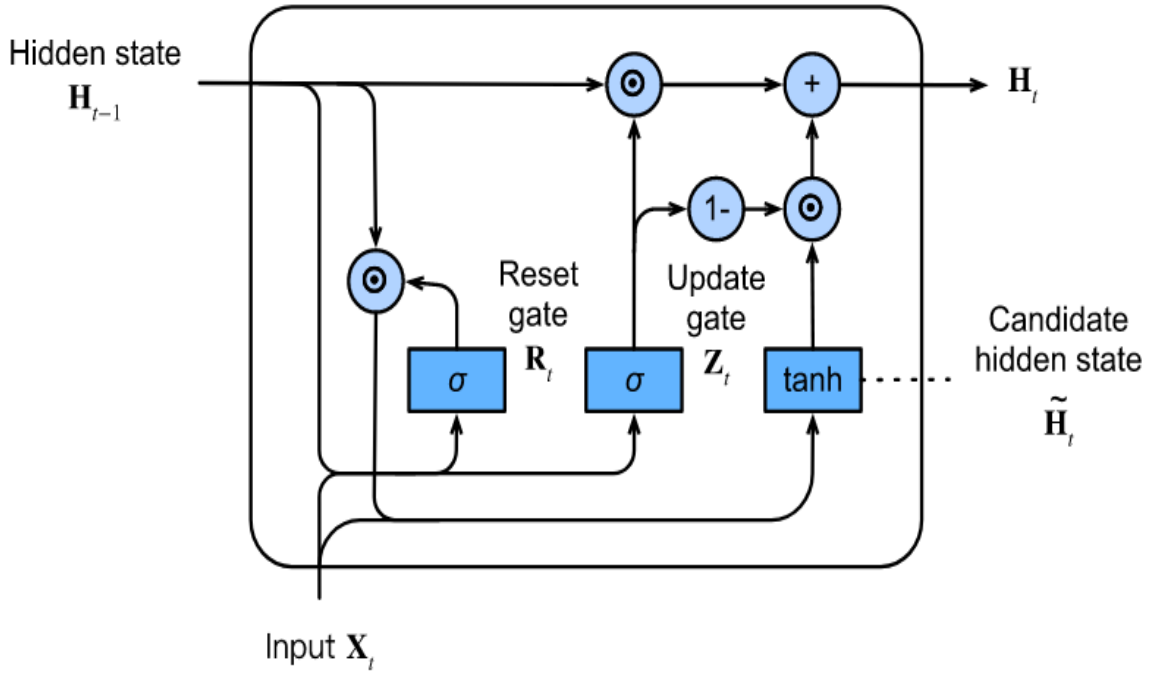



Figure 2.3: An unit structure of GRU[66].

## 2.4 Residual Skip Convolution Neural Network

Convolutional neural network is one that takes full hand in the field of computer vision. Main block of convolution neural networks is the convolution layer. Convolution takes place by striding the convolutional filter along the each pixel of images/ each sample of signal. This element wise multiplication creates the feature maps of the input. The number filter determines the portion of the input the filter gets covered which is called as the receptive filter. The number of strides determines the time steps at which the filter moves.The output dimension of convolutional operation can be determined using,

$$output = (Input - Filtersize + 2 * Zeropadding)/Stride + 1$$

The convolution layer is followed by an activation layer and which is then followed

by a Pooling layer. Most commonly used activation function is the rectifier linear unit(Relu). There are 3 types of pooling Max pooling, Average pooling and min pooling. This helps to reduce the dimension of the input. The activation function eliminates the vanishing gradient problem. Fully connected layer with an Soft-max/Sigmoid completes the convolution neural network The residual connections, skip the in between block and feed the input to the next block. The skip connection helps the model to take the input information without any loss and thus avoiding degradation problems. Skip connections also helps the model to learn deeper with equal number of parameters.

# Chapter 3

# Proposed Work

The proposed work aims to interpret the features learned by the model.To explore our objective state-of-art architectures of deep learning such as LSTM, RNN, GRU, CNN, and RSCNN [23, 26] are implemented in our present work. We considered different types of tachycardia diseases such as atrial fibrillation, ventricular fibrillation, and sinus tachycardia, which have fast beat rhythm as the common characteristic feature is used for the evaluation. To achieve this objective we use the concept of transfer learning. In our present approach, the model trained with one of the tachycardia diseases is tested to detect other different types of tachycardia diseases unseen by the model during training. In case of ECG dataset of lesser samples the proposed work aims to find the suitable feature extraction technique for interpretation and proper classification of the cardiac diseased signal using chebyshev features

## 3.1 Explainable AI for Heart Rate Variability in ECG Signal

### 3.1.1 Data set description

In this work, ECG signal datasets that are publicly available in the Physionet database [21] are used. Dataset for atrial fibrillation disease is taken from the AF classification 2017 PhysioNet CinC challenge which is referred to as the tachycardia dataset one (AF: TD1). The ventricular fibrillation disease dataset which is referred to as the tachycardia dataset two (VF: TD2). This dataset is retrieved from two sources namely Creighton University ventricular tachyarrhythmia (VF: TD2-A) and MIT-BIH ventricular ectopy (VF: TD2-B). MIT ventricular ectopy dataset has ECG signals collected from two leads. Dataset for sinus tachycardia (ST: TD3) is taken from the MIT-BIH arrhythmia database. The number of records of raw ECG signal and the corresponding number of samples based on feature extraction for all the above-mentioned datasets is presented in Table 3.1.

Each sample is a feature vector with a dimension as 169 [22]. This feature vector contains information related to heart rate variability indices and signal quality indices. The heart rate variability (HRV) indicates the changes in the heartbeats per minute. Time-domain features, frequency-domain features, non-linear features along with signal quality indices constitute the feature vector of length 169 [29]. The time-domain features give the fluctuations observed in the HRV over an interval of time. The time intervals may range from 2 minutes to 24 hours. The frequency-domain features give the

Table 3.1: Dataset Description.

| Dataset | Number of records of raw ECG signal | Number of feature extracted samples |
|---|---|---|
| Atrial fibrillation(AF: TD1) (AF dataset CINC Challenge 2017) | 8528 | 64767 |
| Ventricular fibrillation (VF: TD2-A) (Creighton University Ventricular Tachyarrhythmia Database) | 33 | 1426 |
| Sinus tachycardia (ST: TD3) (MIT-BIH Arrhythmia database of PhysioNet) | 59 | 118 |
| Ventricular fibrillation (VF: TD2-B) (MIT-BIH Malignant Ventricular Ectopy Data base) | 22 | 945 |

energy information of the ECG signal. The non-linear features indicate the complexity and nonlinearity within interbeat intervals of the ECG signal. Signal quality indices represent the segment-wise features of the ECG signal [29]. Signals are separated as segments of 10 seconds with an overlap of 50% for constructing feature extracted samples [22] so that no pieces of information are lost. The use of these feature extracted samples reduces the high computational requirement for the deep learning approach. In this, only 70% AF data (AF: TD1) is used for training.

### 3.1.2 Methodology

In this work, we propose a transfer learning approach for the detection and interpretation of the ECG signal. In this approach, the models are trained with one of the tachycardia diseases and tested with all other tachycardia diseases unseen by the model during training. Tachycardia diseases have fast beat rhythm as their common feature. This approach aids to interpret the common characteristic feature of ECG signals cor-

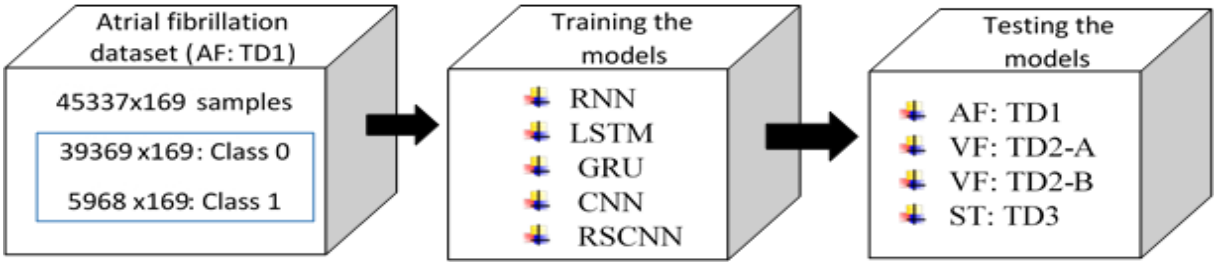



Figure 3.1: Proposed methodology for the detection of tachycardia diseases and for the interpretation of features learned by the model

responding to different types of tachycardia diseases learned by the model. The overall workflow of the methodology for the interpretation of the features learned by the model is shown in Fig 3.1. The proposed method consists of the following steps. Initially, the models are trained with the atrial fibrillation dataset. The AF dataset contains both abnormal and normal cases. Then the trained models using AF dataset are tested with other tachycardia datasets such as atrial fibrillation (AF: TD1), ventricular fibrillation (VF: TD2-A and VF: TD2-B) and sinus tachycardia (ST: TD3) separately. The state of the art deep learning architectures implemented in the present work are RNN, LSTM, GRU, CNN and RSCNN [23, 26, 27].

The benchmarks of deep learning architectures such as RNN, LSTM, GRU, CNN, and RSCNN [23, 26, 27] are contemplated for the study. Details about recurrent neural networks are given in Table 3.2. The input layer of each model is modified to 169x1. As the input signal has a feature vector of size 169x1. Recurrent neural networks considered are RNN, LSTM, GRU which have one hidden layer with 64 units. The second layer (output layer) is dense with the number of neurons same as the number of classes considered. Atrial fibrillation dataset (AF: TD1) which is considered for training

have two classes i.e.the one without atrial fibrillation (Normal: Class 0) and the other with atrial fibrillation (Abnormal: Class 1). So, the final dense layer consists of two neurons with a sigmoid activation function.

Table 3.2: Architecture details of RNN, CNN, LSTM networks

| Architecture | RNN | LSTM | GRU |
|---|---|---|---|
| Input layer | 169x1 | 169x1 | 169x1 |
| Recurrent layer | RNN | LSTM | GRU |
| Output layer | Dense layer (Sigmoid) | Dense layer (Sigmoid) | Dense layer (Sigmoid) |

The number of learnable parameters varies according to the chosen architecture of the model. For the recurrent neural networks the computation of number of learnable parameter is given by

$$paramRNNs = f \times [ns(ns + i) + ns] \tag{3.1}$$

where f is the number of fully connected neural networks, ns is the number of neurons in the hidden layer and i is the input size. In the case of RNN number of fully connected neural networks is 1, for GRU it is 3 and for LSTM it is 4. For dense layers, the number of the learnable parameter is computed by

$$paramDense = [ns \times i + b] \tag{3.2}$$

where b is the bias. Details of the CNN model is shown in Fig 3.2. The CNN model contains a convolution layer with 64 filters of size 3 with stride 1. This convolution layer is accompanied by ReLU (Rectified linear unit) activation function. The output from the convolution layer is mapped into nonlinear output using the activation function for

avoiding the vanishing gradient problem. The model also contains two dense layers: one with 128 neurons and other with 2 neurons which serve as the output layer with a soft-max activation function. Number of the learnable parameter of the convolution neural network is given by

$$paramCNNs = nf \times fs + b \tag{3.3}$$

where nf is the number of filters, fs is the filter size and b is the bias.

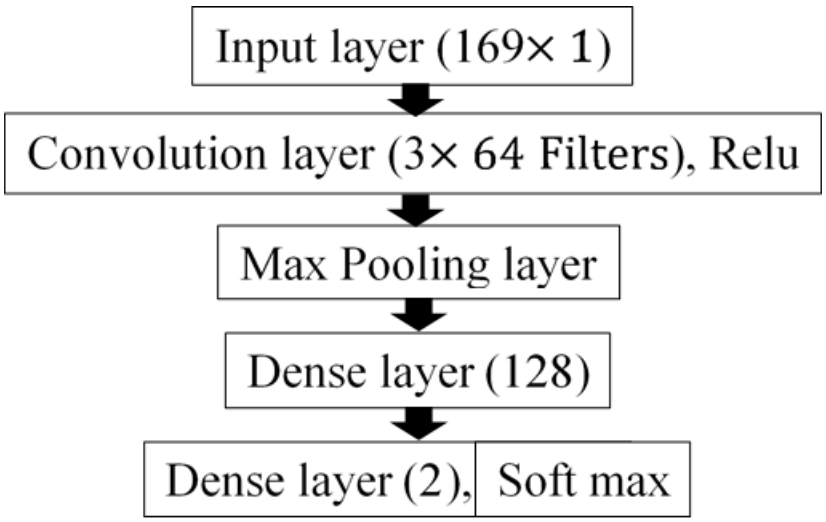



Figure 3.2: 1-D Convolution neural network

Details of the residual convolution neural network (RSCNN) architecture is shown in Fig 3.3. RSCNN contains 13 weighted layers which include 11 convolution layers and 2 dense layers. The first layer is the input layer of size 169x1 which is the same as the size of the feature vector. Convolution layer has 32 filters with 3 as the filter size in each layer. The network has residual blocks. The residual blocks contain two convolution layers with the ReLU activation function. Succeeded by a max-pooling layer for the

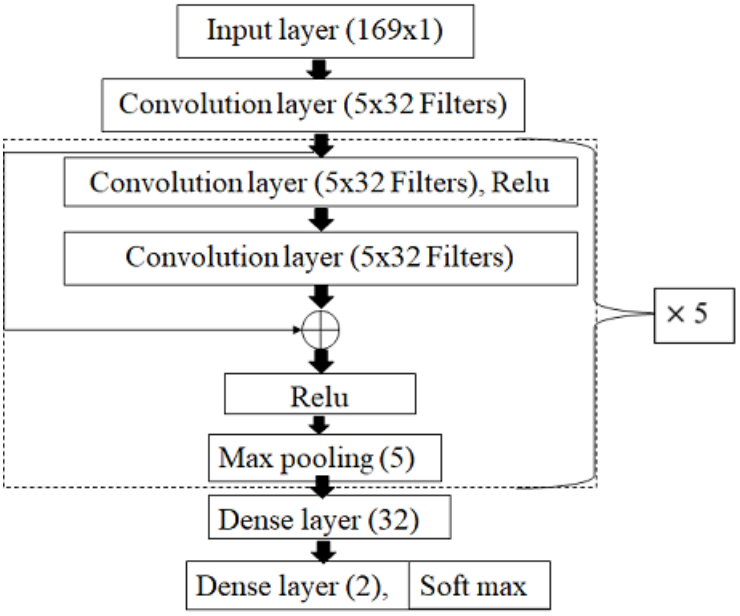



Figure 3.3: Residual Convolution neural network

dimensionality reduction. A skip connection is also included in the residual block. As shown in Fig 3.3, these residual blocks are repeated 5 times. The dense layer of 32 neurons is included after the residual blocks. The final output layer is a dense layer with 2 neurons with a soft-max activation function.

The number of learnable parameters computed for all the benchmark architectures is tabulated in Table 3.3.

Table 3.3: Number of learnable parameters computed for RNN,LSTM,GRU, CNN and RSCNN

| Architectures | Number of learnable parameters |
|---|---|
| RNN | 150170 |
| LSTM | 60,290 |
| GRU | 45,250 |
| CNN | 60,290 |
| RSCNN | 54,914 |

Table 3.4: Experimental results for AF CINC 2017 dataset class 0 (AF: TD1 class 0)

| | LSTM | GRU | RNN | CNN | RSCNN |
|---|---|---|---|---|---|
| Accuracy % | 95.77 | 97.87 | 96.87 | 97.13 | 97.92 |
| Sensitivity | 0.96 | 0.98 | 0.97 | 0.98 | 0.98 |
| F1_score | 0.97 | 0.98 | 0.98 | 0.97 | 0.98 |
| Specificity | 0.957 | 0.96 | 0.96 | 0.97 | 0.97 |

Table 3.5: Experimental results for AF CINC 2017 dataset class 1 (AF: TD1 class 1)

| | LSTM | GRU | RNN | CNN | RSCNN |
|---|---|---|---|---|---|
| Accuracy % | 88.03 | 87.25 | 90.34 | 78.14 | 83.93 |
| Sensitivity | 0.88 | 0.87 | 0.90 | 0.78 | 0.84 |
| F1_score | 0.82 | 0.87 | 0.86 | 0.81 | 0.86 |
| Specificity | 0.88 | 0.90 | 0.90 | 0.78 | 0.83 |

### 3.1.3 Training and testing

All the benchmark deep learning architecture (RNN, LSTM, GRU, CNN, and RSCNN) are trained with 70% of atrial fibrillation data (AF: TD1) for 1000 epochs with the batch size 1000 samples for each architecture. The models are tested with 30% of atrial fibrillation data and also with other tachycardia datasets such as ventricular fibrillation (VF: TD2-A and VF: TD2-B) and sinus tachycardia data (ST: TD3).

Table 3.6: Experimental results for CU-ventricular tachycardia dataset (VF: TD2)

| | LSTM | GRU | RNN | CNN | RSCNN |
|---|---|---|---|---|---|
| Accuracy % | 82.75 | 77.84 | 90.88 | 74.89 | 68.93 |
| Sensitivity | 0.83 | 0.78 | 0.91 | 0.75 | 0.69 |
| F1_score | 0.91 | 0.88 | 0.95 | 0.86 | 0.82 |
| Specificity | 0.83 | 0.77 | 0.90 | 0.75 | 0.68 |

Table 3.7: Experimental results for Sinus Tachycardia (ST: TD3)

| | LSTM | GRU | RNN | CNN | RSCNN |
|---|---|---|---|---|---|
| Accuracy % | 27.97 | 22.88 | 23.73 | 26.50 | 21.19 |
| Sensitivity | 0.28 | 0.23 | 0.24 | 0.26 | 0.21 |
| F1_score | 0.44 | 0.37 | 0.25 | 0.42 | 0.35 |
| Specificity | 0.28 | 0.23 | 0.24 | 0.26 | 0.21 |

Table 3.8: Experimental results for MIT-BIH Malignant Ventricular Ectopy Database(VF: TD21-Lead 1)

| | LSTM | GRU | RNN | CNN | RSCNN |
|---|---|---|---|---|---|
| Accuracy % | 92.17 | 90.69 | 94.71 | 88.67 | 76.61 |
| Sensitivity | 0.92 | 0.91 | 0.95 | 0.89 | 0.77 |
| F1_score | 0.96 | 0.95 | 0.97 | 0.94 | 0.87 |
| Specificity | 0.92 | 0.91 | 0.95 | 0.88 | 0.77 |

Table 3.9: Experimental results for MIT-BIH Malignant Ventricular Ectopy Database(VF: TD21-Lead 2)

| | LSTM | GRU | RNN | CNN | RSCNN |
|---|---|---|---|---|---|
| Accuracy % | 93.86 | 90.16 | 94.18 | 88.67 | 80.32 |
| Sensitivity | 0.94 | 0.90 | 0.94 | 0.89 | 0.80 |
| F1_score | 0.97 | 0.95 | 0.97 | 0.94 | 0.89 |
| Specificity | 0.94 | 0.90 | 0.94 | 0.88 | 0.80 |

| | | **GRU:AF** | | **RNN:AF** | |
|---|---|---|---|---|---|
| True label | 0 | **16514** | **359** | **16345** | **528** |
| | 1 | **326** | **2231** | **247** | **2310** |
| | | 0 | 1 | 0 | 1 |
| | | Predicted label | | Predicted label | |

Figure 3.4: Confusion matrix for GRU and RNN model tested with AF: TD1 data

Table 3.10: Variation in morphology of ECG signal for different type of tachycardia diseases.

| **Disease** | **P wave** | **QRS complex** | **T segment** | **Heart Rate** | **Other features** |
|---|---|---|---|---|---|
| Atrial fibrillation | Absent | Not effected | Not effected | >100 beats/min | Presences of fibrillatory waves in the baseline |
| Ventricular fibrillation | Absent | Absent | Absent | >100 beats/min | Presences of Irregular waves |
| Sinus Tachycardia | Not effected | Not effected | Diminished | >100 beats/min | Similar to normal sinus rhythm |

Table 3.11: Fivefold cross-validation accuracy for CNN and RSCNN

| Models | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| CNN (Fivefold) | 95.48+-0.39% | 0.81 | 0.86 | 0.83 |
| CNN (Single Fold) | 95.32% | 0.851 | 0.781 | 0.815 |
| RSCNN (Fivefold) | 94.96+-0.19% | 0.83 | 0.78 | 0.80 |
| RSCNN (Single Fold) | 96.08% | 0.92 | 0.91 | 0.91 |

Table 3.12: Training and testing time for all the deep learning models used in the present work.

| | LSTM | GRU | RNN | CNN | RSCNN |
|---|---|---|---|---|---|
| Training Time (HH.MM.SS) (AF-TD1) | 08:09:00. | 11:41:54. | 01:46:09. | 06:48:51. | 01:31:04. |
| Testing Time (Seconds) (AF-TD1) | 1.47 | 0.74 | 0.15 | 4.22 | 11.29 |
| Testing Time (Seconds) (CU VF -TD2) | 0.52 | 0.43 | 0.23 | 0.31 | 0.71 |
| Testing Time (Seconds) (ST-TD3) | 0.01 | 0.02 | 0.01 | 0.04 | 0.03 |
| Testing Time (Seconds) (VF-TD21- lead 1) | 0.05 | 0.03 | 0.03 | 0.11 | 0.23 |
| Testing Time (Seconds) (VF -TD21- lead 2) | 0.05 | 0.04 | 0.03 | 0.12 | 023 |

Table 3.13: Performance comparison of the present work with respect to state of art methods for the AF-TD1,VF-TD2,ST-TD3 and VF-TD21 dataset.

| | Accuracy | Sensitivity | F1 score | Specificity |
|---|---|---|---|---|
| **Proposed Methodology (AF-TD1)** | **96.47** | **0.93** | **0.92** | **0.90** |
| J. Chanthercrob et.al (AF -TD1) | - | 0.97 | - | 1.00 |
| Gopika et.al (AF -TD1) | 96.47 | 0.93 | 0.92 | 0.90 |
| **Proposed Methodology (CU_VF-TD2)** | **90.88** | **0.91** | **0.95** | **0.90** |
| Boreiko et.al (VF-TD2) | - | 0.83 | - | - |
| M. A. Sohail et.al (VF-TD2,VF-TD21) | 98.03% | - | - | - |
| **Proposed Methodology (ST-TD3)** | **27.97** | **0.28** | **0.44** | **0.24** |
| Gopika et.al MIT arrhythmia (N,S,V,F,Q) | 97.94 | 0.98 | - | - |
| **Proposed Methodology (VF-TD21-lead 1)** | **94.71** | **0.95** | **0.97** | **0.95** |
| **Proposed Methodology (VF-TD21-lead 2)** | **94.18** | **0.94** | **0.97** | **0.94** |
| Mohanty et.al (VF-TD2/VF-TD21) | 99.18 | 0.97 | - | 0.99 |

Table 3.14: Performance of the deep learning models towards the noisy signal in AF -TD1 dataset

| | LSTM | GRU | RNN | CNN | RSCNN |
|---|---|---|---|---|---|
| Percentage of noisy signal falling into Normal Class | 78.66% | 78.76% | 72.99% | 89.17% | 73.35% |
| Percentage of noisy signal falling into Abnormal Class | 21.34% | 21.24% | 27.01% | 10.83% | 26.65% |

| True label | | RNN:VF | | LSTM:ST | | RNN:VF(Lead I) | | RNN:VF(Lead II) | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | **130** | **1296** | **85** | **33** | **50** | **895** | **55** | **890** |
| | | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| | | Predicted label | | Predicted label | | Predicted label | | Predicted label | |

Figure 3.5: Confusion matrix for RNN model tested with VF: TD2-A, LSTM model tested with ST: TD3 data,RNN model tested with VF: TD2-B data for lead I and lead II

### 3.1.4 Result and analysis

Accuracy,sensitivity,F1-score and Specificity are the evaluation metric used for the performance assessment of all the deep learning architectures implemented in this work. In the biomedical field, the sensitivity score has very high importance. Let's consider the condition of disease as a positive case and the normal as a negative case from the medical perspective. Hence, sensitivity is the measure of the ratio of actual positive detected as positive. Since the disease unidentified is a threat to life, the sensitivity score is considered for the evaluation of the disease detection along with the accuracy.Specificity measures the effectiveness of the to detect the normal cases. The measure of combined precision and sensitivity score is incorporated in the F1-score. The confusion matrix gives the exact idea about the number of samples that are correctly classified and miss classified. The Four parameters used for the analysis of the confusion matrix are True positive (TP), False positive (FP), True negative (TN) and False negative (FN). TP represents the number of samples that are positive and predicted correctly as positive. FP represents the number of samples that are negative and predicted wrongly as pos-

itive. TN represents the number of samples that are negative and predicted correctly as negative. FN represents the number of samples which is positive and predicted wrongly as negative. The accuracy score, sensitivity, F1-score, and Specificity in percentage for the models tested with atrial fibrillation (AF: TD1) for class 0 and class 1, ventricular fibrillation (VF: TD2 and VF: TD21) and sinus tachycardia (ST: TD3) dataset are tabulated in Table 3.4, Table 3.5, Table 3.6, Table 3.7, Table 3.16, and Table 3.9 respectively. Class 0 represents the normal class and class 1 represents the AF (abnormal class). The main purpose of the present work is to determine whether the model trained on one type of tachycardia can detect other types of tachycardia diseases. This is performed in order to interpret the common characteristics of ECG signals that are affected by different types of tachycardia diseases learned by the trained model. Hence, the model trained on AF: TD1 dataset is tested with AF: TD1, VF: TD2-A, VF: TD2-B and ST: TD3.The feature specific differences between three types of tachycardia diseases used in the present work is given in Table 3.18.

In atrial fibrillation data (AF: TD1) a total of 19430 samples were tested in which 16873 are class 0 and 2557 are class 1. For the normal class, the RSCNN model gained an accuracy of 97.92% which is higher than other models and RNN gained an accuracy of 90.34% which is higher than other models for abnormal class. While considering the average accuracy including both classes, the GRU model has performed better than other models with an accuracy of 96.47%. While considering the sensitivity score for the abnormal class, RNN has gained a score of 0.90 which is higher than other models. Thus for AF: TD1 RNN has performed better than other models in detecting the abnormal

class. The confusion matrix for the GRU and RNN model is shown in Fig.3.4. The diagonal elements in the figure represent the TP and TN number of samples.

In the case of ventricular fibrillation, CU ventricular tachycardia dataset (VF: TD2-A) is used for the evaluation. The models are tested using 1426 samples of tachycardia data. The RNN has acquired a percentage accuracy of 90.34% which is higher than other models. Analyzing the sensitivity score of these models, we can understand that RNN has a better score of 0.91 than other models. The confusion matrix for the RNN model which is tested with VF: TD2-A is shown in Fig.3.5. From the figure, it is clear that the number of TP samples is 1296.

For the validation of the above result, the second dataset of ventricular fibrillation disease VF: TD2-B from the MIT-BIH malignant ventricular ectopy database was taken. In this dataset, 945 samples of VF were tested on different models such as LSTM, GRU, RNN, CNN, RSCNN. Among these, RNN has shown better performance with an accuracy of 94.71% for the lead I ECG signal of VF: TD2-B. The RNN model has gained the best sensitivity score of 0.95 when compared with other models. The RNN has shown the same trend for VF: TD2-B lead II data with an accuracy of 94.18% and a sensitivity score of 0.94. From the result acquired from VF: TD2-A and VF: TD2-B dataset, it is clear that RNN had performed well for the detection of VF disease. The confusion matrix of the RNN model of VF: TD2-B for the lead I and lead II is shown in Fig.3.5. It shows that TP samples are 895 and 890 for the lead I and lead II respectively. The next evaluation is done using tachycardia dataset three (ST: TD3) known as sinus tachycardia. The models are evaluated against 118 samples of ST data. The model

acquired sensitivities of 0.28, 0.23, 0.24, 0.19, 0.21 for models such as LSTM, RNN, GRU, CNN, RSCNN respectively and also gained accuracies of 27.97%, 22.88%, 23.73%, 18.64%, 21.19% for LSTM, RNN, GRU, CNN, RSCNN respectively. From these values, it is evident that the models were not able to detect the ST: TD3 disease. The confusion matrix for ST: TD3 is shown in the second matrix of Fig.3.5. From the figure, it is apparent that only 33 samples are TP samples and the remaining samples are miss classified as TN. The performance evaluation in terms of F1-score and specificity shows the same trend as that of accuracy and sensitivity.

From the analysis, we found that RNN has performed better than architectures such as LSTM, GRU, and CNNs. The recurrent neural networks (RNNs) (i.e. RNN, GRU, and LSTM) and convolution neural networks are well known for their performance in biomedical applications. The RNNs have the ability to remember the previous time step and use that information to predict the next. The various RNNs such as RNN, LSTM, and GRU are different from each other due to the presences of gate such as forget gate, input gate and output gate etc. The RNN is a simple feed-forward network with a feedback loop. The LSTM and GRU have additional gates to avoid long term dependency of the previous states. In the case of tachycardia disease detection using the ECG signal, there is a possibility that the GRU and LSTM could miss the important pieces of information while passing signal vectors through different gates. In RNN, the memory cell has the ability to store all the information from the previous state, thus gaining better performance than other architectures. While considering the CNN architecture, the structural information is stored in the convolutional layer. When

compared to RNNs it lacks the capacity to capture timely information. Therefore, CNN could not perform better than RNNs.

The five fold cross-validation results for the CNN and RSCNN models trained with AF-TD1 and tested with the same are shown in Table 3.21. From the results, we analyze that the CNN model achieved approximately equal performances with that of the single fold results. In the case of RSCNN, there is a slight decrease in performance while taking an average performance of five fold testing. From the accuracy, we observe that the RSCNN has a variation of +/- .19 % which makes it approximately equal to the previous results achieved. Five fold results provide an validation of experiments performed.

The time based analysis is incorporated for adding more details to the experiments. The time based analysis includes the time taken by the model to get trained and the time taken by the model to get tested. The time required for each model to get trained with AF-TD1 dataset and get tested with AF-TD1, CU VF -TD2, ST-TD3, VF-TD21- lead 1, VF-TD21- lead 2 data samples is shown in Table 3.22. The LSTM, GRU, and CNN networks are trained for 1000 epochs each. For the networks such as RNN, the model is trained for 298 epochs and RSCNN the model is trained for 105 epochs. From the experimental results, we observe that the RSCNN has taken the least time for training as the number of epochs is less, compared to all other models. Then the second least time is taken by the RNN model since it has the least number of learnable parameters among other models. Even though the GRU has the second least number of learnable parameters, it has taken much greater time than other models while comparing other

models based on accuracy and number of learnable parameters. RNN has achieved an accuracy of 96.87%, 90.34%, 90.88%, 94.71% and 94.18% for AF-TD1 class 0, AF-TD1 class 1, CU_VF-TD2, MIT_VF-TD21- lead 1 and MIT_VF-TD21- lead 2 respectively. RNN model have better performance than any other models in terms of accuracy and time.

The comparison of the present work with respect to the current state art of method is shown in Table 3.19. For the AF CINC challenge dataset(AF-TD1), CU Ventricular fibrillation dataset (VF-TD2), MIT BIH Arrhythmia database(ST-TD3) and Malignant ventricular Ectopy database (VF-TD21) respectively. From the results shown in Table 3.19, we could interpret that the current methodology was able to achieve comparable performance with respect to the state of art methods. An exception was found in the case of ST disease. This difference in the performance of sinus tachycardia is because, the current methodology employs a model trained with AF disease to detect other tachycardia diseases such as VF, ST. The model which is trained with AF was not able to detect ST segments. The main feature of ST disease which is the upright P waves is different from the AF and VF disease features. The features of ST don't share common feature distribution with the other tachycardia disease dataset (AF and VF). The difference in feature distribution failed the model trained by the AF dataset to detect ST disease.

In order to check the noise robustness, the model is tested with the noisy segments from the AF CINC challenge dataset. These noisy segments are not annotated. The results are tabulated in Table 3.20. All the deep learning models classify the majority

of the noisy segments into the normal class. The expected result was the reverse, as the noisy segments are considered the deviation from the normal class. This enforce the direction of training the deep learning algorithms to detect multi-class with the inclusion of unlabelled noisy segments along with the normal and abnormal. This can be considered as the future scope of the present work.

From the experimental results and analysis, we observed that the model trained with AF was able to detect VF and failed to detect ST. The expected results were that, if the model has learned the fast beat rhythm, it must be able to detect the other types of tachycardia diseases.From this, we were able to interpret that, even though AF, VF and ST fall under the common disease type called tachycardia, the features learned by the model were common to AF and VF which was not applicable to ST. The presence of Upright P-wave is the characteristics of an ECG signal specific to ST disease. This explains that the model just did not capture the fast beat rhythm, which is the coarse level feature to detect all the three different types of tachycardia diseases. Instead, it captured the disease specific feature, which differentiates the ST from AF and VF. This analysis from the experiments conducted led to "Explainable AI".

The findings based on the experiments conducted and the results obtained are given below

- Among all the benchmark deep learning architectures implemented for the different tachycardia disease detection, RNN was able to perform better based on our proposed transfer learning approach.

- Even though RNN was able to detect two types of tachycardia diseases namely atrial fibrillation and ventricular fibrillation, it failed to detect sinus tachycardia. This may be due to the absence of P-wave characteristics in the trained model (using AF: TD1). ST has an up-right distinct P-wave that differentiates it from atrial and ventricular fibrillation.

## 3.2 Performance Improvement of Deep Residual Skip Convolution Neural Network for Atrial Fibrillation Classification

The proposed method uses a deep residual skip convolution neural network for the classification of the atrial fibrillation and normal ECG samples. The same atrial fibrillation dataset from the CINC challenge is used for the study. The input of the RSCNN is a feature vector of size $169 \times 1$. For training the CNN, we used Adam optimizer and categorical cross-entropy as the loss function. The hyperparameters such as the number of epochs and learning rates are tuned in such a way that the model was able to learn the features perfectly. The proposed methodology is illustrated in Fig 3.6. The model is trained with 70% of the AF dataset. The trained model is evaluated with 30% of the AF dataset. The metrics such as accuracy, F1 score, precision, and recall are used for the evaluation of the model.

### 3.2.1 Experimental Results

The number of epochs is fixed according to the performance of the model to classify AF. The model was trained for 300 epochs and with a learning rate of 0.001 (default). The

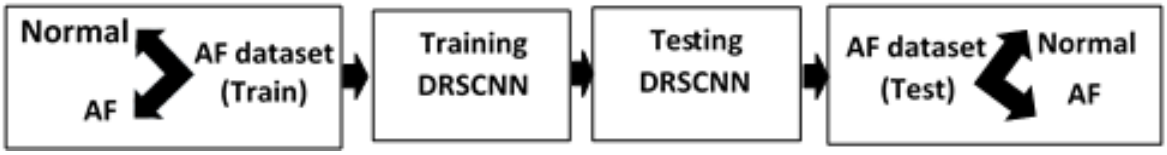



Figure 3.6: Proposed methodology for performance improvement of DRSCNN for AF classification

deviation of training accuracy and loss of the model concerning the epoch are shown in Fig 3.7 and Fig 3.8.

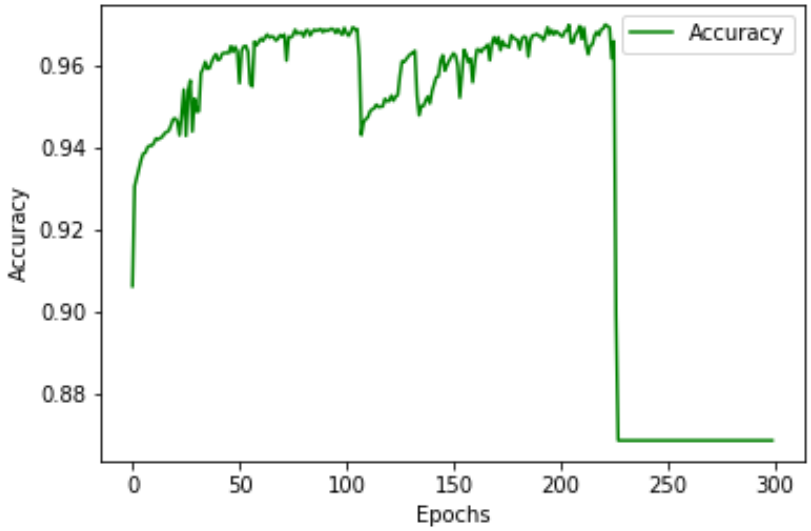



Figure 3.7: The variation in accuracy according to the increase in number of epochs for AF classification.

From Fig 3.7 and Fig 3.8, it is evident that the training loss decreased from infinity to 0.07 till $105^{th}$ epoch and the training accuracy increased from 0 to 97%. After $105^{th}$ epoch, the training accuracy decreased for a few epochs and again there is an increase in accuracy till $227^{th}$ epoch. Later, there is a steep decrease in accuracy. The training loss doesn't have any vigorous change after $105^{th}$ epoch till $227^{th}$ epoch. But after $227^{th}$

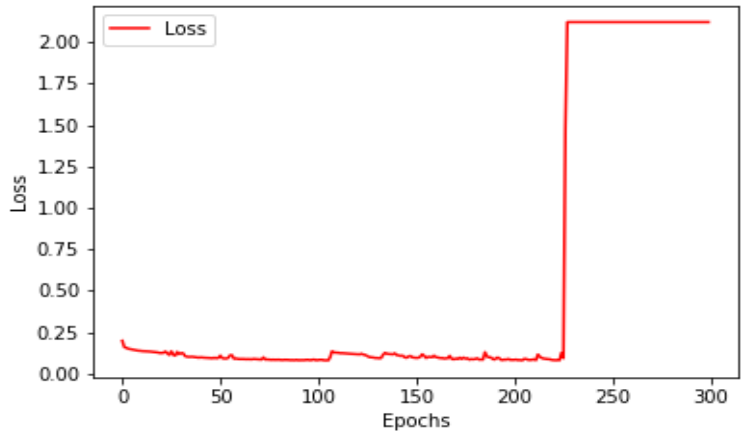



Figure 3.8: The variation in loss according to the increase in number of epochs for AF classification.

epoch, there is a steep steady increase in loss. This is due to the overfitting of the model i.e., the model is trained more so that it has started to generalize all the given data. Such cases affect the overall performance of the model. Thus the model weight in $105^{th}$ epoch which showed the best training accuracy of 97% and loss of 0.07 was used for testing. The evaluation of the model efficacy is carried out using metrics such as accuracy and weighted average F1 score, precision, and recall. The evaluation metric acquired by the model together with the values acquired by the existing method is shown in Table 3.15. The proposed model acquired an accuracy of 96.08% and weighted average F1 score of 0.96 which is higher than other prevailing methods. The confusion matrix containing the true positive, false positive, true negative and false negative is shown in Fig 3.9. From the confusion matrix, it is clear that the model classified 98% of the normal data as normal (True negative) and 84% of the AF data as abnormal (True positive).

The receiver operating characteristics (ROC) and area under the curve (AUC) is

Table 3.15: Evaluation metric results obtained for the AF classification using DRSCNN and performance of previously existing method

| Work | Accuracy | F1 score | Recall | Precision |
|---|---|---|---|---|
| Proposed | 96.08 % | 0.96 | 0.96 | 0.96 |
| Gopika et.al[37] | 50.00 % | 0.67 | 0.50 | 0.25 |

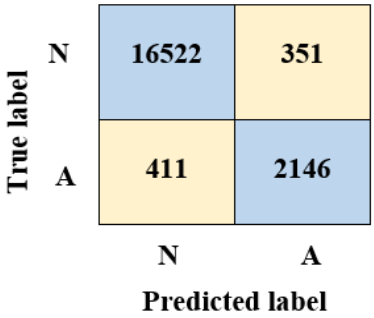



Figure 3.9: Confusion matrix for AF classification obtained by hyper parameter tuning of DRSCNN model.

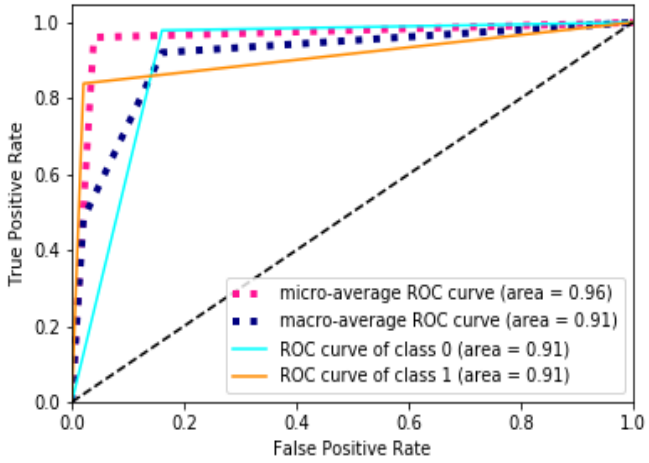



Figure 3.10: ROC curve and AUC for AF classification obtained by hyper parameter tuning of DRSCNN model

one another parameter that measures the efficiency of the model. In the ROC curve, if the graph goes more towards the upper left corner it indicates better performance. AUC is the area covered by the curve and it measures the quality of performance.

Higher the area, better the performance. The ROC curve and AUC are shown in Fig 3.10. The AUC for the AF classification is 0.96. From the results, we understand that the RSCNN model classified the AF data with 96% accuracy. The performance comparison of RSCNN with the existing approaches is shown in Table 3.16 From the table, it's evident that the performance of the proposed approach is 4% higher than the existing methods in the literature in the case of normal class classification and has shown equal performance in case of classification of abnormal class.

Table 3.16: Performance comparison for the proposed method and the method implemented in the literature for AF classification

| Authors | F1_score Normal | F1_score AF |
|---|---|---|
| Proposed | 0.98 | 0.86 |
| Teijeiro et.al[38] | 0.92 | 0.86 |
| Datta et.al[39] | 0.92 | 0.82 |
| Zabihi et.al[40] | 0.91 | 0.84 |
| Hong et.al[41] | 0.92 | 0.85 |
| Kamaleswaran et.al[42] | 0.91 | 0.82 |
| Plesinger et.al[43] | 0.92 | 0.82 |
| Kropf et.al[44] | 0.91 | 0.84 |
| Rizwan et.al[45] | 0.91 | 0.78 |
| Gopika et.al (GRU)[37] | 0.96 | 0.87 |
| Gopika et.al (DRSCNN)[37] | 0.33 | 0.00 |

Unlike the recurrent neural networks used by Gopika et.al [37] which was trained for 1000 epochs, the present work attains the benchmark accuracy by training the RSCNN network for just 105 epochs. The number of learnable parameters of RSCNN is less than LSTM and the same in the case of GRU. Even though the number of learnable parameters for RNN is less than RSCNN the number of epochs used to train the RSCNN is lesser than the number of epochs used to train the RNN model. Hence

the highlight of the present work is the attainment of benchmark accuracy using the lighter model based on the number of learnable parameters and trained for less number of epochs. This was feasible due to the beat segmented features containing the heart rate variability in time and frequency domain existing in the literature [33].

## 3.3 Detection of Cardiac Disease for Less Number of ECG samples using Chebyshev Coefficients

### 3.3.1 Chebyshev Interpolation System

A Chebfun is a function which represents every functions in an interval [-1,1]. Chebfun is defined based on the fact that the smooth function can be symbolized by using the polynomial interpolations in chebyshev points. The number of chebyshev points are stored automatically to machine precision using an adaptive technique [14]. The Chebfun system only stores minimum chebyshev points for representing a signal even with a large number of samples. The chebyshev points are defined by

$$x_j = -cos(j\pi/N), 0 \leq j \geq N, \tag{3.4}$$

The polynomial interpolant is defined as a unique polynomial at any data values at the chebyshev points and a chebyshev series is defined as an expansion represented as

$$f(x) = \sum_{n=0}^{\infty} a_k T_k(x) \tag{3.5}$$

$a_k$ is known as the chebyshev coefficients. Peculiarity of the Chebyshev is that we will get a continuous function while using the chebyshev functions in the MATLAB.

The chebfun of the signal 's' is computed using the command

$$f = Chebfun(s) \tag{3.6}$$

f is the cheb function then, the command

$$C = chebcoeffs(f) \tag{3.7}$$

returns its chebyshev coefficient. The truncation of the valid coefficient are computed by using the command

$$t = chebfun(f, trunc, m) \tag{3.8}$$

m represents the number of the coefficients to which the function is to be truncated.

In the literature, the Cheb function (Chebfun) is used in the speech epoch extraction purpose [56] and in the electrical system [57]. B.Ganga et.al [56] used the Chebfun for the accurate epoch extraction of the telephonic speech signal. The proposed method gained an improvement in performance by incorporating changes in the prevailed zero frequency filtering by including Chebyshev interpolation method. Neethu Mohan et.al [57] proposed a combination of variational mode decomposition and chebyshev interpolation for the estimation of power system frequency and the amplitude.

### 3.3.2 Dataset Description

ICBHI 2019 scientific challenge cardiovascular disease dataset is used for this study [65]. The dataset contains both generalized normal and abnormal cardiac vascular disease cases. The dataset contains 24 normal cases (negative cases) and 13 abnormal cases (positive cases).

### 3.3.3 Feature Extraction Technique based on Chebfun

The main feature that is considered for this study is the approximated signal that is extracted from the original signal by using the Chebfun. The signal is initially downsampled from the sampling frequency of 44100Hz to a sampling frequency of 250 Hz. The signal is then transformed into a Chebyshev function. From the Chebyshev function the Chebfun coefficients are extracted. The main advantage of the Chebfun is that we don't need to consider the whole signal coefficients. The Chebfun provides the truncated coefficients of the signal.The number of coefficients varies for each signal as a number of samples for each signal are different. For making a uniform number of samples in each signal we can either reduce the number of coefficients to the minimum number available or we can zero pad the signal with a number of zeros to match each signal length. Here the Chebfun coefficients considered are fixed according to the minimum coefficient required for the signal reconstruction. The Chebfun coefficients are fixed to 3500 coefficients for this study.

### 3.3.4 Methodology

The aim of this study is to analyze whether the Chebfun feature can replace conventional feature extraction techniques. Using the ability of Chebfun coefficients to reconstruct the original signal the study aims to interpret the features learned by the model. For this study, we have considered the feature extraction techniques such as the Pan-Tompkin algorithm [58], Hamilton algorithm [59], Engelse and Zeelenberg algorithm [60], Christov algorithm [61], stationary wavelet transform [62], and matched

filter [63] . Due to the minimum number of sample sizes, the analysis is conducted using the state-of-art machine learning algorithms such as SVM, Logistic Regression, Decision Tree, and AdaBoost. The best model parameters for the study are computed by running the grid search over the model parameters. For better validation of the result using a small dataset, instead of conventionally splitting the dataset into training, testing, and validation sets we consider the leave-one-out cross-validation (LOOCV). In this type of validation the model is validated against each of the data samples. The workflow of the validation set up is, first, the model will get trained for all the available data samples except one. The model is getting tested by predicting the label for one data sample that is untrained. Similarly this procedure gets repeated for all the data samples by shifting one after the other. The model accuracy is the average of accuracy of the model prediction for all the data samples. The methodology followed for the study is shown in the Fig3.11. The features from the ECG signal of the SC dataset is extracted using Chebyshev features , and other conventional features and is fed to the classifiers for making predictions. These prediction predictions are further analysed for the study.

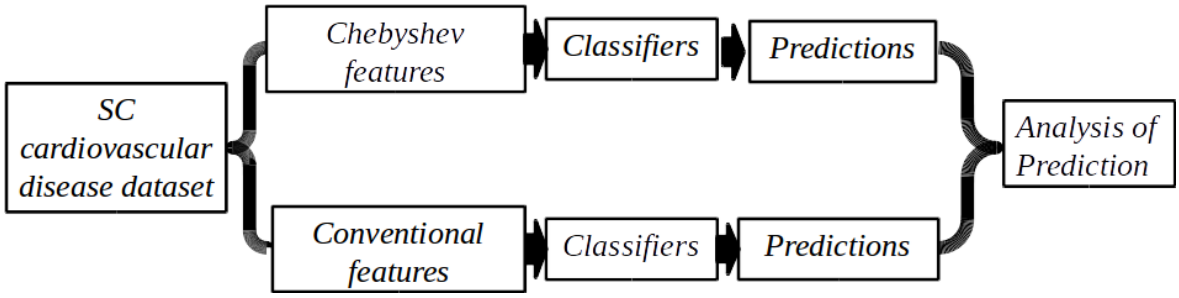



Figure 3.11: The methodology of comparing the performance of the Chebyshev features with different conventional features.

Table 3.17: The PSNR value computed between the original and the ECG signal reconstructed from the truncated Cheb coefficients.

| Data | PSNR(dB) | Data | PSNR(dB) |
|---|---|---|---|
| Normal_signal_1 | 60.755 | Abnormal_signal_1 | 48.872 |
| Normal_signal_2 | 58.010 | Abnormal_signal_2 | 48.718 |
| Normal_signal_3 | 57.695 | Abnormal_signal_3 | 48.033 |
| Normal_signal_4 | 56.662 | Abnormal_signal_4 | 47.938 |
| Normal_signal_5 | 54.642 | Abnormal_signal_5 | 47.332 |

### 3.3.5 Experiments and Results

The proposed work is evaluated using SC cardiovascular disease dataset. As a foot in the door of analysis the PSNR value is computed between the original and the reconstructed ECG signal from the truncated Chebfun coefficients. The PSNR value is computed for the validation of whether the truncated cheb coefficients are sufficient for capturing all the structural information of the ECG signal. Few samples of the PSNR values that are obtained from the computations are shown in Table 3.17.

The PSNR value of reconstructed ECG signal with reference to the original signal ranges from 47.332 to 60.755. From the PSNR value there is a clear picture portrayed that the reconstruction is proper for the ECG signal. The truncated Chebfun coefficient could capture all the required signal information. For the further validation for the proper retrieval of the structural information of the ECG signal, the signal is plotted as an image. The images of the original signal and its reconstruction with the Chebfun for the healthy and diseased cases are shown in Fig 3.12 and Fig 3.13 respectively.

From Fig 3.12 and Fig 3.13, it's clear that the Chebfun is able to capture all structural variation of the signal for healthy (Normal) and diseased (Abnormal) cases. For

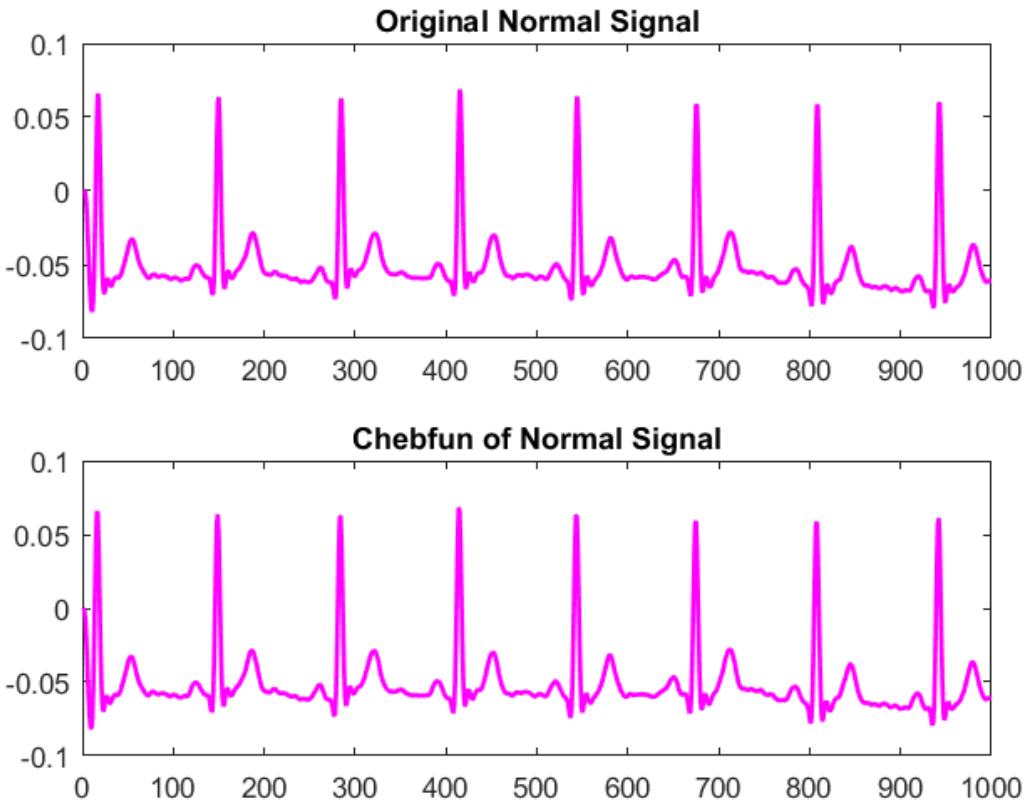



Figure 3.12: The original normal signal and the reconstructed signal using Chebfun coefficients.

the analysis of performance the proposed chebfun feature extraction technique is compared to the prevailed conventional feature extraction such as the Pan-Tompkin algorithm, Hamilton algorithm, Engelse and Zeelenberg algorithm (EngZee), Christov algorithm, stationary wavelet transform, and matched filter. The different models used for the analysis are SVM, Logistic Regression, Decision Tree, and AdaBoost. In a biomedical application scenario, validation of disease detection is unacceptable with accuracy measures alone. For a more accurate evaluation of the Chebyshev system, we have considered the accuracy, precision, recall, and specificity as evaluation metrics.

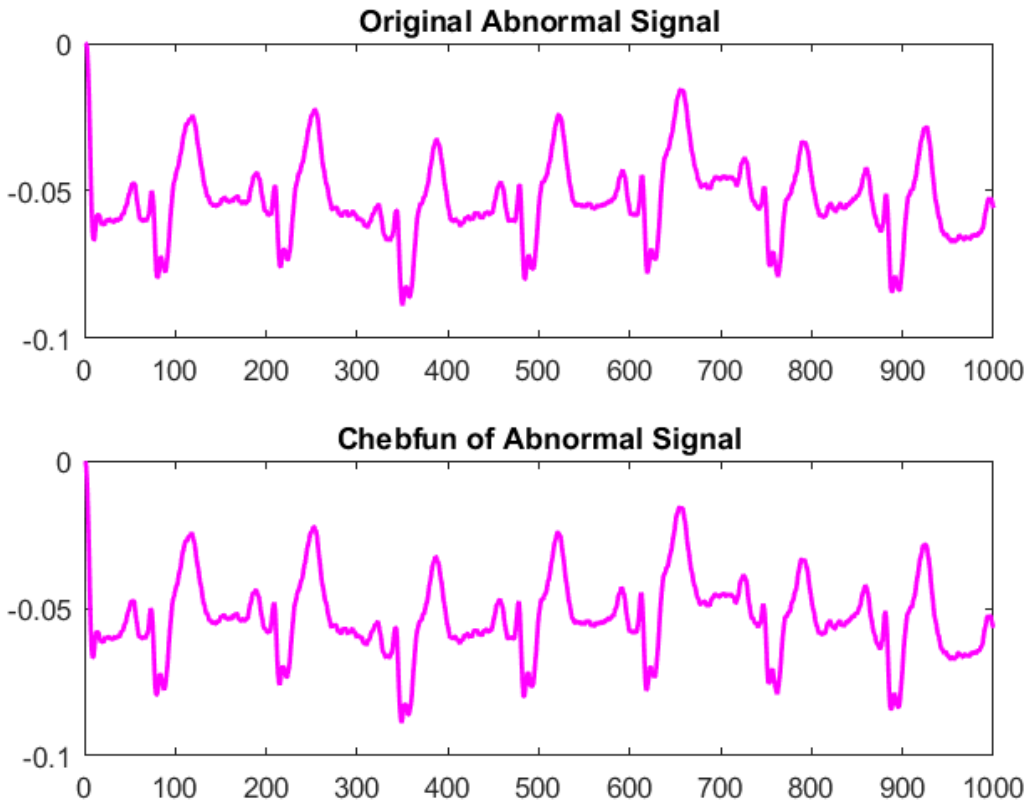



Figure 3.13: The original abnormal signal and the reconstructed signal using Chebfun coefficients.

Accuracy defines the overall performance. Sensitivity/Recall defines the ability of the model to detect the positive cases (diseased cases). Precision defines the ability of the model to detect the relevant positive cases. Specificity defines the ability of the model to detect the negative cases (healthy cases). Comparison of different conventional features with Chebyshev features for the detection of cardiac disease in SC data by using different machine learning algorithms in terms of accuracy (%), precision, recall, and specificity are shown in Table 3.18, Table 3.19, Table 3.20, Table 3.21 respectively.

From the tables we can infer that Chebyshev features using SVM classifiers have

Table 3.18: Comparison of different conventional features with Chebyshev features for the detection of cardiac disease in SC data by using different machine learning algorithms in terms of accuracy (%)

| | **Cheby-shev** | **Pan Tompkins** | **Hamil-ton** | **Chris-tov** | **Eng-Zee** | **Matched filter** | **SWT** |
|---|---|---|---|---|---|---|---|
| SVM | **81.1** | 59.5 | 62.2 | 59.5 | 70.3 | 67.6 | 64.9 |
| LR | 64.9 | **70.3** | 67.6 | 64.9 | **70.3** | 64.9 | **70.3** |
| Decision Tree | 59.5 | 59.5 | 59.5 | 45.9 | 67.6 | 62.2 | **73.0** |
| AdaBoost | 62.2 | 62.2 | 54.1 | 48.6 | 67.6 | 62.2 | **73.0** |

Table 3.19: Comparison of different conventional features with Chebyshev features for the detection of cardiac disease in SC data by using different machine learning algorithms in terms of precision

| | **Cheby-shev** | **Pan Tompkins** | **Hamil-ton** | **Chris-tov** | **Eng-Zee** | **Matched filter** | **SWT** |
|---|---|---|---|---|---|---|---|
| SVM | **0.800** | 0.333 | 0.333 | 0.00 | 0.750 | 0.600 | 0.500 |
| LR | 0.000 | **0.750** | 0.571 | 0.000 | **0.750** | 0.500 | **0.625** |
| Decision Tree | 0.400 | 0.429 | 0.400 | 0.182 | 0.556 | 0.467 | **0.615** |
| AdaBoost | **0.444** | 0.462 | 0.000 | 0.200 | 0.556 | 0.444 | **0.615** |

Table 3.20: Comparison of different conventional features with Chebyshev features for the detection of cardiac disease in SC data by using different machine learning algorithms in terms of recall/sensitivity

| | **Cheby-shev** | **Pan Tompkins** | **Hamil-ton** | **Chris-tov** | **Eng-Zee** | **Matched filter** | **SWT** |
|---|---|---|---|---|---|---|---|
| SVM | **0.615** | 0.154 | .077 | 0.00 | 0.231 | 0.231 | 0.077 |
| LR | 0.000 | **0.231** | 0.308 | 0.000 | **0.231** | 0.231 | **0.385** |
| Decision Tree | 0.308 | 0.462 | 0.308 | 0.154 | 0.385 | 0.538 | **0.615** |
| AdaBoost | **0.308** | 0.462 | 0.000 | 0.154 | 0.385 | 0.308 | **0.615** |

Table 3.21: Comparison of different conventional features with Chebyshev features for the detection of cardiac disease in SC data by using different machine learning algorithms in terms of specificity

| | **Cheby-shev** | **Pan Tompkins** | **Hamil-ton** | **Chris-tov** | **Eng-Zee** | **Matched filter** | **SWT** |
|---|---|---|---|---|---|---|---|
| SVM | **0.917** | 0.834 | 0.917 | 0.917 | **0.958** | 0.917 | **0.958** |
| LR | **1.00** | **0.958** | 0.875 | **1.000** | **0.958** | 0.875 | **0.875** |
| Decision Tree | 0.750 | **1.00** | 0.75 | 0.625 | 0.834 | 0.666 | **0.833** |
| AdaBoost | **0.792** | 0.708 | 0.834 | 0.666 | 0.834 | 0.792 | **0.792** |

outperformed all other conventional feature extraction techniques using all other classification algorithms in terms of accuracy, precision and recall. One other inference is that, in case of specificity even though the SVM classifier using Chebyshev features cannot outperform other conventional features, it has equivalent performance with other models.

For the further evaluation of the experiment we have included the confusion matrix. The confusion metric for the SVM model using Chebyshev features and conventional features are given in the Fig 3.14 From all this analysis and with reference to Fig 3.12

| **Predicted values** | | **Chebyshev** | | **Pan-Tompkins** | | **Hamilton** | | **Christov** | | **Engzee** | | **Matched filter** | | **SWT** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **0** | **22** | **2** | **20** | **4** | **22** | **2** | **22** | **2** | **23** | **1** | **22** | **2** | **23** | **1** |
| | **1** | **5** | **8** | **11** | **2** | **12** | **1** | **13** | **0** | **10** | **3** | **10** | **3** | **12** | **1** |
| | | **0** | **1** | **0** | **1** | **0** | **1** | **0** | **1** | **0** | **1** | **0** | **1** | **0** | **1** |
| | **Actual values** | | | | | | | | | | | | | | |

Figure 3.14: The confusion matrix of the SVM model predictions using the chebyshev features and conventional features..

and Fig 3.13. we can infer that Chebyshev polynomial approximation can capture all the

structural pieces of information of ECG signals which indicate the disease characters. In most cardiovascular diseases, the disease characteristics can be present in other portions of the signal such as Q, S, T waves along with R peak. In the case of other conventional features, they capture only the peak information [9] and so there is a chance of missing out other structural variation that indicates the diseases. In such cases it leads to the failure of the model to distinguish between normal and abnormal signals. The peculiar characteristics of the Chebfun which exploit the capacity to capture all the structural information of the disease in ECG signal helps the model to distinguish between the normal and the abnormal cases than the conventional features.

The proposed work is further extended to check the possibilities of chebyshev features for the analysis of the different tachycardia disease. The possibilities are evaluated on the AF dataset from the CINC challenge. For this analysis we artificially created a small sample dataset by taking only 50 samples from each normal and atrial fibrillation cases. The same procedure which we followed for extracting the Chebfun features from SC data are repeated for AF dataset. The minimum Chebfun coefficients are fixed to 2497 for every sample.

The minimum number of Chebfun coefficients required for the analysis is validated using the PSNR value computation. The samples of PSNR of the AF ECG samples are shown in Table 3.22.

From the PSNR values we could infer that the original ECG signal of the AF dataset is highly noisy. For the detailed analysis we have plotted the reconstruction of the normal and abnormal signal of the AF dataset. Fig 3.15 and Fig 3.16 shows the

Table 3.22: The PSNR value of the reconstructed signal using chebfun coefficients with reference to the original signal

| Data | PSNR | Data | PSNR |
|---|---|---|---|
| 1 | -15.086 | 6 | -42.856 |
| 2 | -15.892 | 7 | -43.234 |
| 3 | -15.977 | 8 | -44.928 |
| 4 | -16.506 | 9 | -47.12 |
| 5 | -16.715 | 10 | -51.653 |

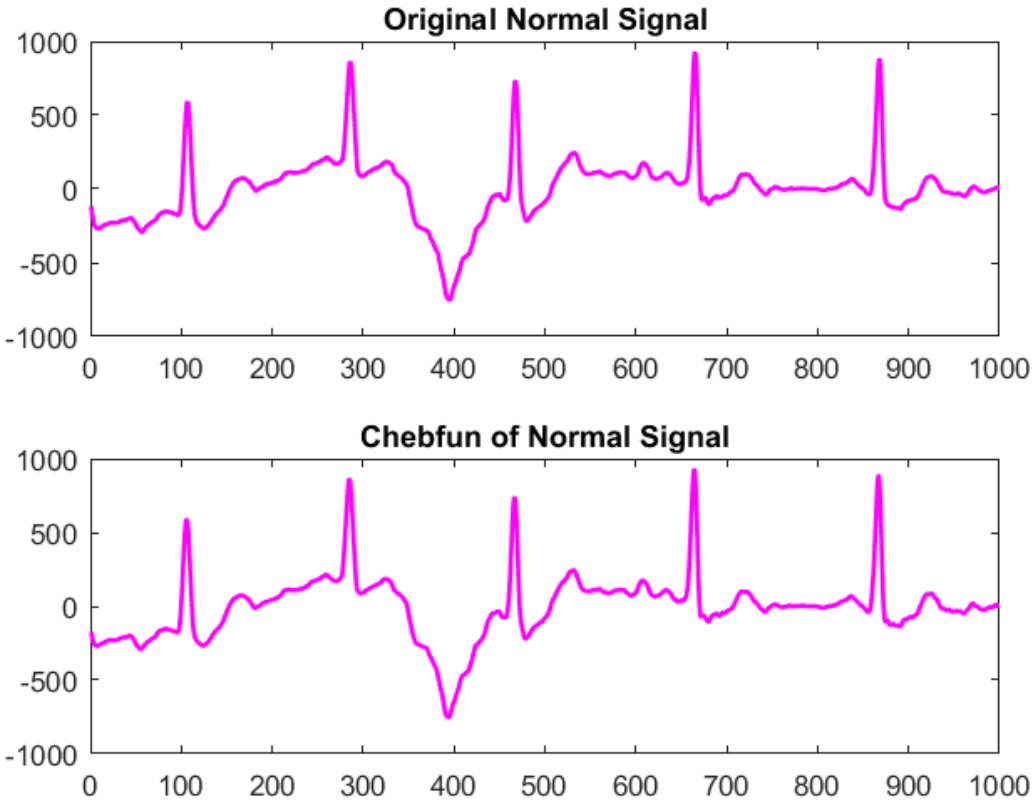



Figure 3.15: The original normal signal and the reconstructed signal using Chebfun coefficients for AF data.

reconstructed signal and its original signal.

The performance of the different machine learning models such as SVM,LR, Decision tree, adaboost using the chebyshev features and other conventional features are

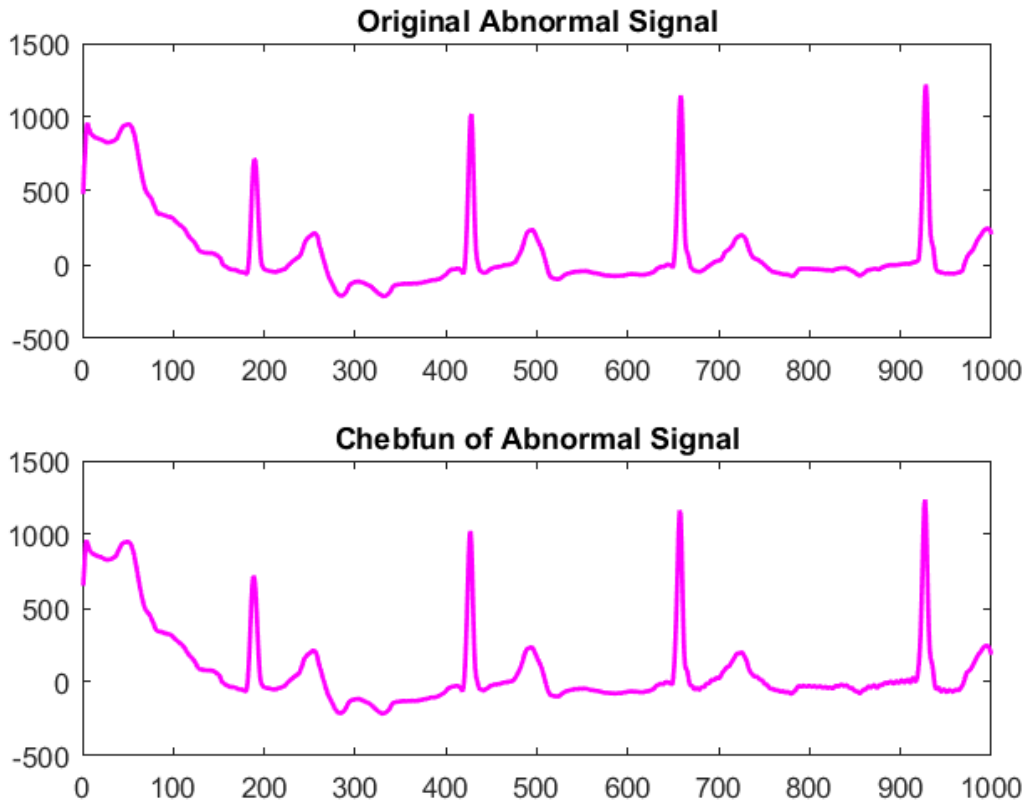



Figure 3.16: The original abnormal signal and the reconstructed signal using Chebfun coefficients for AF data

tabulated in Table 3.23, Table 3.24, Table 3.25 and Table 3.26.

From the results it's clear that the Chebfun couldn't outperform other conventional feature extraction techniques but gained approximately equal performance. The interpretation for this model performance is done by analysing the signal quality. Already from the PSNR value there is a clear picture that the signal is noisy. In the Fig 3.15 in the time stamp near 400 we could see that the signal is going down steeply which is not the case in an normal ECG signal. This shows that normal ECG signal of AF data is corrupted with the unwanted noises which may resemble the diseased condi-

Table 3.23: Comparison of different conventional features with Chebyshev features for the detection of cardiac disease in AF data by using different machine learning algorithms in terms of accuracy (%)

| | **Chebyshev** | **Pan Tompkins** | **Hamilton** | **Christov** | **Eng-Zee** | **Matched filter** | **SWT** |
|---|---|---|---|---|---|---|---|
| SVM | **51.00** | 64.00 | **68.00** | 30.00 | 60.00 | 55.00 | 64.00 |
| LR | 55.00 | **59.00** | **68.00** | 67.00 | **67.00** | 58.00 | **60.00** |
| Decision Tree | **71.00** | 69.00 | **72.00** | 66.00 | 65.00 | 56.00 | **58.00** |
| AdaBoost | **71.00** | 54.00 | 64.00 | 56.00 | 58.00 | 61.00 | **58.00** |

Table 3.24: Comparison of different conventional features with Chebyshev features for the detection of cardiac disease in AF data by using different machine learning algorithms in terms of precision

| | **Chebyshev** | **Pan Tompkins** | **Hamilton** | **Christov** | **Eng-Zee** | **Matched filter** | **SWT** |
|---|---|---|---|---|---|---|---|
| SVM | **0.510** | 0.640 | **0.661** | 0.308 | 0.609 | 0.547 | 0.621 |
| LR | 0.551 | **0.582** | **0.645** | **0.707** | **0.707** | 0.577 | **0.586** |
| Decision Tree | **0.684** | 0.694 | **0.775** | 0.638 | 0.653 | 0.560 | **0.565** |
| AdaBoost | **0.733** | 0.537 | 0.630 | 0.562 | 0.583 | 0.604 | **0.569** |

Table 3.25: Comparison of different conventional features with Chebyshev features for the detection of cardiac disease in AF data by using different machine learning algorithms in terms of recall/sensitivity

| | **Chebyshev** | **Pan Tompkins** | **Hamilton** | **Christov** | **Eng-Zee** | **Matched filter** | **SWT** |
|---|---|---|---|---|---|---|---|
| SVM | **0.520** | **0.640** | **0.740** | 0.320 | 0.560 | 0.580 | 0.720 |
| LR | 0.540 | **0.640** | **0.800** | **0.580** | **0.580** | 0.600 | **0.680** |
| Decision Tree | **0.780** | 0.680 | **0.620** | 0.740 | 0.640 | 0.560 | **0.700** |
| AdaBoost | **0.660** | 0.580 | **0.680** | 0.540 | 0.560 | 0.640 | **0.660** |

Table 3.26: Comparison of different conventional features with Chebyshev features for the detection of cardiac disease in AF data by using different machine learning algorithms in terms of specificity

| | **Cheby-shev** | **Pan Tompkins** | **Hamil-ton** | **Chris-tov** | **Eng-Zee** | **Matched filter** | **SWT** |
|---|---|---|---|---|---|---|---|
| SVM | **0.50** | **0.64** | **0.62** | 0.28 | 0.58 | 0.52 | 0.56 |
| LR | 0.56 | **0.54** | **0.56** | **0.62** | **0.76** | 0.56 | **0.52** |
| Decision Tree | **0.64** | 0.70 | **0.82** | 0.58 | 0.66 | 0.560 | **0.46** |
| AdaBoost | **0.76** | 0.50 | **0.60** | 0.58 | 0.70 | 0.58 | **0.5** |

tion. From the analysis we could interpret that the model may have confused between the normal and abnormal classes because of the presences of these irregularities and thus leading to small drop in performance in comparison with the conventional feature extraction technique such as Hamilton. In the case of the Hamilton feature detector it's the QRS complex features that they are more focussing on and thus neglecting the baseline features of the ECG signal which is not the case in Chebfun features. The Chebfun captures all the structural information of the ECG signal and thus misleading the model due to the noise signal. The Chebfun feature gained a better performance than other conventional feature extraction techniques in SC dataset. In case AF dataset the Chebfun features performance is slightly lower compared to Hamiltion but this can be due to the misleading irregularities of the AF dataset's ECG signals.

# Chapter 4

# Conclusion

In this work, we proposed the transfer learning approach to interpret the features learned by the model to detect different types of tachycardia diseases. This is attained by training the five different deep learning architectures with one of the tachycardia diseases (AF: TD1). The features learned by the model using one of the tachycardia diseases are tested with all other types of tachycardia diseases namely ventricular fibrillation and sinus tachycardia. The experimental results and analysis have shown that the RNN model performed better than other standard deep learning models such as LSTM, GRU, CNN, and RSCNN. The model was able to detect the atrial (AF: TD1) and ventricular type of tachycardia diseases (VF: TD2-A and VF: TD2-B) but failed in the case of ST (TD3). In the case of atrial fibrillation and ventricular fibrillation, it is the absence of P-wave and the presence of fibrillatory waves are the features that enabled the model to detect diseases distinct from the sinus tachycardia. The characteristic feature for sinus tachycardia is the upright P-wave, which the model failed to capture when trained with one of the types of tachycardia diseases called atrial fibrillation. Thus, the present work led to 'Explainable AI', that interprets the model

used to detect different types of tachycardia diseases, which have fast beat rhythm as a common characteristic of the input ECG signals. In case of ECG dataset with less samples we have analysed the possibilities of the Chebyshev feature extraction for cardiac disease detection. From the experiments and the result we could conclude that for the SC dataset the Chebyshev features extracted from the ECG signal using the SVM classifier have outperformed all other conventional feature extraction techniques such as Pan-Tompkins, Hamiltion, Engzee, Christov, matched filter and switched wavelet transform. The Chebfun gained an accuracy of 81.1%, precision of 0.800, recall of 0.615 and a specificity of 0.958. From the experiments it's also possible to infer that the Chebyshev coefficients could capture all the structural variations of the ECG signal for the better identifications of the normal and abnormal ECG signals. In case of AF data the Chebyshev features couldn't outperform other conventional feature extraction technique. This drop in performance in AF dataset can due to the irregularities in the normal ECG signal which when fed to the model makes less distinguishable between the abnormal signal.

The major findings of this work includes,

- This work interprets features learned by the model for the detection of tachycardia diseases leading to Explainable AI. Our main finding is: although all type of tachycardia diseases have fast beat rhythm as the common feature, there lies features specific to a disease. For example, sinus tachycardia, which is one of the types of tachycardia has positive upright P-wave, which was not captured

by the deep learning models trained with the ECG signals containing one of the tachycardia diseases, called atrial fibrillation (AF).

- This paper detects different types of tachycardia diseases using a single model by using the transfer learning approach.
- The proposed method finds the RNN is the best model for the detection of tachycardia diseases in different models considered.
- The chebyshev function performs better than other conventional feature extraction technique for SC data.
- The chebyshev function couldn't perform better for AF dataset. This can be due to the irregularities in normal signal of the AF data which may mislead the algorithms.

The future work that can be extended from the current work are

- To find more specific feature that is corresponding to the P-Wave and to do performance analysis based on that features
- To explore the possibilities of Chebyshev function for the interpretation of different classes in cardiac disease.
- To remove the different factors of noises and check the possibility of Chebyshev function.

- To remove the barrier of imbalance between the normal and abnormal class and explore the influence of that in cardiac disease detection

# References


1. Coleman, Elliott H., and Louis A. Soloff. "Incidence of significant coronary artery disease in rheumatic valvular heart disease." The American journal of cardiology 25.4 (1970): 401-404.

2. Buettner, Ricardo, and Marc Schunter. "Efficient machine learning based detection of heart disease." Proc. IEEE Healthcom (2019).

3. https://www.hrsonline.org/patient-resources/heart-diseases-disorders

4. Keating, Mark T., and Michael C. Sanguinetti. "Molecular and cellular mechanisms of cardiac arrhythmias." Cell 104.4 (2001): 569-580.

5. Acharya, U. Rajendra, et al. "A deep convolutional neural network model to classify heartbeats." Computers in biology and medicine 89 (2017): 389-396.

6. Acharya, U. Rajendra, et al. "Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network." Information sciences 405 (2017): 81-90.

7. Rajpurkar, Pranav, et al. "Cardiologist-level arrhythmia detection with convolutional neural networks." arXiv preprint arXiv:1707.01836 (2017).

8. Bär, Frits W., et al. "Differential diagnosis of tachycardia with narrow QRS complex (shorter than 0.12 second)." The American journal of cardiology 54.6 (1984): 555-560.

9. Eraut, Dennis, and David B. Shaw. "Sinus bradycardia." British heart journal 33.5 (1971): 742.

10. Kaplan, Benjamin M., et al. "Tachycardia-bradycardia syndrome (so-called "sick sinus syndrome"): pathology, mechanisms and treatment." The American journal of cardiology 31.4 (1973): 497-508.

11. Braunwald, Eugene. "Heart disease." (1988).

12. Kutlu, Yakup, Gokhan Altan, and Novruz Allahverdi. "Arrhythmia classification using waveform ECG signals." International conference on advanced technology & sciences. (2016).

13. Subramanian, Balambigai. "ECG signal classification and parameter estimation using multiwavelet transform." (2017).

14. Popa, Rustem, and Iuliana Cornei. "A method for acquisition and processing of electrophysiological signals on computer. The case of ECG signal." The Annals of "Dunarea de Jos" University of Galati (1996): 62-66.

15. Das, Resul, Ibrahim Turkoglu, and Abdulkadir Sengur. "Effective diagnosis of heart disease through neural networks ensembles." Expert systems with applications 36.4 (2009): 7675-7680.

16. Wang, Yang, et al. "A short-time multifractal approach for arrhythmia detection based on fuzzy neural network." IEEE Transactions on Biomedical Engineering 48.9 (2001): 989-995.

17. Martis, Roshan Joy, et al. "Computer aided diagnosis of atrial arrhythmia using dimensionality reduction methods on transform domain representation." Biomedical signal processing and control 13 (2014): 295-305.

18. Martis, Roshan Joy, et al. "Application of higher order statistics for atrial arrhythmia classification." Biomedical signal processing and control 8.6 (2013): 888-900.

19. Khadra, Labib, Amjed S. Al-Fahoum, and Saed Binajjaj. "A quantitative analysis approach for cardiac arrhythmia classification using higher order spectral techniques." IEEE Transactions on Biomedical Engineering 52.11 (2005): 1840-1845.

20. Li, Qiao, Cadathur Rajagopalan, and Gari D. Clifford. "Ventricular fibrillation and tachycardia classification using a machine learning approach." IEEE Transactions on Biomedical Engineering 61.6 (2013): 1607-1613.

21. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals Circulation. 101(23) (2003) :e215-e220.

22. Andreotti, Fernando, et al. "Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ECG." 2017 Computing in Cardiology (CinC). IEEE, (2017).

23. Kachuee, Mohammad, Shayan Fazeli, and Majid Sarrafzadeh. "Ecg heartbeat classification: A deep transferable representation." 2018 IEEE International Conference on Healthcare Informatics (ICHI). IEEE, (2018).

24. Tsipouras, Markos G., Dimitrios I. Fotiadis, and D. Sideris. "An arrhythmia classification system based on the RR-interval signal." Artificial intelligence in medicine 33.3 (2005): 237-250.

25. Saoudi, N., et al. "A classification of atrial flutter and regular atrial tachycardia according to electrophysiological mechanisms and anatomical bases. A Statement from a Joint Expert Group from the Working Group of Arrhythmias of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology." European heart journal 22.14 (2001): 1162-1182.

26. Sujadevi, V. G., K. P. Soman, and R. Vinayakumar. "Real-Time Detection of Atrial Fibrillation from Short Time Single Lead ECG Traces Using Recurrent Neural Networks." The International Symposium on Intelligent Systems Technologies and Applications. Springer, Cham, (2017).

27. Kiranyaz, Serkan, Turker Ince, and Moncef Gabbouj. "Real-time patient-specific ECG classification by 1-D convolutional neural networks." IEEE Transactions on

Biomedical Engineering 63.3 (2015): 664-675.

28. Gopika, P., et al. "Performance Improvement of Residual Skip Convolutional Neural Network for Myocardial Disease Classification." International Conference on Intelligent Computing and Communication Technologies. Springer, Singapore, (2019).

29. Shaffer, Fred, and J. P. Ginsberg. "An overview of heart rate variability metrics and norms." Frontiers in public health 5 (2017): 258.

30. Lepage, Ronan, et al. "ECG segmentation and P-wave feature extraction: application to patients prone to atrial fibrillation." 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol. 1. IEEE, (2001).

31. Greenwald, Scott David. "The development and analysis of a ventricular fibrillation detector". Diss. Massachusetts Institute of Technology,(1986).

32. SHEN, WIN-KUANG, et al. "Is sinus node modification appropriate for inappropriate sinus tachycardia with features of postural orthostatic tachycardia syndrome?." Pacing and Clinical Electrophysiology 24.2 (2001): 217-230.

33. F. Andreotti, O. Carr, M. A. F. Pimentel, A. Mahdi and M. De Vos, "Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ECG," 2017 Computing in Cardiology (CinC), Rennes, (2017): pp. 1-4.

34. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals (2003). Circulation.101.23 (2000): e215-e220.

35. M. Kachuee, S. Fazeli and M. Sarrafzadeh, "ECG Heartbeat Classification: A Deep Transferable Representation," 2018 IEEE International Conference on Healthcare Informatics (ICHI), New York, NY,(2018), pp. 443-444.

36. Shaffer, Fred, and J. P. Ginsberg. "An overview of heart rate variability metrics and norms." Frontiers in public health, 5 (2017): 258.

37. Gopika, P., Sowmya V et.al. "Transferable approach for cardiac disease classification using deep learning." Deep learning for biomedical and health informatics(BHI)(2020): 285.

38. T. Teijeiro, C. A. García, D. Castro and P. Félix, "Arrhythmia classification from the abductive interpretation of short single-lead ECG records," 2017 Computing in Cardiology (CinC), Rennes, (2017), pp. 1-4.

39. S. Datta et al., "Identifying normal, AF and other abnormal ECG rhythms using a cascaded binary classifier," 2017 Computing in Cardiology (CinC), Rennes, (2017), pp. 1-4.

40. M. Zabihi, A. B. Rad, A. K. Katsaggelos, S. Kiranyaz, S. Narkilahti and M. Gabbouj, "Detection of atrial fibrillation in ECG hand-held devices using a random

forest classifier," 2017 Computing in Cardiology (CinC), Rennes, (2017), pp. 1-4.

41. S. Hong et al., "ENCASE: An ENsemble ClASsifiEr for ECG classification using expert features and deep neural networks," 2017 Computing in Cardiology (CinC), Rennes, (2017), pp. 1-4.

42. Kamaleswaran, Rishikesan, Ruhi Mahajan, and Oguz Akbilgic. "A robust deep convolutional neural network for the classification of abnormal cardiac rhythm using single lead electrocardiograms of variable length." Physiological measurement 39.3 (2018): 035006.

43. Plesinger, Filip, et al. "Parallel use of a convolutional neural network and bagged tree ensemble for the classification of Holter ECG." Physiological measurement 39.9 (2018): 094002.

44. Kropf, Martin, et al. "Cardiac anomaly detection based on time and frequency domain features using tree-based classifiers." Physiological measurement 39.11 (2018): 114001.

45. Rizwan, Muhammed, Bradley M. Whitaker, and David V. Anderson. "AF detection from ECG recordings using feature selection, sparse coding, and ensemble learning." Physiological measurement 39.12 (2018): 124007.

46. Gopika, P., et al. "Performance Improvement of Residual Skip Convolutional Neural Network for Myocardial Disease Classification." International Conference

on Intelligent Computing and Communication Technologies. Springer, Singapore, (2019):pp. 226–234.

47. Sujadevi V.G., Soman K.P., Vinayakumar R. "Real-Time Detection of Atrial Fibrillation from Short Time Single Lead ECG Traces Using Recurrent Neural Networks." Intelligent Systems Technologies and Applications. Cham: Springer International Publishing, (2018), pp. 212–221.

48. Go, Alan S., et al. "Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the Anticoagulation and Risk Factors in Atrial Fibrillation (ATRIA) Study." Jama 285.18 (2001): 2370-2375.

49. Ganesan, Anand N., et al. "Long-term outcomes of catheter ablation of atrial fibrillation: a systematic review and meta-analysis." Journal of the American Heart Association 2.2 (2013): e004549.

50. McManus, David D., et al. "A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation." Heart Rhythm 10.3 (2013): 315-319.

51. Pan, Jiapu, and Willis J. Tompkins. "A real-time QRS detection algorithm." IEEE transactions on biomedical engineering 3 (1985): 230-236.

52. Desai, Usha, et al. "Machine intelligent diagnosis of ECG for arrhythmia classification using DWT, ICA and SVM techniques." 2015 Annual IEEE India Confer-

ence (INDICON). IEEE, (2015): pp.1–4

53. Kalidas, Vignesh, and Lakshman Tamil. "Real-time QRS detector using stationary wavelet transform for automated ECG analysis." 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE). IEEE, (2017):pp.457–461

54. Kiranyaz, Serkan, Turker Ince, and Moncef Gabbouj. "Real-time patient-specific ECG classification by 1-D convolutional neural networks." IEEE Transactions on Biomedical Engineering 63.3 (2016): 664-675.

55. Xia, Yong, et al. "Detecting atrial fibrillation by deep convolutional neural networks." Computers in biology and medicine 93 (2018): pp.84-92.

56. Gowri, B. Ganga, K. P. Soman, and D. Govind. "Improved Epoch Extraction from Telephonic Speech Using Chebfun and Zero Frequency Filtering." Interspeech. (2018): pp.2152–2156

57. Mohan, Neethu, and K. P. Soman. "Power system frequency and amplitude estimation using variational mode decomposition and Chebfun approximation system." 2018 twenty fourth national conference on communications (NCC). IEEE, (2018):pp.1–6

58. Pan, Jiapu, and Willis J. Tompkins. "A real-time QRS detection algorithm." IEEE transactions on biomedical engineering 3 (1985): 230-236.

59. Hamilton, Pat. "Open source ECG analysis." Computers in cardiology. IEEE, (2002):101–104

60. Engelse, W. A. H., and C. Zeelenberg. "A single scan algorithm for QRS-detection and feature extraction." Computers in cardiology 6.1979 (1979): 37-42.

61. Christov, Ivaylo I. "Real time electrocardiogram QRS detection using combined adaptive threshold." Biomedical engineering online 3.1 (2004): 28.

62. Kalidas, Vignesh, and Lakshman Tamil. "Real-time QRS detector using stationary wavelet transform for automated ECG analysis." 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE). IEEE, (2017): 457–461.

63. Śmigiel, Sandra, and Tomasz Marciniak. "Detection of QRS Complex with the Use of Matched Filtering." Innovations in Biomedical Engineering. Springer, Cham, (2017):310-322.

64. Platte, Rodrigo B., and Lloyd N. Trefethen. "Chebfun: a new kind of numerical computing." Progress in industrial mathematics at ECMI 2008. Springer, Berlin, Heidelberg, (2010):pp.69-87.

65. Rocha, B.M., Filos, D., Mendes, L., Serbes, G., Ulukaya, S., Kahya, Y.P., Jakovljevic, N., Turukalo, T.L., Vogiatzis, I.M., Perantoni, E. and Kaimakamis, E. An open access database for the evaluation of respiratory sound classification algorithms. Physiological measurement 40.3 (2019): pp.035001.

66. Chen, Gang. "A gentle tutorial of recurrent neural network with error backpropagation." arXiv preprint arXiv:1610.02583 (2016).

# List of Publications based on this research work

1. Sanjana K, Sowmya V, Gopalakrishnan E A, and Soman K P , *Performance Improvement of Deep Residual Skip Convolution Neural Network for Atrial Fibrillation Classification* , 8th International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA 2020) , Springer (Presented)

2. Sanjana K, Sowmya V, Gopalakrishnan E A, and Soman K P *Explainable AI for Heart Rate Variability in ECG Signal*, Journal Paper on Healthcare Technology Letter, IET digital library (First revision submitted on June 16, 2020.)

3. Sanjana K, B.Ganga Gowri, Sowmya V, Gopalakrishnan E A, and Soman K P *Detection of Cardiac Disease for Less Number of ECG samples using Chebyshev Coefficients*,Book chapter, Efficient Data Handling For Massive Internet Of Medical Things- Healthcare Data Analytics, Book series: Springer- Internet of Things - Technology, Communications and Computing (Abstract accepted,Full chapter submission, Dead line:July 15th)

# DEEP LEARNING APPROACHES FOR BIOMEDICAL IMAGE SEGMENTATION

A THESIS

*Submitted by*

**VYSHNAV M T**
**(CB.EN.P2CEN18021)**

*in partial fulfillment for the award of the degree of*

**MASTER OF TECHNOLOGY**
**IN**
**COMPUTATIONAL ENGINEERING AND NETWORKING**


**Center for Computational Engineering and Networking**

**AMRITA SCHOOL OF ENGINEERING**

**AMRITA VISHWA VIDYAPEETHAM**

COIMBATORE - 641 112 (INDIA)

**June - 2020**

# AMRITA SCHOOL OF ENGINEERING
# AMRITA VISHWA VIDYAPEETHAM
COIMBATORE - 641 112


# BONAFIDE CERTIFICATE

This is to certify that the thesis entitled **"DEEP LEARNING APPROACHES FOR BIOMEDICAL IMAGE SEGMENTATION"** submitted by **VYSHNAV M T (Register Number - CB.EN.P2CEN18021),** for the award of the **Degree of Master of Technology** in the **"COMPUTATIONAL ENGINEERING AND NETWORKING"** is a bonafide record of the work carried out by him under our guidance and supervision at Amrita School of Engineering, Coimbatore.

| | | |
|---|---|---|
| Dr. Sowmya V<br>Project Guide<br>Assistant Professor (Sr.Gr.), CEN | Dr. E. A. Gopalakrishnan<br>Project Co-Guide<br>Assistant Professor (SG), CEN | Dr. K. P. Soman<br>Project Co-Guide<br>Head, CEN |
| Mr. Vijay Krishna Menon<br>Project Co-Guide<br>Assistant Professor (Sr.Gr.), CEN | Mr. Sajith Variyar V. V.<br>Project Co-Guide<br>Assistant Professor, CEN | |

*Submitted for the university examination held on ... 20/6/2020 ...*

**INTERNAL EXAMINER** **EXTERNAL EXAMINER**

# AMRITA SCHOOL OF ENGINEERING

# AMRITA VISHWA VIDYAPEETHAM

COIMBATORE - 641 112

**CENTER FOR COMPUTATIONAL ENGINEERING AND NETWORKING**

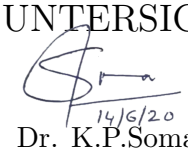
## DECLARATION

I, **VYSHNAV M T (Register Number - CB.EN.P2CEN18021),** hereby declare that this thesis entitled **"DEEP LEARNING APPROACHES FOR BIOMEDICAL IMAGE SEGMENTATION"**, is the record of the original work done by me under the guidance of **Dr. Sowmya V**, Assistant Professor (Sr.Gr.), **Dr. E. A. Gopalakrishnan**, Assistant Professor (SG), **Mr. Vijay Krishna Menon**, Assistant Professor (Sr.Gr.), **Mr. Sajith Variyar V. V**, Assistant Professor, **Dr. K.P.Soman**, Professor and Head, Centre for Computational Engineering and Networking, Amrita School of Engineering, Coimbatore. To the best of my knowledge this work has not formed the basis for the award of any degree/diploma/ associateship/fellowship/or a similar award to any candidate in any University.

**Place:** Coimbatore **Signature of the Student**

**Date:** 12/6/2020

COUNTERSIGNED



Dr. K.P.Soman
Professor and Head
Center for Computational Engineering and Networking

# Contents

# Acknowledgement

First of all, I thank Almighty for the immeasurable blessings for smooth completion of my project. I am immensely pleased to express my sincere obligation to my project guide **Dr. Sowmya V**, Assistant Professor (Sr.Gr.) and my project co-guides **Dr. E. A. Gopalakrishnan**, Assistant Professor (SG), **Mr. Vijay Krishna Menon**, Assistant Professor (Sr.Gr.), **Mr. Sajith Variyar V. V**, Assistant Professor and **Dr. K.P.Soman**, Professor and Head, Centre for Computational Engineering and Networking, Amrita School of Engineering, Coimbatore, for their valuable guidance, dedication and encouragement for the successful completion of this project.

I am grateful to the entire staff of CEN, and research scholars of the Department for their timely cooperation. I avail this opportunity to thank my friends for their whole-hearted support in the difficult stages of the project.

I am greatly indebted to my loving family members for being the motivating forces behind the completion of this dissertation. Also I am grateful for their invaluable help, moral support and encouragement throughout the course of study.

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| SGCE | Single Cell Gel Electrophoresis Assay |
| DEP | Diethyl Phthalate |
| H2O2 | Hydrogen Peroxide |
| dpf | Days Post Fertilization |
| MM | Multiple Myeloma |
| RBCs | Red Blood Cells |
| WBCs | White Blood Cells |
| CNN | Convolutional Neural Network |
| RCNN | Regional Convolutional Neural Network |
| FCN | Fully Convolutional Network |
| RPN | Region Proposal Network |
| ROI | Region of Interest |
| YOLO | You Only Look Once |
| SSD | Single Shot MultiBox Detector |
| IOU | Intersection Of Union |
| AI | Artificial Intelligence |
| DNA | Deoxyribonucleic Acid |
| AIMS | Amrita Institute of Medical Sciences |
| CT | Computer Tomography |
| PET | Positron Emission Tomography |
| MRI | Magnetic Resonance Imaging |
| BMP | BIT MAP |
| DL | Deep Learning |
| Val | Validation |
| TP | True Positive |
| FN | False Negative |
| FP | False Positive |
| TN | True Negative |

# Abstract

The recent years has witnessed a substantial progress in the research of computer vision and its application in medical imaging. The automated computer vision techniques has become a vital component in several healthcare applications. Although there were several researches on medical imaging using classical image processing techniques, most of them were subjective to the threshold values or other parameter values chosen by the user. The deep learning based computer vision algorithms are data driven techniques that can capture any spatial or temporal variations in the data and hence are not subjective to user. The application of computer vision techniques for medical imaging can provide useful information for diagnosis and treatment. The AI based technology can assist doctors for better diagnosis, prescribe correct treatments and monitor the progress of several diseases. This work mainly focuses on exploring the effectiveness of deep learning based detection and segmentation algorithms for tasks such as detection of valid and invalid comets in a comet assay image and for the detection and segmentation of multiple myeloma cancer cells from bone marrow aspirate images. Single cell gel electrophoresis assay (SGCE) or comet assay is a most frequently used method to measure the degree of DNA damage in eukaryotic cells or tissues. The result of the comet assay is a set of comet images which is analysed to measure the level of DNA damage. Currently the comet assay image analysis is performed using semi-automatic and automatic open source softwares. Most of these softwares faces the problem of either higher false negative or higher false positive. The chapter 2 proposes a deep learning based object detection algorithm for detection of valid and invalid comets. The dataset is collected from AIMS Nano science department which has total of 3000 comet assay images for different concentrations of compound tested. From the experiments performed it is observed that the proposed method gave competing results with open

source softwares open comet and hicomet. Multiple myeloma cancer is caused by the abnormal growth of plasma cells in the bone marrow. The most commonly used method for diagnosis of multiple myeloma is Bone marrow aspiration, where the aspirate slide images are either observed visually or passed onto existing digital image processing software for the detection of myeloma cells. The chapter 3 explores the effectiveness of deep learning based object detection/segmentation algorithms such as Mask-RCNN and U-net for the detection of multiple myeloma. The manual polygon annotation of the current dataset is performed using VGG image annotation software. The deep learning models were trained by monitoring the train and validation loss per epoch and the best model was selected based on the minimal loss for the validation data. From the comparison results obtained for both the models, it is observed that Mask-RCNN has competing results than U-net and it addresses most of the challenges existing in multiple myeloma segmentation.

# Chapter 1

# Introduction

Computer vision is an interdisciplinary field of research that focuses on extracting meaningful descriptors of objects from digital images or videos. It can be interpreted as a methodology where the computers can learn high level information from real world objects by mimicking the human visual system. The computer vision includes several stages such as acquiring, processing, analysing and understanding of images, which is then used to extract high-level information from images. The computer vision has a wide range of applications such as autonomous vehicles, medical diagnosis, geological surveying, satellite imaging, quality control and other basic scientific research problems. Furthermore, the computer vision has potential application to all problems where the images plays an important role. The most commonly used computer vision techniques include edge detection, shape description, region growing, image detection and segmentation.

The recent years has witnessed a rapid growth in researches in the field of computer vision and medical imaging. This is made possible with the abundant availability of

data and advances in the computing power. With the advances in the Deep Learning (DL) methodologies, it has become a prominent framework for data analysis in medical imaging. The deep learning models has the ability to learn complex non-linear relations in the data and can capture any spatial or temporal variations present. The use of DL algorithms has proven to be highly flexible and effective in medical imaging tasks such as Brain Imaging, Ophthalmology, Pulmonology, Pathology and Cardiology. The advent of Convolutional Neural Networks (CNNs) has made a huge impact in the field of computer vision and image data analysis. The CNNs have the capability to extract informative, translational invariant features from the images. The Deep CNNs are used as backbone networks for feature extraction and finally outputs a pixel-wise classification or regression information for several tasks such as object detection or image segmentation. Some of the most commonly used architectures are the Regional Convolutional Neural Network (RCNN) family [59, 60, 58, 54], and U-Net [55]. The use of computer vision techniques in healthcare can help doctors diagnose the patients to make faster and better decisions. Also, it can reduce the misdiagnosis rate substantially, as any misdiagnosis in the medical field can cause loss of money and most importantly the life of patients. Furthermore, it can help in early detection of fatal illnesses such as cancer with high certainty and thus be useful for timely treatment and can save countless lives in the long run.

One of the most common challenge for any medical imaging task is the availability of large number of annotated datasets and is considered a limitation for the develop-

ment of new applications in the field. This is because, the annotation of medical data requires many hours of strong attention by medical experts to provide pixel-level annotation for several tasks. The current work mainly focuses on the application of deep learning based detection and segmentation algorithms on various microscopic medical images to automate the diagnosis by expert doctors, clinicians and pathologists.

## 1.1 Literature

The use of automated computer vision based detection and segmentation algorithms can effectively reduce the valuable time of expert doctors and other trained personals to provide better diagnosis and treatment. The most commonly used medical imaging techniques for computer vision applications are x-ray, computer tomography (CT), ultrasound, MRI and positron emission tomography (PET). Some of the recent works in medical imaging using computer vision includes [1, 2] where they have proposed a CNN based image registration for the volume registration of CT and MR images. The works [3, 4] has shown that the networks when trained with a combination of model-labelled and human-labelled data outperforms the original architecture. The authors of [5] have used a combination of CNN and handcrafted features and applied classical machine learning for the classification of prostate lesions. While [6, 7] has shown how the networks can be trained for segmentation with higher-level annotation than pixel-wise labels. A review on semi-supervised learning approaches applied on medical imaging was done by [8]. Ghafoorian et al. [9] have made a study on the optimal

ways that can be used to apply transfer learning on MRI image segmentation tasks. The authors of [10, 11] has proposed a model where the shape priors are learned using auto-encoders and is then combined with CNNs as regularizers for several tasks such as cardiac MRI, ultrasound segmentation , kidney ultrasound segmentation, and super resolution. Agrawal et al. [12] have used encoded anatomical information as input to the network for modelling bone landmarks. The local shape descriptors are obtained by geodesic patch extraction method and is trained using Siamese CNN network.

Some of the other works has explored on the integration of data from different modalities for a single task. Liu et al. [13] have proposed a multi branch CNN that will combine data from different modalities such as brain patches, patient information for brain disease diagnosis. Similarly, [14] performed brain tumor segmentation from MRI data by splitting a CNN network into various channels with sparse interaction for different MRI modalities. Ge et al. [15] used a CNN network for skin disease identification with multimodal inputs such as macroscopic mole images and dermatoscopy images. While [16] have proposed an ensemble learning technique for lung nodule classification by integrating different inputs. Dai et al. [17] have suggested a solution for combining patient textual information in CNN along with fundus images for diabetic retinopathy detection.

## 1.2 Objectives

The main objectives of this work are:

- To detect valid and invalid comets in a comet assay image using deep learning based object detection algorithm.
- To detect and segment plasma cell or myeloma cell from bone marrow aspiration microscopic images using deep learning based object detection algorithm.

## 1.3 Thesis Overview

The rest of the thesis is organized as follows. Chapter 2 presents the work on the detection of valid and invalid comets from a comet assay image using deep learning based object detection algorithm. This chapter includes the introduction, literature, methodology and experimental results of the work on comet assay detection. In chapter 3, the detection and segmentation of multiple myeloma cells from bone marrow aspirate images using deep learning based detection and segmentation algorithms are explained in detail. This chapter includes the introduction, literature, methodology and experimental results of the work on multiple myeloma detection. Finally chapter 4 briefly describes the conclusions of both the works done and their respective future works.

# Chapter 2

# Comet Assay - Detection OF Valid Comets Using Object Detection Algorithm

## 2.1 Introduction

Image processing techniques applied to medical images can extensively reduce the effort and time of doctors, medical practitioners and scientists in the field of genetics. It is important to preserve the DNA structure in cells to effectively transfer the genetic information to the upcoming generations. Most of the new-born diseases like cancer, immune deficiencies are caused by alteration in the DNA structure. Chemical and radiation genotoxins destruct the genetic information in the cells making it important to measure the level of DNA damage [18]. The SGCE (single cell gel electrophoresis) assay or comet assay is a sensitive and effective method in measuring the degree of DNA damage in eukaryotic cells or tissues [19, 20]. In SGCE, cells treated with a DNA damaging agent are lysed and loaded onto an agarose gel, which is then stained with a dye. The electrophoresis releases the segments of the broken strand of the damaged

DNA in the surroundings outside of the cell. The appearance of the damaged DNA becomes similar to the tail of a comet, whose head is the nucleus of each single cell. The comet tail formation is obtained due to the migration of the damaged DNA strands towards the anode [21]. As the dose of the DNA damaging agent increases, the comet head grows dimmer and the tail grows longer and brighter. The comet assay images are obtained using a camera attached to the optical microscopes and DNA damage is quantified by analysing the images.

## 2.2 Literature survey

Damage analysis techniques for comet assays include visual scoring and digital image processing methods. Initially, the damage analysis of DNA was based on visual perception of the researchers and this approach is prone to human error as it involves grouping of comets based on the levels of damage observed. A software that analyses the images overcomes these difficulties by automating the process of quantification on DNA damage by measuring the parameters such as Head intensity, Tail intensity, Head DNA percentage, Tail DNA percentage, Tail moments. The availability of such software as open source is limited. CASPLab [22] is one such semi-automatic software which accepts only TIFF images and requires manual selection of comets and threshold. CASPLab localizes the head and tail of the comet based on the selected threshold and measures the damage parameters. CometScore [23] overcomes the manual selection of comets but requires the selection of threshold value. CometScore software fails to measure highly damaged comets as it considers the diameter of the head is the same as the height of the

comet and also restricts itself to BMP images. The OpenComet [24] is a fully automated software which validates the comet based on convexity ratio and symmetricity check. This limits the resolution of damage detection which inturn leads to false negatives i.e., missing the asymmetric comets and heavily damaged comets. In hicomet [25], the detection of heavily damaged comets (aptosis) is included using adaptive thresholding technique and further classification of comets as normal,necrosis and aptosis is implemented. Hicomet classifies the comets based on Histogram of Gradients (HOG) features extracted with 4 different classifiers such as Support Vector Machine(SVM),Neural Network(NN),Adaboost and Classification and Regression Tree(CART). Though hicomet detects all the types of comet, it also includes some of the non-comets as comets which results in false positives. Apart from the open source software, there exist other classification algorithms for damage analysis of the comet assays. De Souza et al. [26] experimented various machine learning approaches based on features extracted from the images to classify the degree of damage of the comets. Another image based classification approach by Namuduri et al. [27] quantifies the DNA damage of the comet assays using deep transfer learning approach. The DNA damage analysis is a combination of detecting valid comets, damaged comets and measuring the damage parameters.

The existing open source algorithms for detecting the valid comets are based on traditional approaches and the types of the comet are not explicitly given except hicomet [25]. Furthermore, quantification of damaged comets is implemented with the help of intensity profile of detected comets. As the existing approaches are limited to sym-

metric and less-damaged comets from the assays, deep learning approaches to detect, classify and measure the comets as individual modules is proposed in this paper. The main goal of this paper is to provide a data driven approach in order to analyse the comet assays and to assist on research experiments. One of the recent work on damage analysis of DNA cells has recommended the use of Convolutional Neural Networks (CNN) architectures with sufficient data as future work [26]. In this paper, the valid comets are detected using object detection algorithm. Faster R-CNN [28] is the latest technique in the family of R-CNN based object detection algorithms [29, 30] and it is less time consuming than its previous versions Fast R-CNN and RCNN. The main difference of Faster R-CNN with its previous versions is the use of a region proposal network (RPN) for generating region of interests. The main goal of this work is to provide a data driven approach to analyse the comet assays and detect valid and invalid comets with lesser false negatives and false positive, which in turn will improve the quantification of comets.

## 2.3 Problem Statement

From the literature, it is evident that most of the open source software's faces the problem of high false negatives or high false positives. Thus the detection of valid and invalid comets from a comet assay image is a crucial step towards comet assay analysis.

## 2.4 Objectives

The main objective of the current project is to detect valid and invalid comets from a comet assay image using deep learning based object detection algorithm, such that it reduces the number of false negative and false positives.

## 2.5 Motivation

As per the literature, most of these open source software's face the problem of having either high false negatives or high false positives. Figure 2.1 show the comparison of open source software's open comet and hicomet. Figure 2.1 (a) shows that open comet fails to detect a highly damaged comet, which is detected by hicomet. This shows the high false negative rate of open comet. Figure 2.1 (b) shows that hicomet detects invalid comets as comets, increasing the false positive rate. The proper detection of valid comets is thus a crucial step for applications of comet assay images. In this work, a deep learning based object detection algorithm is proposed for the detection of valid and invalid comets from a comet assay image, which reduces the false negative and false positives.

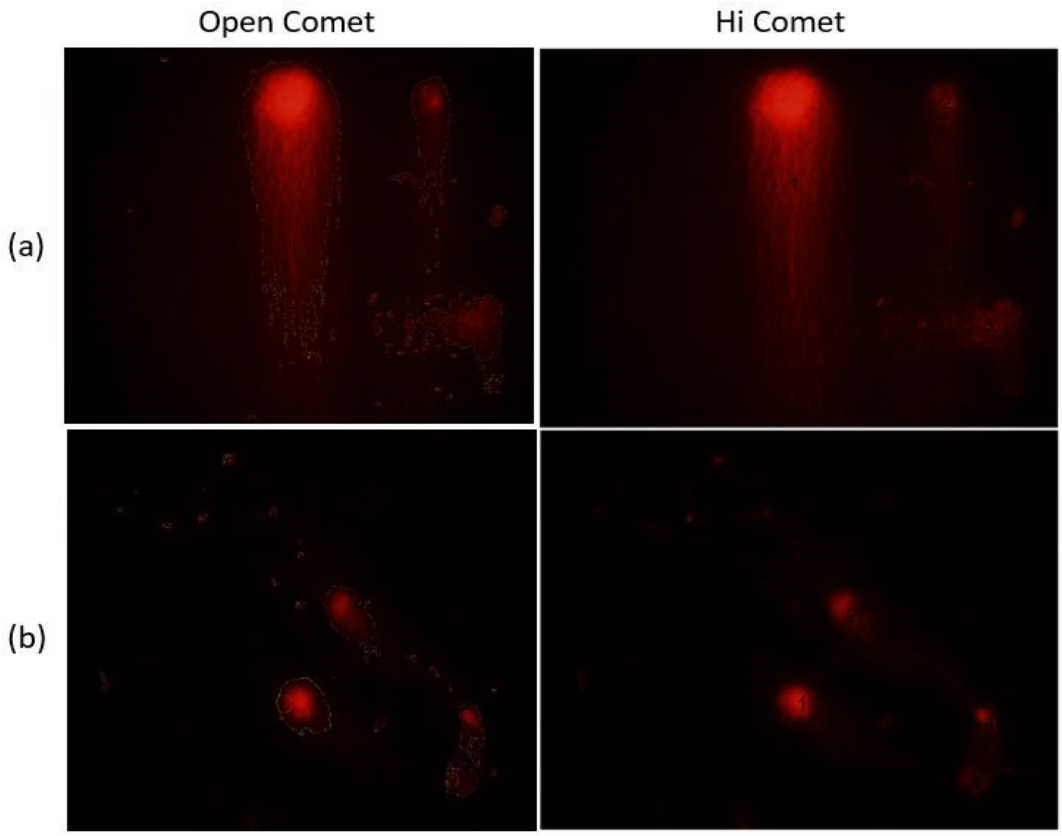



Figure 2.1: Comparison of Open comet and Hi comet results on comet assay images

## 2.6 Methodology

Literature has explored various image processing techniques for the detection of valid comets. These techniques are mostly based on some assumptions and threshold values, making it more vulnerable to errors. The current work acts as one of the modules of a complete comet assay tool that was developed for comet assay analysis. The tool has mainly three modules such as detection of valid comets using Faster RCNN, classification of valid comets using a classifier model and quantification of damaged comets using a key point detection model as shown in 2.2.

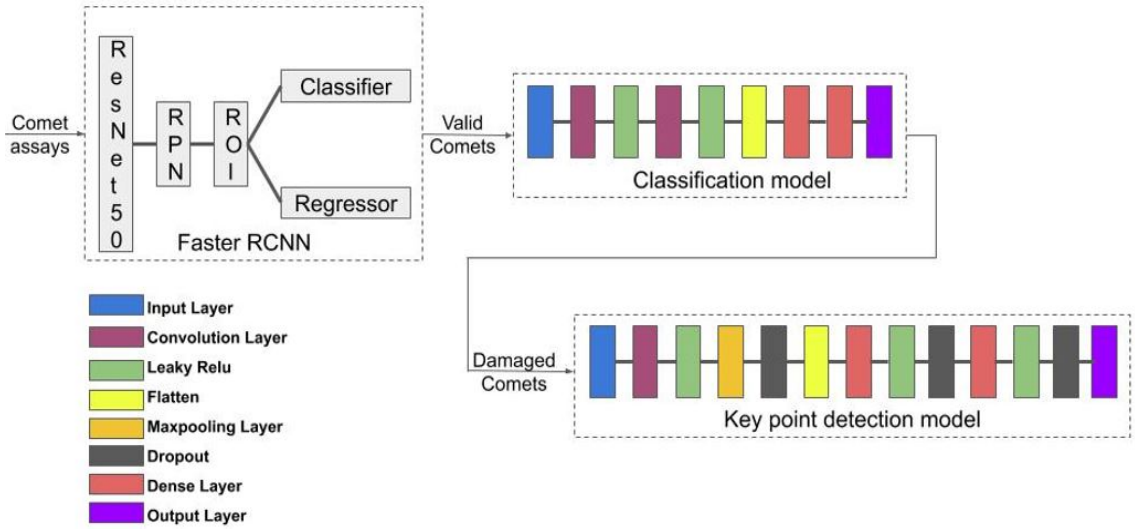



Figure 2.2: Block diagram of comet assay analysis tool

In this work the first module of the comet assay tool is explained in detail, which proposes a deep learning based objection detection algorithm namely Faster R-CNN [28] for detection of valid and invalid comets. Object detection is a computer vision task which identifies location, presence and type of one or more objects in a given image. The algorithm is a data driven algorithm which by itself learns different spacial and temporal patterns from the raw images without the help of any image pre-processing techniques, thus making it more reliable.

### 2.6.1 Faster-RCNN (Regional Convolutional Network)

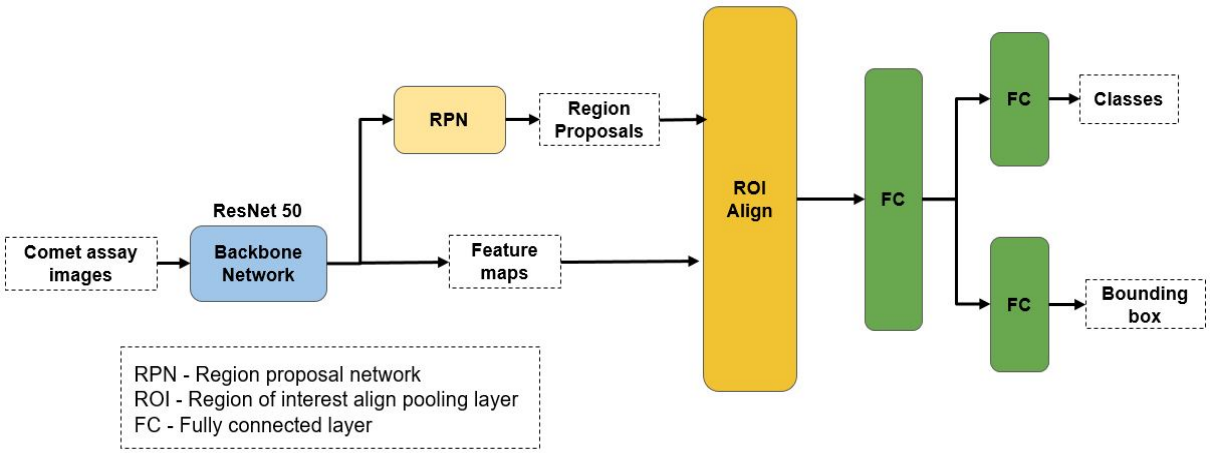



Figure 2.3: Block diagram for the proposed detection method using Faster-RCNN

In faster R-CNN, the input image is passed to a deep convolutional network which generates features maps of the image. The region proposal network acts on the feature maps to give the object proposal along with the objectness score. These object proposals are then resized to the same size by passing through a region of interest (ROI) pooling layer. Finally, the resized object proposals are used to classify and predict bounding box for objects. The classification is performed by a fully connected layer with softmax activation and bounding box are predicted by a regression layer on top of the softmax layer.

In this module, we have used the implementation of Faster R-CNN from [31], where the input to the object detection algorithm is the raw comet assay images and its corresponding bounding box annotation files. The bounding box of valid and invalid comets are annotated with the help of domain experts using ImageJ annotation tool [32]. The predicted results of the object detection algorithm will have corresponding classes and

bounding box coordinates of the detected comets.

The proposed model is experimentally chosen based on the performance on detection of valid comets. The results of the proposed model is then compared with the results of open source softwares such as open comet and hicomet based on the evaluation metrics true positives, false positives, false negatives and true negatives.

### 2.6.2 Image Annotation

Input to the object detection algorithm is the raw comet assay images and corresponding bounding box annotation files. The bounding box annotations for valid and invalid comets are done with the help of domain experts using Image J annotation tool [32]. The annotation files for each comet assay image is saved as a .txt file with the name of the image and has class label and the bounding box co-ordinates. The annotations has to be further formatted to Faster R-CNN input annotation format to input the algorithm. Figure 2.4 and figure 2.5 shows the sample image for manual annotation work done. The annotation has to be manually done using Image J tool by simultaneously comparing with ground truth annotations form the domain experts.

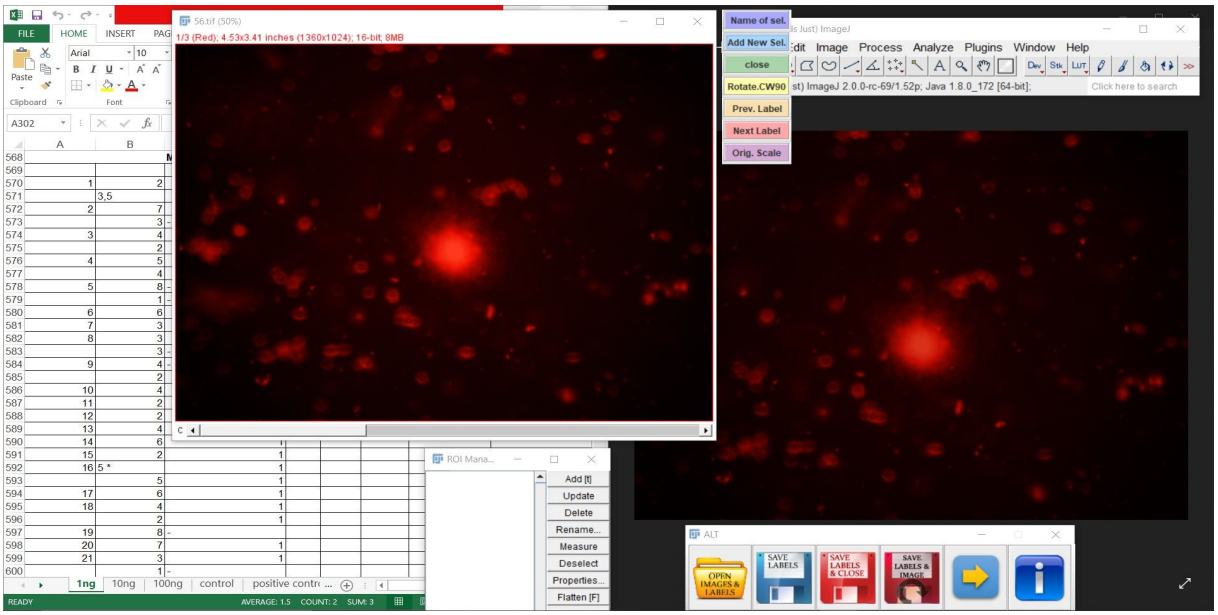


Figure 2.4: Illustrates the demo on manual annotation of comet assay images using Image J tool

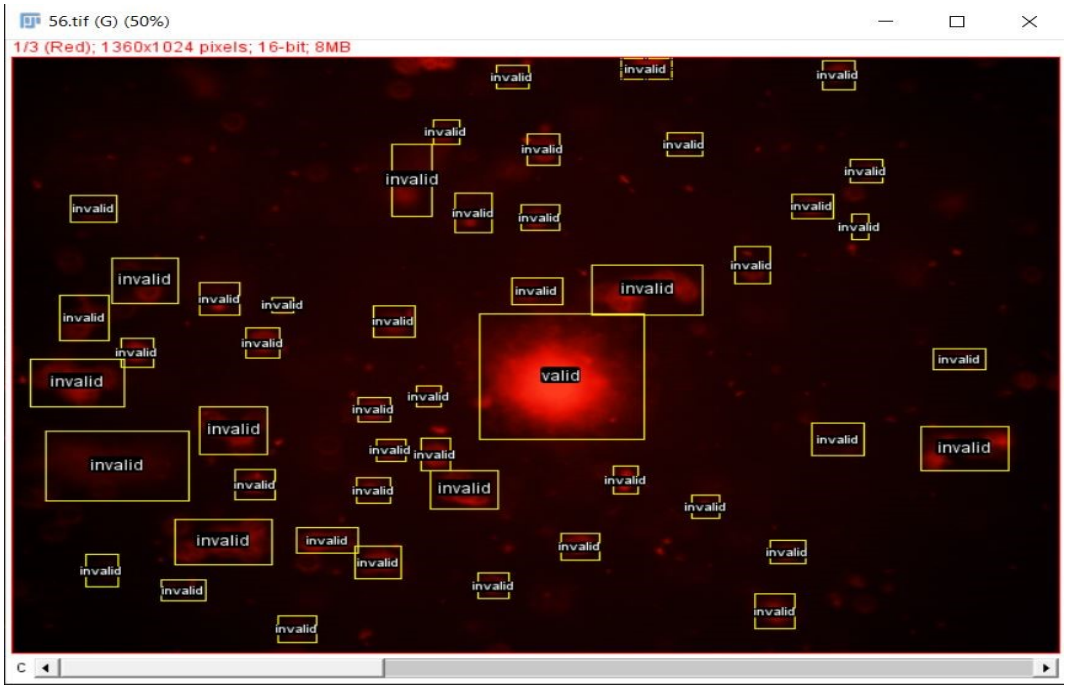



Figure 2.5: Illustrates the sample annotated image

## 2.7 Experiments Analysis and Discussions

### 2.7.1 Dataset Description

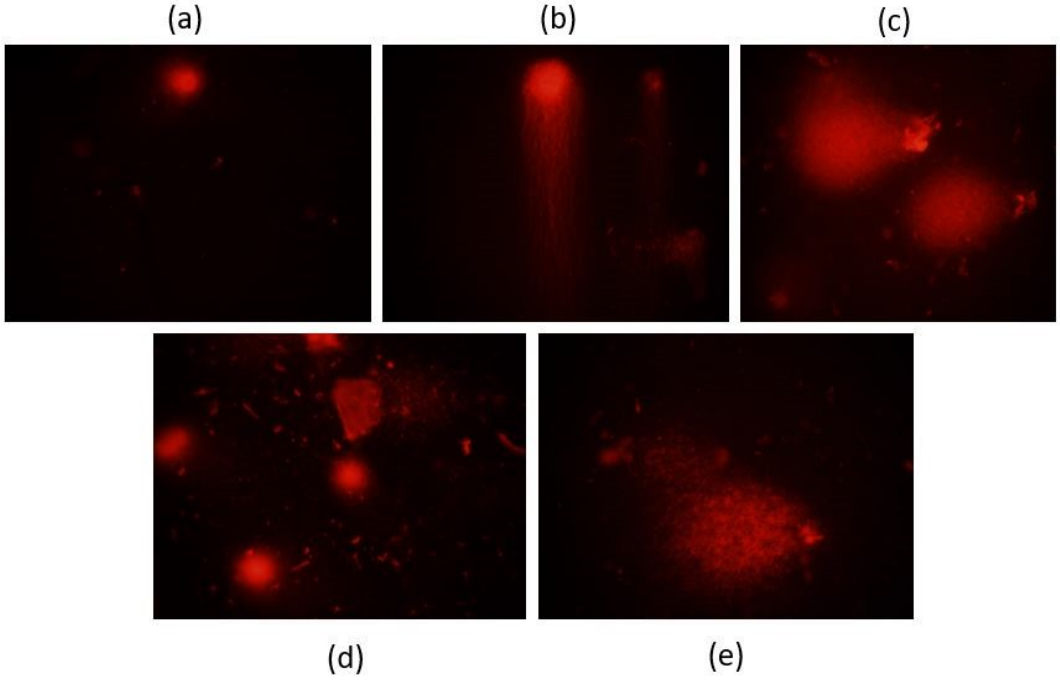



Figure 2.6: Describes dataset with different concentrations: (a) Control, (b) 1ng/ml, (c) 10ng/ml, (d) 100ng/ml, (e) Positive control

Dataset was obtained from AIMS Nano science Department. The comet assay experiment was performed on a fish breed "Japanese Medaka". The eggs of the fishes were collected initially by breeding the fishes. The eggs were stored in water for larvae formation. The larvae formation happens at 7 days post fertilization (7dpf) and the experimentation was started at 18 dpf. First the larvae is treated with water (controls). Then the experimentation was carried out using different concentrations (1ng/ml, 10ng/ml, 100ng/ml) of the compound diethyl phthalate (DEP). Finally experimentation was done on hydrogen peroxide (H2O2 – positive control) which gives

maximum damage. The cells are then lysed and stained using ethyledium bromide dye. As per the experimentation, the dataset is divided into 5 folders namely control, 1ng/ml, 10ng/ml, 100ng/ml and positive control, each having 5 folders of male and female experimented cell images. For each concentration male and female cell folders were labelled as female1, female2, female3, female4, female 5 and male1, male2, male3, male4, male5. The dataset has total 3000 images with each concentration folder having 600 images. Figure 2.6 shows the sample images from each concentration folders of the dataset.

### 2.7.2 Experimental Results and Discussions

The dataset is divided into train and test set by taking 80/20 split from each of the 5 concentration folders, where 80% of the images from each concentration is taken for training (40 folders – 2400 images) and remaining 20% for testing (10 folders – 600 images). The Faster R-CNN model is trained by transfer learning approach [33] from scratch using pre-trained weights of a blood cell detection model [34, 35] as its initial weights. The faster R-CNN model is trained for 500 epochs and is validated on test set at every 100 epochs to select the proposed model.

The result analysis is performed on the test set folders at every 100 epochs of the trained model. Figure 2.7 (a) shows the sample result image for the test set folder positive control female 1. The green bounding box in the image shows the valid comets detected by the algorithm and blue bounding box shows the invalid comets detected by

the algorithm with corresponding probability values for each class.

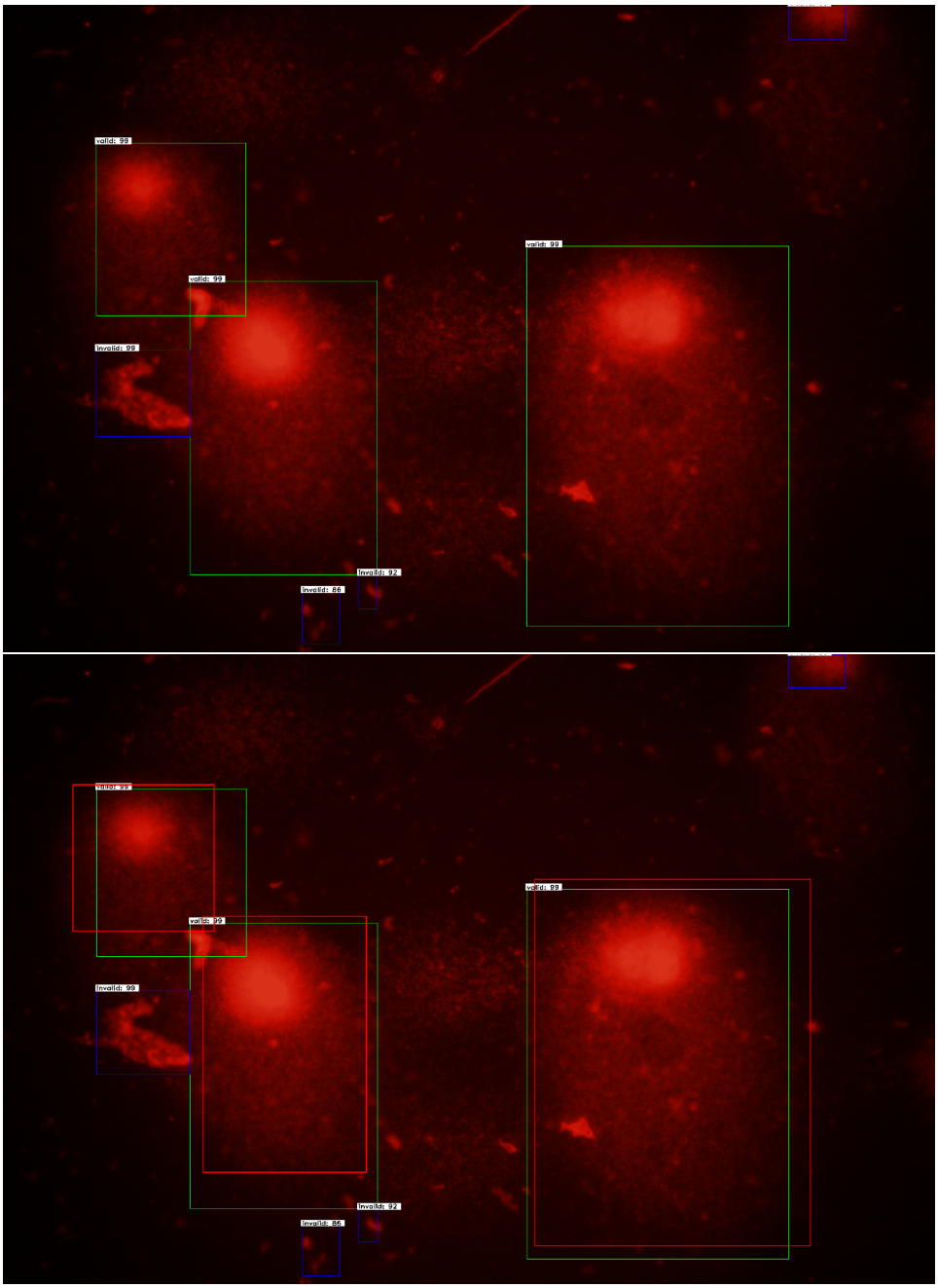


Figure 2.7: Illustration sample result image for the test set folder positive control female 1: (a) predicted result, (b) ground truth annotation drawn on predicted result.

Figure 2.7 (b) shows the corresponding ground truth annotation bounding box drawn on the resulting image. The red bounding box in the image indicates the ground truth annotation for valid comets. The comparison with ground truth annotation is done using Intersection of union (IOU) metric [36, 37]. IOU is a metric that evaluates the similarity of the predicted bounding boxes to the ground truth bounding boxes. It calculates the ratio of overlapped region to the original ground truth area. The IOU values range from 0 to 1. The best IOU threshold for the proposed work is selected based on the performance of the model for the test set for every 100 epochs. The best IOU threshold is set by comparing the test set results for each 100 epochs. The IOU calculation is done with only valid comets detected, as in a comet assay image everything other than valid ground truth comets are considered to be invalid comets. The performance of the experimented models are compared based on the analysis of the confusion matrix computed for different IOU values such as 0.5, 0.6, 0.7, and 0.8.

Figure 2.8 show the variation of evaluation metrics such as true positive (tp), false negative (fn), false positive (fp) and true negative (tn) for each 100 epochs for the test set folder positive control female 1. For all the epochs from 100-500 epochs, it is observed that false negative is very less and true positive is very high for an IOU of 0.5, whereas false positive and true negative values remains same for all the IOU values. The best IOU threshold value is hence set as 0.5.

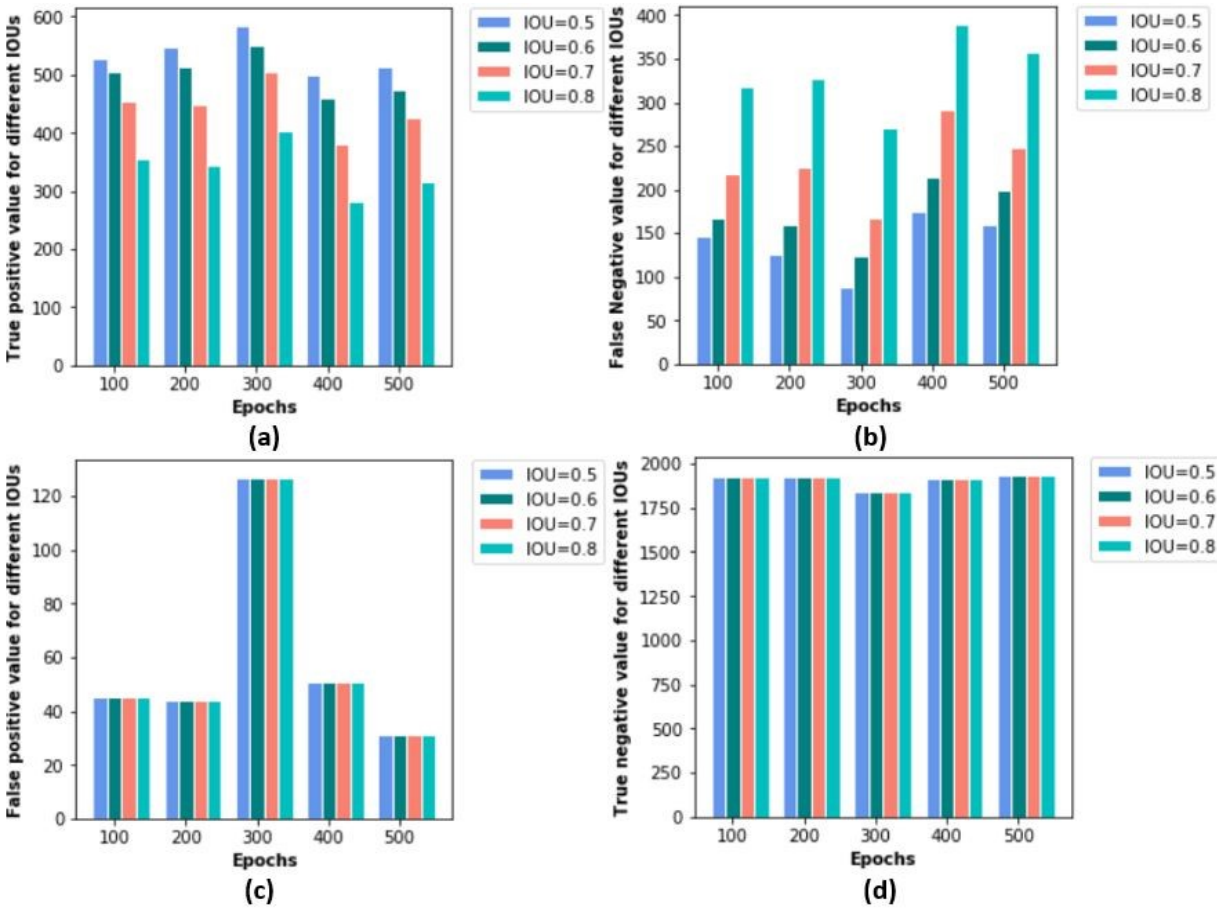



Figure 2.8: Comparison of experimented model on a)True positives b)False Negatives c)False positives and d)True negatives

The proposed model is selected by further analysis on the variation of evaluation metrics for each 100 epochs by fixing the IOU threshold value as 0.5. From figure 2.8, it is observed that number of false negatives is very less and true positives is very high at 300 epochs, but the number of false positives is relatively higher for 300 epochs. In the current problem of detection of DNA damage in cells, the model is considered to be efficient if the number of false negatives is lesser, irrespective of the false positive value. This is because, the model must not miss any of the valid comets, though

it detects the invalid as valid. The proposed model is hence selected to be at 300 epochs.

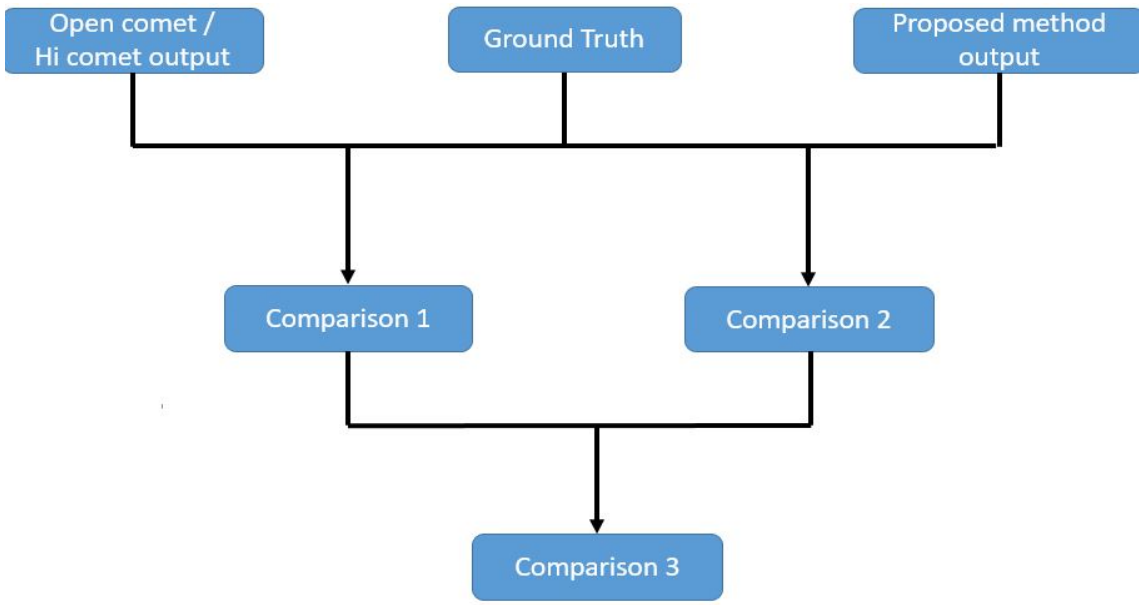



Figure 2.9: Block diagram for results comparison.

The results obtained from the selected model of IOU as 0.5 and epoch 300 is now compared with the ground truth annotations, detection by open comet and hicomet to show the efficiency and reliability of the proposed method. The result comparison is done for all the test set folders using confusion matrix. Figure 2.9 shows the block diagram of results comparison for the proposed method. In comparison 1, the confusion matrix for the open comet and hicomet outputs with respect to the ground truth annotations is computed for all the test set folders. In comparison 2, the confusion matrix for the proposed method output with respect to the ground truth annotations is computed for all the test set folders. Finally in comparison 3 to show the efficacy of the proposed method, the confusion matrices obtained from the comparison 1 and com-

parison 2 are compared based on the evaluation metrics true positive, false negative, false positive and true negative.

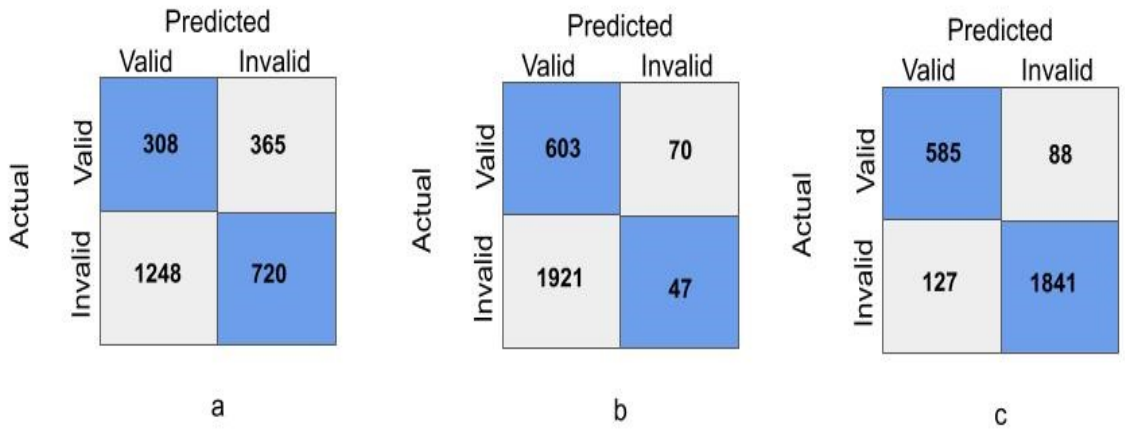



Figure 2.10: Confusion matrix of a) Open comet b) Hi comet and c) Proposed model

The comparison is carried on test set of 600 comet assays consisting of 673 valid comets and 1968 invalid comets. The confusion matrix of the proposed model, open comet and hicomet results are shown in figure 2.10. From figure 2.10, it is observed that true positives are high in confusion matrix of hicomet and true negatives are high in confusion matrix of the proposed model. Though hicomet has less false negatives and high true positives, the comets are detected irrespective of the scope of groundtruth valid comets and this may further lead to false measurements of valid comets. The proposed model overcomes the above limitation by considering the comets as valid only if it covers at least fifty percentage area of the groundtruth valid comets (IOU $> 0.5$). This makes the proposed model more reliable on validation of a comet. The false positives in the confusion matrix shows the detection of invalid comets as valid comets. The proposed model detects less false positives than open comet and hicomet. These

comets can be further validated based on the classification and quantification modules present in the comet assay analysis tool developed.

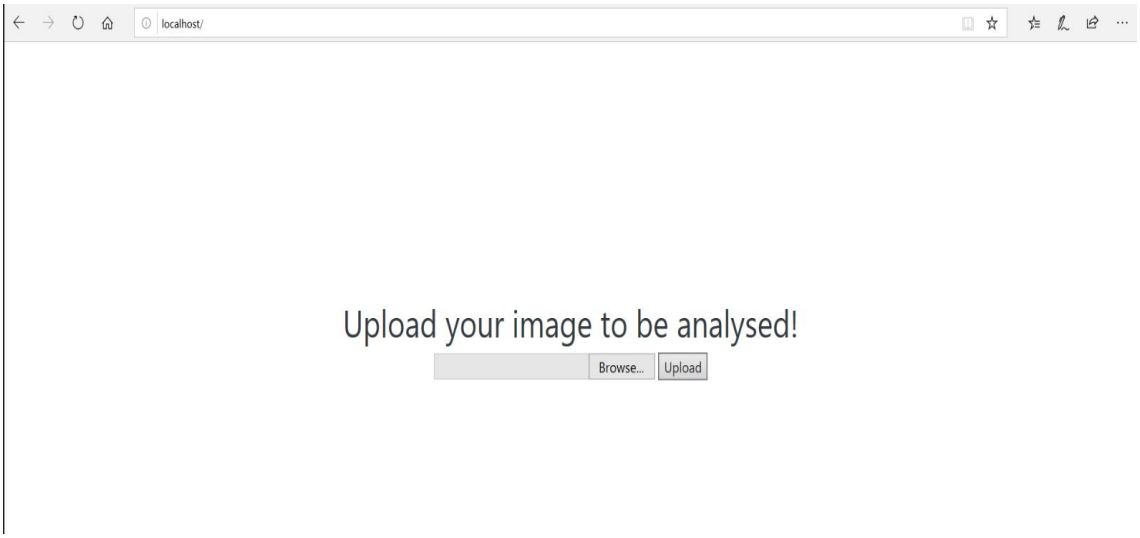



Figure 2.11: Online software tool for comet assay analysis - uploading image

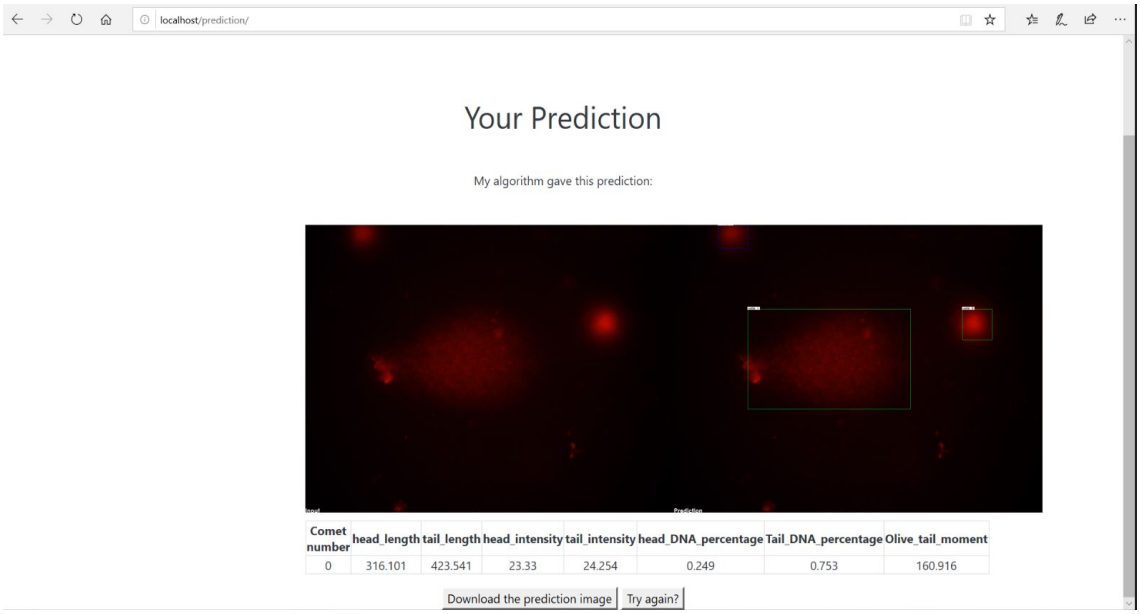



Figure 2.12: Prediction results of online software tool for comet assay analysis

Figures 2.11 and 2.12 shows the comet assay analysis tool that was developed as a web application. Figure 2.11 shows the web page where the comet assay image has to be loaded for the qualitative and quantitative analysis. The prediction results obtained from the comet assay analysis tool for the uploaded image are shown in figure 2.12.

## 2.8 Acknowledgement

I acknowledge Dr.Bindhu Paul, Assistant Professor, and team, Laboratory of Reproductive and Developmental Biology, Amrita Centre for Nano sciences & Molecular Medicine, Amrita Institute of Medical Sciences, Amrita Vishwa Vidyapeetham, Kochi for the access to the comet assay data set on analysis of compound diethyl phthalate and for help in manual annotation of the data set.

# Chapter 3

# Deep Learning Based Approach For Multiple Myeloma Detection

## 3.1 Introduction

Cancer is a condition when cell grows out of control. Multiple myeloma (MM) is a type of cancer caused by the abnormal growth of myeloma cells (plasma cells) in the bone marrow [38]. Bone marrow is a soft tissue in bones which is home to blood cells such as red blood cells (RBCs), platelets and various white blood cells (WBCs) that are essential for the immune system [39]. The B lymphocytes or B cells are a group of WBC present in the bone marrow that respond to any bacterial or viral infections in the body. As the B cells respond to infections, they mature and get converted into plasma cells. Plasma cells helps in fighting infections by producing antibodies that will prevent infection and diseases [40]. In Multiple myeloma, the abnormal plasma cells gets accumulated in the bone marrow and produce abnormal proteins which interferes with the normal production of blood cells in bone marrow. As the myeloma cells overcrowd the healthy blood cells, it reduces the count of RBC, WBC and platelets leading

to low immunity. These abnormal proteins generated by myeloma cells can also cause damage to kidneys and bones in the long run.

One of the recent studies has found that out of all cancers, 1% is caused by Multiple myeloma and contributes 2% to the cancer deaths. Multiple myeloma is considered to be highly curable when detected in early stages. The myeloma cells are usually distinguished from normal plasma cells based on their histologic and cytologic features [38]. One of the most common method used by pathologists for multiple myeloma diagnosis is bone marrow aspiration, where a needle is injected into the bone to extract blood samples which is then lysed onto a slide and stained. From the obtained aspiration slides, the percentage of plasma cells present in the bone marrow is estimated via microscopic observation for the diagnosis of multiple myeloma [41]. The revised guidelines for the diagnosing and managing MM is provided by [42]. Detection of myeloma cells from bone marrow aspirates include visual inspection and digital image processing methods. Even though advanced image processing techniques were recently introduced for cancer cell detection, the visual microscopic inspection is still considered as the golden standard for multiple myeloma detection. However, the manual microscopic method is time consuming and is highly subjective to human error. The limitations of the visual microscopic method can be overcome by the use of digital image processing methods, where the detection of myeloma cells in the bone marrow aspirate slides can be automated and further diagnostic decisions can be made by expert pathologists. This will reduce the diagnostic time of the patient as well as the workload on the pathologist experts.

## 3.2 Literature survey

Detection, segmentation and further classification of cells from microscopic images is a most recent research topic and has several approaches proposed by researchers using advanced image processing and machine learning techniques. Some of the recent works include [43] where the authors proposed an automatic cancerous cell detection method in which the segmentation of nucleus is performed by a fuzzy c-means clustering algorithm. The principal component analysis (PCA) is then applied to the rich set of features obtained from nucleus. Finally, classification was done by ensembling SVM classifiers for different parameters. Another work [44] studied the importance of coherence-controlled holographic microscopy using quantitative phase features for the classification of cancer tissue cells and observed that classification based on quantitative phase information outperformed classification based on morphometric features. While the authors of [45] have used shallow CNN followed by a recurrent layer for the detection of tuberculosis, intestinal parasite eggs and malaria in microscopic images. Meanwhile the authors of [46] have used classical machine learning algorithms for malaria parasite detection, where the images were divided into patches and feature were extracted from the patch images. In [47], the authors used a combination of image processing methods such as thresholding, k-means, and modified watershed algorithm for automatic segmentation of cells and nuclei from microscopic images. The algorithm consists of three stages including segmentation of WBCs, extraction of nuclei and separation of clustered cells. Whereas [48] proposes a new method for counting bone marrow cells which includes

localization of cells by color transformation and stepwise averaging method, then segmentation of nucleus and cytoplasm using ostu's method, region growing method, color weakening transformation and k-means algorithm. The clustered cells were separated using water shed algorithm followed by classification using support vector machines. The authors of [49] has given a survey on the different computer-aided methods available for segmenting blood smear images. In [50], the authors used a combination of typical features with three new features such as minimum thickness to minimum convex thickness of nucleus mask, perimeters of smaller nucleus after splitting, minimum thickness of nucleus mask to hausdorff distance of the nucleus and cell. The classification of chronic myeloid leukemia cells was then performed using a novel decision tree algorithm.

Even though there are several researches in the area of detection and segmentation of cell from microscopic images, the application of these approaches to MM specific task is very limited. The initial work on MM detection was proposed by Saeedizadeh et. al. [38], where contrast enhancement was initially performed on the images. The nucleus and cytoplasm were detected using thresholding. The connected cells were then splitted using bottleneck and watershed algorithms. Finally features are extracted from cytoplasm and nucleus such as nucleus eccentricity, nucleus-cell ratio and classification was performed using SVM classifier with a precision of 96.52%. Meanwhile [51], addresses color normalization of stains and cell segmentation for microscopic images and was an initial step for the development of an automatic plasma cell segmentation tool – PCSeg for microscopic images [52]. The PCSeg tool, initially performs a sta-

tistical modelling and computes separability index for the region of interests in the image and then removes the unstained cells. Further the nucleus and cytoplasm are extracted using a multiphase level set method and finally the cluster cells are segmented using watershed along with circular hough transform utilizing k-means based nuclei mask. Another recent work [53] introduced Convolutional neural network (CNN) for detection of MM. The authors applied median filter and contrast enhancement for the image pre-processing. Finally features are extracted from the images using CNN model Alexnet and classification into normal and blast cells was performed using SVM classifier.

The segmentation and classification of cells from microscopic images using advanced image processing and machine learning techniques has gained huge traction in building automated tools. The machine learning and deep learning algorithms learns morphological features of objects in the image which are invisible to human eyes. In this paper, the effectiveness of most commonly used deep learning based segmentation algorithms such as Mask-RCNN (regional convolutional neural network) [54] and U-net [55] on multiple myeloma detection and segmentation are explored. Moreover one of the recent works [56] have compared the performance of Mask-RCNN and U-net for a nuclei segmentation task.

## 3.3 Problem Statement

The diagnosis of multiple myeloma is currently performed by searching for myeloma cells in the bone marrow aspirate slides through a microscope [38]. The detection by visual inspection is a subjective and time-consuming task for pathologists. Most of the existing approaches are based on classical image processing techniques and is subjective to the thresholding values chosen. Moreover, the detection and segmentation of plasma cells is a challenging task due to following problems:

- Segmentation of plasma cells requires segmentation of both cytoplasm and nucleus.
- The color contrast of nucleus and cytoplasm and the cytoplasm with the adjacent background is less due to overstaining or understaining.
- Most of the plasma cells are clustered together such as nuclei-nuclei cluster, nuclei-cytoplasm cluster or cytoplasm-cytoplasm cluster.
- Microscopic image has more than one type unstained or stained cells.

The challenges specific to plasma cell segmentation mentioned above are shown in figure 3.1.

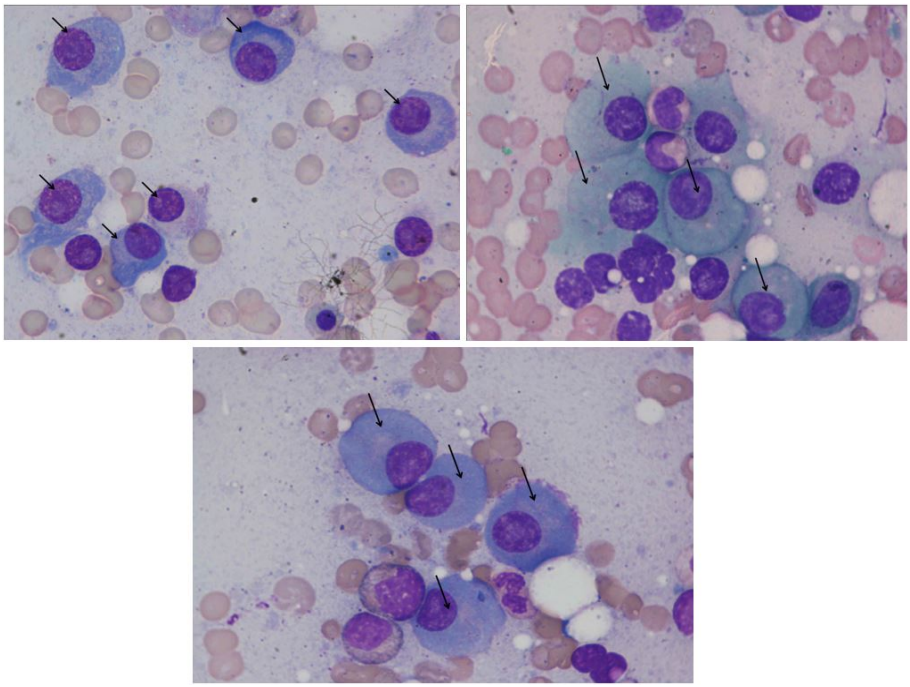


Figure 3.1: Challenges of plasma cell segmentation in Multiple myeloma dataset

## 3.4 Objectives

The main objective of the current project is to detect and segment plasma cell or myeloma cell from bone marrow aspiration microscopic images using deep learning based object detection algorithm.

## 3.5 Motivation

From the literature it is evident that most of the researches in the area of MM detection were based on conventional image processing techniques and to our knowledge so far

there is no work related to the MM detection from bone marrow aspirate images using deep learning based segmentation algorithms. Also the authors of [52] have quoted their future work as plasma cell detection and segmentation using deep learning methods. Furthermore, the Previous work in Chapter 2 on detection of valid and invalid comets from a comet assay image using object detection algorithm was able to detect most of the valid comets with respect to the groundtruth. Thus the literature and the competing results obtained on comet assay data gave us motivation to explore the performance of object detection algorithm on plasma cell detection.

## 3.6 Methodology

In this work, deep learning based segmentation algorithms namely Mask-RCNN and U-net are used for segmentation of myeloma cells from the bone marrow aspirate slides. The Mask-RCNN is an instance segmentation algorithm that identifies pixel wise boundary of objects belonging to a particular class and segments them into different instances, while the U-net algorithm performs semantic segmentation by segmenting all objects belonging to a class into single instance. However both the algorithms follow a data driven approach where the algorithm learns different spatial and temporal patterns from the image without relying on any conventional image processing techniques. The evaluation metrics used for selecting the best model are accuracy, precision, recall, specificity, negative predictive value, false discovery rate and f1-score.

### 3.6.1 Mask-RCNN (Regional Convolutional Network)

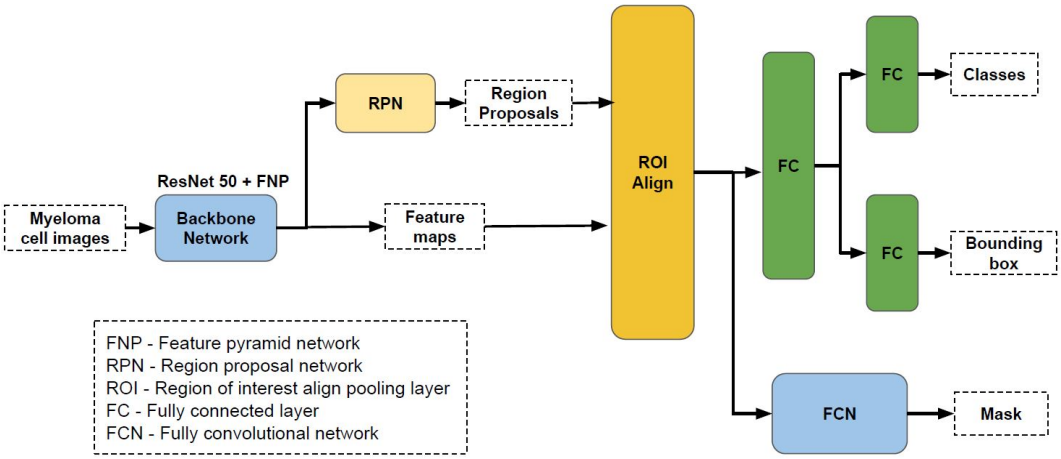



Figure 3.2: Block diagram for the proposed myeloma cell detection using Mask-RCNN

Mask-RCNN is a successor of Faster-RCNN [58] model and it is the latest in the family of R-CNN [59, 60]. The Mask-RCNN is build on top of Faster R-CNN framework, where it extends the Faster R-CNN with an extra branch for the object mask prediction which runs in parallel with the existing bounding box regression branch. The Faster-RCNN generates the class and bounding box coordinates for every objects in the image and a FCN (Fully Convolutional Network) generates pixel-wise masks for objects. In this work, the implementation of Mask-RCNN is taken from matterplot [61, 62]. Figure 3.2 explains the block diagram of the proposed myeloma cell detection using Mask-RCNN. The input image is initially passed to a deep convolutional network (backbone network - Resnet50 or Resnet101), which generate feature maps of the image. The RPN (Region Proposal Network) acts upon these feature maps to create object proposals along with objectness score. These object proposals are then passed through a Region of interest (ROI) pooling layer to resize all proposals to same size. Finally, the object

proposals are given to a fully connected layer that will classify and predict bounding box for the objects. The fully connected layer has a softmax layer for classification and a regression layer on top of it for bounding box prediction. The FCN then generates masks for the positive regions that are selected by classifier layer. The result predicted by Mask-RCNN will have corresponding class and polygon coordinates of the detected myeloma cells.

### 3.6.2 U-net

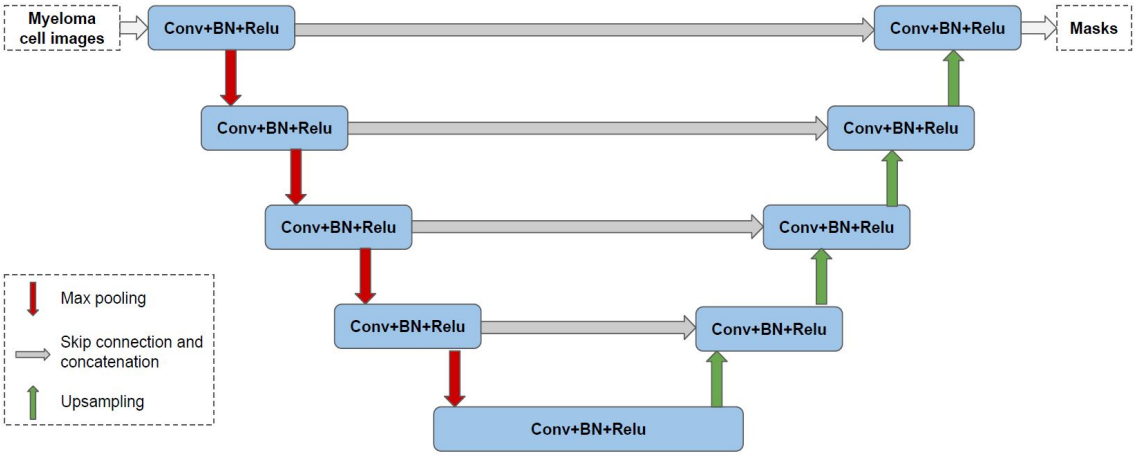



Figure 3.3: Block diagram for the proposed myeloma cell detection using U-net

The U-net algorithm is an end-to-end FCN that has proven its efficacy on biomedical image segmentation tasks. The U-net has a symmetric network with skip connections among each downsampling and upsampling paths with concatenation. The skip connections helps in passing the local information from downsampling to global information in upsampling which gives a combination of local and contextual information to predict

better segmentation. Figure 3.3 shows the block diagram of the proposed myeloma cell detection using U-net. The U-net mainly has 3 parts: (a) Encoder / downsampling path: which is a stack of convolution and max pool layers to capture the context in the image. This is usually a pretrained network (VGG/ResNet); (b) Bottleneck path: that has two convolutional layers with batch normalization and dropout. This path acts as a tunnel between the encoder and decoder; (c) Decoder/ upsampling path: which consists of transposed convolutions or deconvolution to enable precise localization. The image is upsized and is concatenated with corresponding image from encoder path. This is done to ensure that prediction is precise by combining information from previous layers. In this work the implementation of U-net is taken from [63].

### 3.6.3 Image Annotation

The input to the segmentation algorithms are the raw microscopic images of bone marrow aspirate slides and corresponding polygon annotations as a '.json' file. The polygon annotations for myeloma cells are done manually with reference to the annotation file provided along with the dataset using VGG Image Annotation software [57]. Figure 3.4 shows the sample images of manual annotation work done.

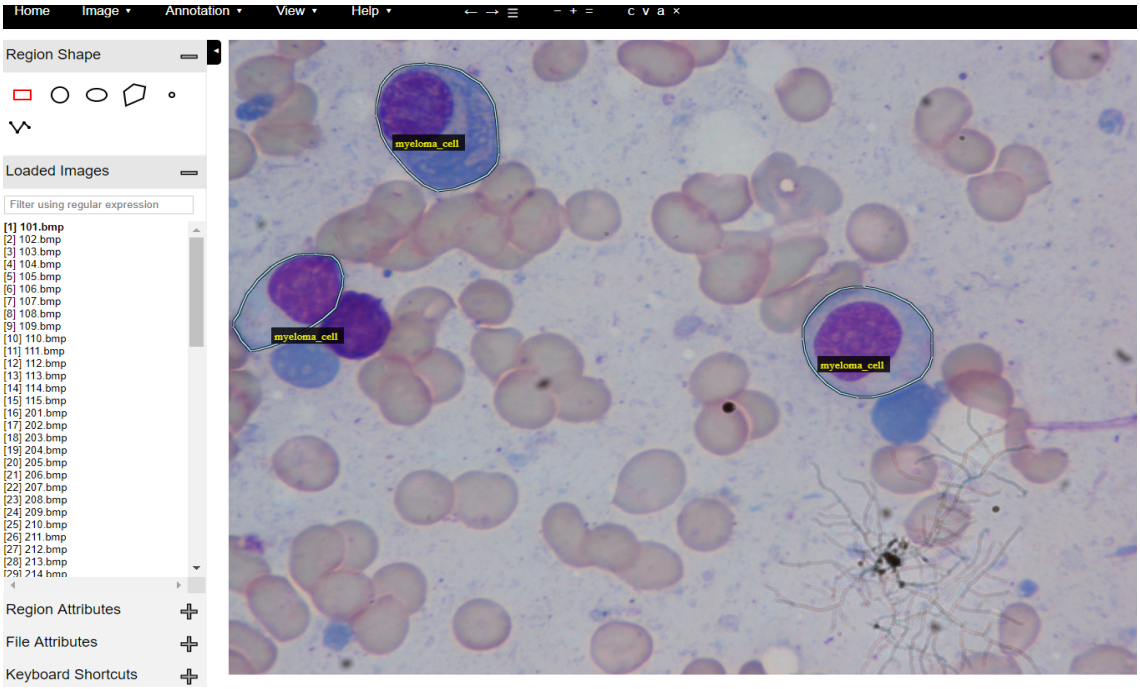



Figure 3.4: Manual Annotation of Myeloma cells using VGG Image Annotation software

## 3.7 Experiments Analysis and Discussions

### 3.7.1 Dataset Description

The current work uses dataset obtained from the cancer imaging archive (TCIA) [64, 65]. The dataset contains microscopic images of bone marrow aspirate slides captured according to the standard guidelines [66] from patients diagnosed as multiple myeloma. The bone marrow smears were prepared by staining with Jenner-Giemsa stain [51, 67]. The images were captured at 1000x magnification with the help of a Nikon eclipse-200 microscope equipped with digital camera. The dataset has a total of 85 microscopic images of 5 subjects diagnosed as multiple myeloma. All the images were of size 2560x1920

pixels saved as raw BMP format. The myeloma cells in each image slides of the dataset is annotated by expert pathologist with arrow marks indicating the abnormal plasma cells or myeloma cells and is provided as a pdf format file. Figure 3.5 shows sample images of the dataset with black arrow marks indicating the actual myeloma cells.

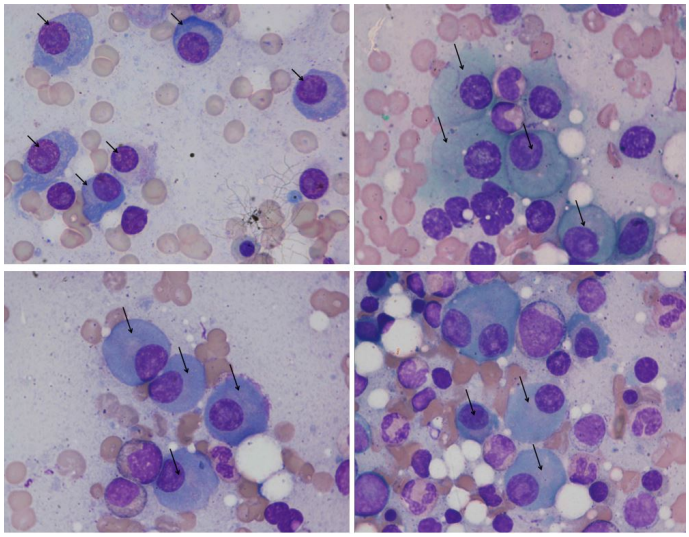


Figure 3.5: Sample images of the dataset

Table 3.1: Datasplit proportions used for the experiments

| **Datasplit Ratio** | **Number of images** | | |
|---|---|---|---|
| | **Train** | **Val** | **Test** |
| **60/40** | 51 | 14 | 20 |
| **70/30** | 60 | 10 | 15 |
| **80/20** | 68 | 7 | 10 |

For the current work, in order to identify the best datasplit ratio, experiments are performed on different splits of the dataset such as 60/40, 70/30 and 80/20 where 60%, 70%, 80% are for training and remaining 40%, 30%, 20% are split in the ratio of 60/40 into test and validation (val) sets. Table 3.1 gives the number of images in train/val/test sets for each of the datasplit ratios. The number of multiple myeloma cells and other

WBC's in test set of each split are 71 & 264 for 60/40 split, 58 & 229 for 70/30 split and 41 & 142 for 80/20 split respectively.

### 3.7.2 Experimental Results and Discussions

In this section, the different experiments performed on the dataset to choose the best trained model is explained in detail. The Mask-RCNN model is trained on different train/val/test data split proportions such as 60/40, 70/30 and 80/20. Since, the current dataset has lesser number of images, the models are trained by transfer learning approach [68] from scratch with initial weights as coco weights [69]. Further, image augmentation is also performed to boost the performance of the models.

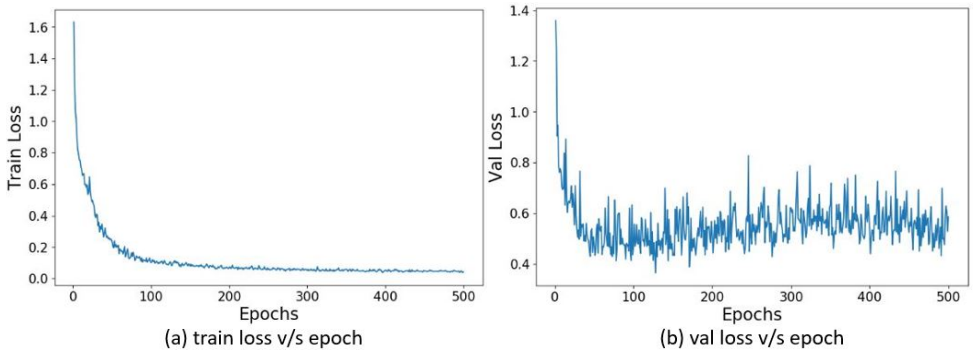



Figure 3.6: Loss v/s epoch graphs for 60/40 datasplit ratio using Mask-RCNN model

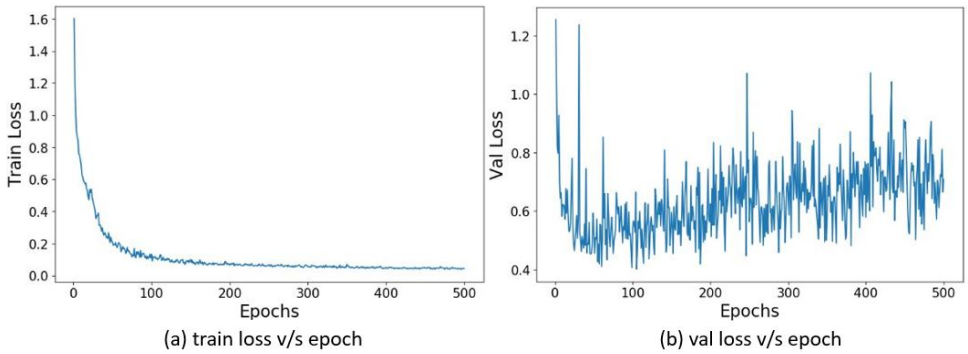



Figure 3.7: Loss v/s epoch graphs for 70/30 datasplit ratio using Mask-RCNN model

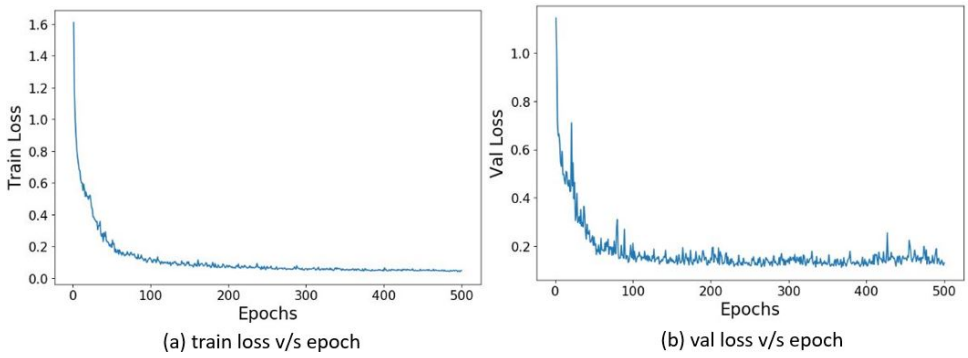



Figure 3.8: Loss v/s epoch graphs for 80/20 datasplit ratio using Mask-RCNN model

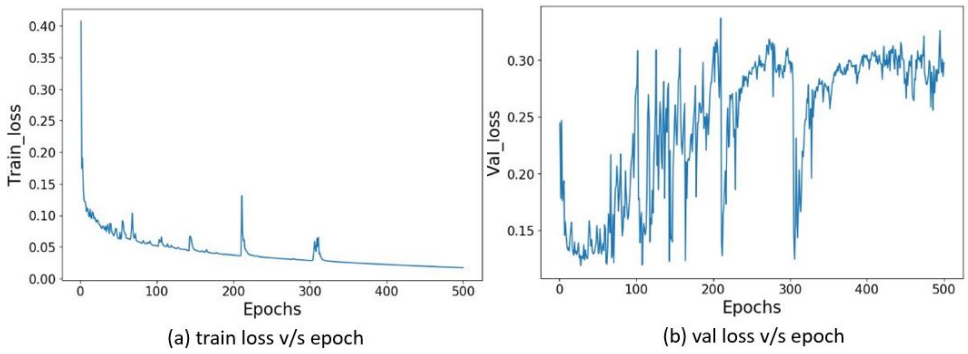



Figure 3.9: Loss v/s epoch graphs for 80/20 datasplit ratio using U-net model

The training is performed by monitoring the training and validation loss with respect to the number of epochs. Figure 3.6 shows the train, val v/s number of epochs graph for 60/40 split. The train loss decreases gradually and saturates to 0.05 by 400 epochs, but a huge variation is seen in val loss without saturation. Similarly figure 3.7 shows the train, val v/s number of epochs graph for 70/30 split, which follows similar trend as 60/40 split. However in figure 3.8, the train and val loss for 80/20 split gives better graph than other splits, where the val loss also decreases gradually with lesser deviations. Thus by comparing the graphs obtained for different data split proportions, it can be concluded that for the current dataset with lesser number of images, 80/20 split to be

the best datasplit proportion.

Further experiments are performed by selecting 80/20 split as the best datasplit ratio. In order to compare the performance of Mask-RCNN model, the U-net model is trained on the 80/20 split by monitoring the train, val loss v/s number of epochs. Figure 3.9 shows the loss per epoch graph obtained for U-net model, where the training loss reduces gradually and saturates whereas the val loss does not saturate.

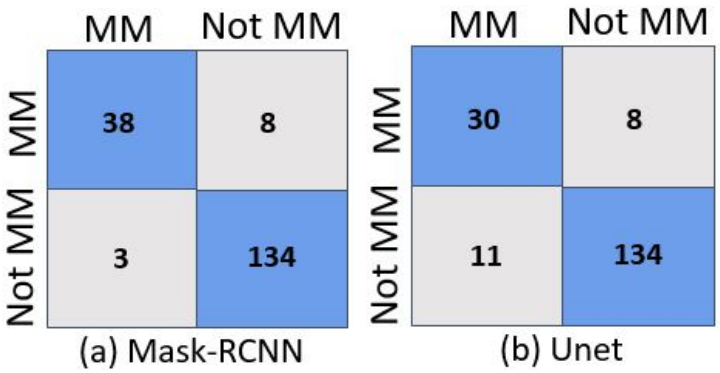
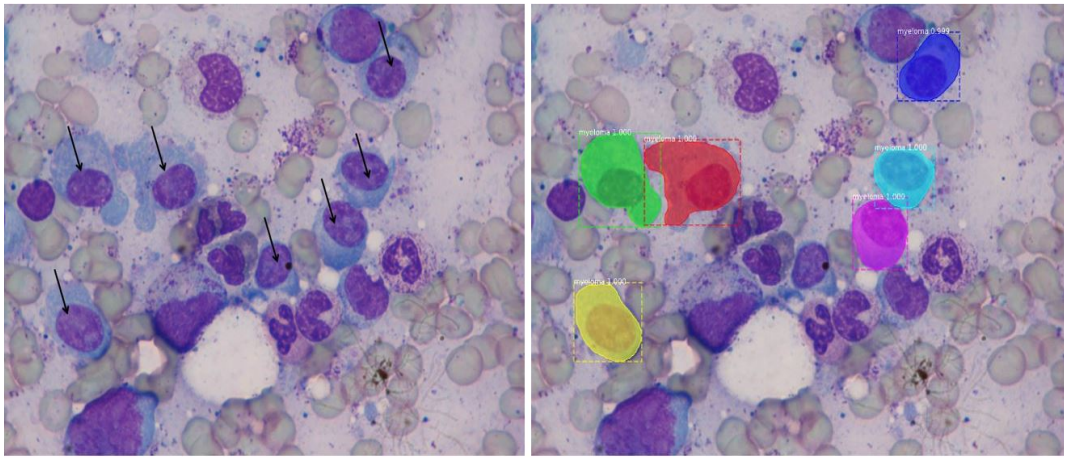



Figure 3.10: Confusion matrices obtained for Mask-RCNN and U-net models

The best trained model is selected based on the least val loss. The confusion matrix for the test set of 80/20 split is calculated for both Mask-RCNN and U-net models by comparing with the groundtruth annotations. Figure 3.10 shows the confusion matrix obtained for the test set for both models on respective best trained models. The Mask-RCNN model has lesser false negatives when compared to U-net model, but the false positive obtained for both models are same. Table 3.2 shows the evaluation metrics calculated for both models. From the confusion matrices and evaluation metrics obtained, it is observed that Mask-RCNN outperforms the U-net model.

Table 3.2: Evaluation metrics calculated for test set on best trained models of Mask-RCNN and U-net

| **Metrics** | **80/20 split** | |
|---|---|---|
| | **Mask-RCNN (%)** | **U-net (%)** |
| **accuracy** | 93.99 | 89.62 |
| **precision** | 82.61 | 78.95 |
| **recall** | 92.68 | 73.17 |
| **specificity** | 94.37 | 94.37 |
| **negative predictive value** | 97.81 | 92.41 |
| **fasle discovery rate** | 17.39 | 21.05 |
| **f1-score** | 87.36 | 75.95 |

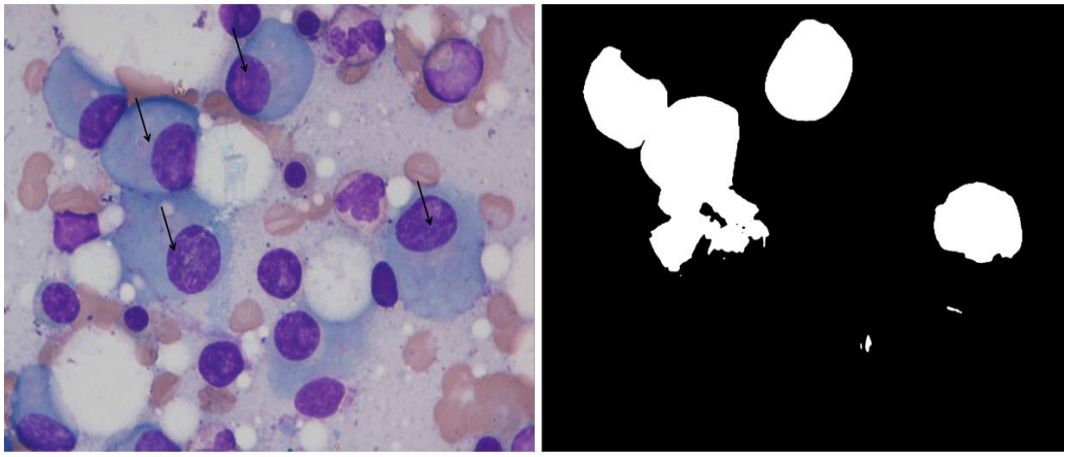


(a) Ground truth annotations (a) Mask-RCNN predictions

Figure 3.11: Mask-RCNN predictions for test set on best trained model

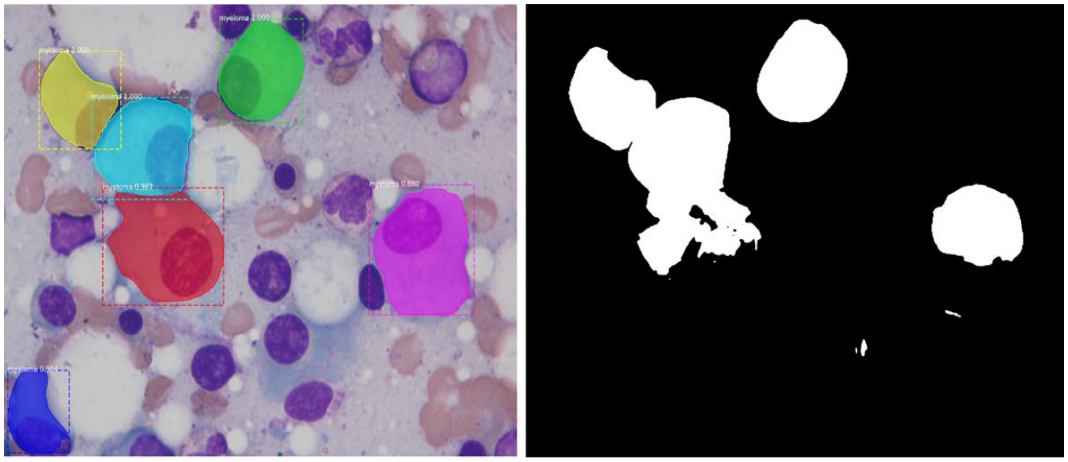
Figure 3.11 shows the sample detection of multiple myeloma cells for the test set data by the best trained model of Mask-RCNN. Figure 3.11 (a) shows the groudtruth annotations by the expert pathologists and figure 3.11 (b) shows the predictions of Mask-RCNN with the coloured portions in the image identified as the myeloma cells. Similarly, the predictions of the best trained model of U-net is shown in figure 3.12, where figure 3.12 (a) shows the groundtruth annotations and figure 3.12 (b) shows the

respective U-net predictions.



(a) Groundtruth annotations (b) Unet predictions

Figure 3.12: U-net predictions for test set on best trained model



(a) Mask-RCNN predictions (b) Unet predictions

Figure 3.13: Comparison of Mask-RCNN and U-net predictions

The comparison of predicted results of Mask-RCNN and U-net models are shown in figure 3.13, where the coloured segments in figure 3.13 (a) shows the predictions of Mask-RCNN and the white segments in figure 3.13 (b) shows the predictions of U-net.

It is observed that U-net model is unable to separate out the clustered plasma cells, while Mask-RCNN clearly segments out different plasma cells clustered together. Thus from the comparison, it can be concluded that Mask-RCNN model has competing results when compared to U-net model for the detection of multiple myeloma cancer cells.

# Chapter 4

# Conclusion & Future Work

The application of deep learning based computer vision techniques has now become a vital component in several researches in the medical filed. This can act as a helping hand for the doctors and clinicians to provide better diagnosis and treatment. In this work, we explore the effectiveness of deep learning based detection and segmentation algorithms on microscopic image datasets such as comet assay images and multiple myeloma images.

The damage analysis of DNA from comet assays is presently performed using image processing software tools. The existing open source softwares are limited to accurately detect, classify and measure the valid comets. The chapter 2 proposes a data driven deep learning approach using Faster R-CNN algorithm for the detection of valid and invalid comets from a comet assay image. The proposed model for valid comet detection is selected by analysing the variation in evaluation metrics for different values of IOU and training epochs. The experimental results shows that the proposed method gave competing results when compared with open source softwares open comet and hicomet.

With the proper identification of valid and invalid comets, it can improve the further quantification of the comets. The proposed method works well for best IOU set as 0.5, but the detection of valid comets reduces when IOU is increased. The detection can be further improved by tuning the sliding window size of RPN network in the Faster R-CNN model. Also, algorithms like Mask-RCNN, YOLO or SSD can be used as future work for this module.

The detection of multiple myeloma cells is essential for the diagnosis of multiple myeloma cancer. In chapter 3, we propose a deep learning based algorithm Mask-RCNN for the detection and segmentation of multiple myeloma cells from bone marrow aspirate images. The polygon mask annotations for the current dataset are done with the help of VGG Image annotation software and groundtruth annotations. Since the dataset is limited to lesser number of images, we have performed data augmentation and have used transfer learning approach for training the model. The model was trained by monitoring the train and validation loss per epoch and the proposed model was selected based on the minimal loss for the validation data. The performance of proposed model is compared with most popular biomedical segmentation algorithm U-net. The comparison results shows that Mask-RCNN outperforms U-net with competing results and also addresses most of the challenges that exist in multiple myeloma segmentation. Even though Mask-RCNN is able to address most of the challenges specific to myeloma cell segmentation, it failed to detect some of the under-stained cells where the colour contrast of cytoplasm and background is lesser. Inorder to establish the supremacy

of the proposed method, the experiments has to be replicated on a huge dataset and can then be made as a real time tool for detection of multiple myeloma cancer, thus reducing the effort of pathologists. Further the results of the Mask-RCNN model can be improved by tuning the hyper parameters such as anchor boxes, detection minimum confidence etc.

Hence, the obtained results from both the works shows that CNN based object detection algorithms such as Faster-RCNN and Mask-RCNN gives competing results when compared to classical image processing techniques used in literature.

# References

[1] Uzunova, H., Wilms, M., Handels, H., & Ehrhardt, J. (2017, September). Training CNNs for image registration from few samples with model-based data augmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 223-231). Springer, Cham.

[2] Uzunova, H., Wilms, M., Handels, H., & Ehrhardt, J. (2017, September). Training CNNs for image registration from few samples with model-based data augmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 223-231). Springer, Cham.

[3] Roy, A. G., Conjeti, S., Sheet, D., Katouzian, A., Navab, N., & Wachinger, C. (2017, September). Error corrective boosting for learning fully convolutional networks with limited data. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 231-239). Springer, Cham.

[4] Uzunova, H., Wilms, M., Handels, H., & Ehrhardt, J. (2017, September). Training CNNs for image registration from few samples with model-based data augmentation.

In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 223-231). Springer, Cham.

[5] Karimi, D., & Ruan, D. (2017, September). Synergistic combination of learned and hand-crafted features for prostate lesion classification in multiparametric magnetic resonance imaging. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 391-398). Springer, Cham.

[6] Dubost, F., Bortsova, G., Adams, H., Ikram, A., Niessen, W. J., Vernooij, M., & De Bruijne, M. (2017, September). GP-Unet: Lesion detection from weak labels with a 3D regression network. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 214-221). Springer, Cham.

[7] Dubost, F., Bortsova, G., Adams, H., Ikram, A., Niessen, W. J., Vernooij, M., & De Bruijne, M. (2017, September). GP-Unet: Lesion detection from weak labels with a 3D regression network. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 214-221). Springer, Cham.

[8] Cheplygina, V., de Bruijne, M., & Pluim, J. P. (2019). Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. Medical image analysis, 54, 280-296.

[9] Ghafoorian, M., Mehrtash, A., Kapur, T., Karssemeijer, N., Marchiori, E., Pesteie, M., ... & Fedorov, A. (2017, September). Transfer learning for domain adaptation in mri: Application in brain lesion segmentation. In International conference on med-

ical image computing and computer-assisted intervention (pp. 516-524). Springer, Cham.

[10] Oktay, O., Ferrante, E., Kamnitsas, K., Heinrich, M., Bai, W., Caballero, J., ... & Kainz, B. (2017). Anatomically constrained neural networks (ACNNs): application to cardiac image enhancement and segmentation. IEEE transactions on medical imaging, 37(2), 384-395.

[11] Ravishankar, H., Venkataramani, R., Thiruvenkadam, S., Sudhakar, P., & Vaidya, V. (2017, September). Learning and incorporating shape models for semantic segmentation. In International conference on medical image computing and computer-assisted intervention (pp. 203-211). Springer, Cham.

[12] Agrawal, P., Whitaker, R. T., & Elhabian, S. Y. (2017, September). Learning Deep Features for Automated Placement of Correspondence Points on Ensembles of Complex Shapes. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 185-193). Springer, Cham.

[13] Liu, M., Zhang, J., Adeli, E., & Shen, D. (2017, September). Deep multi-task multi-channel learning for joint classification and regression of brain status. In International conference on medical image computing and computer-assisted intervention (pp. 3-11). Springer, Cham.

[14] Fidon, L., Li, W., Garcia-Peraza-Herrera, L. C., Ekanayake, J., Kitchen, N., Ourselin, S., & Vercauteren, T. (2017, September). Scalable multimodal convolu-

tional networks for brain tumour segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 285-293). Springer, Cham.

[15] Ge, Z., Demyanov, S., Chakravorty, R., Bowling, A., & Garnavi, R. (2017, September). Skin disease recognition using deep saliency features and multimodal learning of dermoscopy and clinical images. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 250-258). Springer, Cham.

[16] Xie, Y., Xia, Y., Zhang, J., Feng, D. D., Fulham, M., & Cai, W. (2017, September). Transferable multi-model ensemble for benign-malignant lung nodule classification on chest CT. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 656-664). Springer, Cham.

[17] Dai, L., Sheng, B., Wu, Q., Li, H., Hou, X., Jia, W., & Fang, R. (2017, September). Retinal microaneurysm detection using clinical report guided multi-sieving CNN. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 525-532). Springer, Cham.

[18] Shah, S. U. (2012). Importance of Genotoxicity & S2A guidelines for genotoxicity testing for pharmaceuticals. IOSR Journal of Pharmacy and Biological Sciences, 1(2), 43-54.

[19] McKelvey-Martin, V. J., Green, M. H. L., Schmezer, P., Pool-Zobel, B. L., De Meo, M. P., & Collins, A. (1993). The single cell gel electrophoresis assay (comet assay):

a European review. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 288(1), 47-63.

[20] Collins, A. R. (2004). The comet assay for DNA damage and repair. Molecular biotechnology, 26(3), 249.

[21] Rojas, E., Lopez, M. C.,& Valverde, M. (1999). Single cell gel electrophoresis assay: methodology and applications. Journal of Chromatography B: Biomedical Sciences and Applications, 722(1-2), 225-254.

[22] Koca, K., Lankoff, A., Banasik, A., Lisowska, H., Kuszewski, T., Gd, S.,...& Wojcik, A. (2003). A cross-platform public domain PC image-analysis program for the comet assay. Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 534(1-2), 15-20.

[23] CometScoreCometScoringSoftware,Version1.5,TriTekCorp. Available: http://autocomet.com/ł, Accessed:5.10.12.

[24] Gyori, B. M., Venkatachalam, G., Thiagarajan, P. S., Hsu, D.,& Clement, M. V. (2014). OpenComet: an automated tool for comet assay image analysis. Redox biology, 2, 457-465.

[25] Lee, T., Lee, S., Sim, W. Y., Jung, Y. M., Han, S., Won, J. H., ...& Yoon, S. (2018). HiComet: a high-throughput comet analysis tool for large-scale DNA damage assessment. BMC bioinformatics, 19(1), 44.

[26] de Souza, M. R., Ruschel, R., Susin, A., Boeira, J. M., Guimares, L. V.,Parraga, A. (2018, July). A Framework for Automatic Recognition of Cell Damage on Microscopic Images using Artificial Neural Networks.In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 636-639). IEEE.

[27] Namuduri, S., Narayanan, B. N., Karbaschi, M., Cooke, M.,& Bhansali, S. (2019, September). Automated quantification of DNA damage via deep transfer learning based analysis of comet assay images. In Applications of Machine Learning (Vol. 11139, p. 111390Y). International Society for Optics and Photonics.

[28] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91-99).

[29] Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).

[30] Girshick, R., Donahue, J., Darrell, T., Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).

[31] https://github.com/kbardool/keras-frcnn , accessed on September 2019.

[32] Cai, Z., Vallis, K. A., & Reilly, R. M. (2009). Computational analysis of the number, area and density of Y-H2AX foci in breast cancer cells exposed to 111In-DTPA-

hEGF or Y-rays using Image-J software. International journal of radiation biology, 85(3), 262-271.

[33] Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10), 1345-1359.

[34] https://www.analyticsvidhya.com/blog/2018/11/implementation-faster-r-cnn-python-object-detection/, accessed on September 2019.

[35] https://drive.google.com/file/d/1OmCKlUEYmTjg_jaaN-IQm81eHROU-Gyl/view, accessed on September 2019.

[36] Zitnick, C. L., Doll'ar, P. (2014, September). Edge boxes: Locating object proposals from edges. In European conference on computer vision (pp. 391-405). Springer, Cham.

[37] http://ronny.rest/tutorials/module/localization 001/iou/ , accessed on September 2019.

[38] Saeedizadeh, Z., Mehri Dehnavi, A., Talebi, A., Rabbani, H., Sarrafzadeh, O., & Vard, A. (2016). Automatic recognition of myeloma cells in microscopic images using bottleneck algorithm, modified watershed and SVM classifier. Journal of microscopy, 261(1), 46-56.

[39] Guyton, A. C., & Hall, J. E. (2006). Textbook of medical physiology 11th ed. Philadelphia, Perm: Elsevier Saunders.

[40] Theml, H., Diem, H., & Haferlach, T. (2004). Color atlas of hematology. Thieme.

[41] Minges Wols, H. A. (2001). Plasma cells. eLS, 1-11.

[42] Smith, A., Wisloff, F., Samson, D., & UK Myeloma Forum, Nordic Myeloma Study Group and British Committee for Standards in Haematology. (2006). Guidelines on the diagnosis and management of multiple myeloma 2005. British journal of haematology, 132(4), 410-451.

[43] MoradiAmin, M., Memari, A., Samadzadehaghdam, N., Kermani, S., & Talebi, A. (2016). Computer aided detection and classification of acute lymphoblastic leukemia cell subtypes based on microscopic image analysis. Microscopy research and technique, 79(10), 908-916.

[44] Strbkova, L., Zicha, D., Vesely, P., & Chmelik, R. (2017). Automated classification of cell morphology by coherence-controlled holographic microscopy. Journal of biomedical optics, 22(8), 086008.

[45] Simon, A., Vinayakumar, R., Sowmya, V., Soman, K. P., & Gopalakrishnan, E. A. A. (2019). A deep learning approach for patch-based disease diagnosis from microscopic images. In Classification Techniques for Medical Image Analysis and Computer Aided Diagnosis (pp. 109-127). Academic Press.

[46] Saiprasath, G. B., Babu, N., ArunPriyan, J., Vinayakumar, R., Sowmya, V., & Soman, K. P. (2019). Performance comparison of machine learning algorithms for malaria detection using microscopic images.

[47] Ghane, N., Vard, A., Talebi, A., & Nematollahy, P. (2017). Segmentation of white blood cells from microscopic images using a novel combination of K-means clustering and modified watershed algorithm. Journal of medical signals and sensors, 7(2), 92.

[48] Liu, H., Cao, H., & Song, E. (2019). Bone marrow cells detection: A technique for the microscopic image analysis. Journal of medical systems, 43(4), 82.

[49] Acharya, V., & Kumar, P. (2019). Detection of acute lymphoblastic leukemia using image segmentation and data mining algorithms. Medical & biological engineering & computing, 57(8), 1783-1811.

[50] Ghane, N., Vard, A., Talebi, A., & Nematollahy, P. (2019). Classification of chronic myeloid leukemia cell subtypes based on microscopic image analysis. EXCLI journal, 18, 382.

[51] Gupta, R., Mallick, P., Duggal, R., Gupta, A., & Sharma, O. (2017). Stain color normalization and segmentation of plasma cells in microscopic images as a prelude to development of computer assisted automated disease diagnostic tool in multiple myeloma. Clinical Lymphoma, Myeloma and Leukemia, 17(1), e99.

[52] Gupta, A., Mallick, P., Sharma, O., Gupta, R., & Duggal, R. (2018). PCSeg: Color model driven probabilistic multiphase level set based tool for plasma cell segmentation in multiple myeloma. PloS one, 13(12).

[53] Tehsin, S., Zameer, S., & Saif, S. (2019, January). Myeloma Cell Detection in Bone Marrow Aspiration Using Microscopic Images. In 2019 11th International Conference on Knowledge and Smart Technology (KST) (pp. 57-61). IEEE.

[54] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).

[55] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.

[56] Vuola, A. O., Akram, S. U., & Kannala, J. (2019, April). Mask-RCNN and U-net ensembled for nuclei segmentation. In 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) (pp. 208-212). IEEE.

[57] Dutta, A., & Zisserman, A. (2019, October). The VIA annotation software for images, audio and video. In Proceedings of the 27th ACM International Conference on Multimedia (pp. 2276-2279).

[58] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems (pp. 91-99).

[59] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).

[60] Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).

[61] https://github.com/matterport/Mask_RCNN , accessed on January 2020.

[62] Johnson, J. W. (2018). Adapting mask-rcnn for automatic nucleus segmentation. arXiv preprint arXiv:1805.00500.

[63] https://github.com/ShawDa/unet-rgb, accessed on February 2020.

[64] Gupta, R., & Gupta, A. (2019). MiMM SBILab Dataset: Microscopic Images of Multiple Myeloma [Data set]. The Cancer Imaging Archive0. https://doi.org/10.7937/tcia.2019.pnn6aypl

[65] Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., ... & Tarbox, L. (2013). The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. Journal of digital imaging, 26(6), 1045-1057.

[66] Rajkumar, S. V., Dimopoulos, M. A., Palumbo, A., Blade, J., Merlini, G., Mateos, M. V., ... & Landgren, O. (2014). International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. The lancet oncology, 15(12), e538-e548.

[67] Gupta, A., Duggal, R., Gupta, R., Kumar, L., Thakkar, N., & Satpathy, D. GCTI-SN: geometry-inspired chemical and tissue invariant stain normalization of microscopic medical images. Under review.

[68] Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10), 1345-1359.

[69] https://github.com/matterport/Mask_RCNN/releases, accessed on January 2020.

# List of Publications based on the research work

1. Vyshnav M T, Sowmya V, Gopalakrishnan E.A, Sajith Variyar V. V, Vijay Krishna Menon, Soman K.P, *Deep Learning Based Approach For Multiple Myeloma Detection* , In 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT - 2020), IEEE (accepted)

# PARKINSON'S DISEASE CLASSIFICATION USING MR IMAGES SYNTHESIZED BY GENERATIVE MODELS

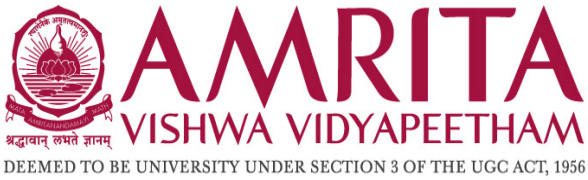
A THESIS

*Submitted by*

**Yamini Madan**
**(CB.EN.P2CEN19016)**

*in partial fulfillment for the award of the degree of*

**MASTER OF TECHNOLOGY**
**IN**
**COMPUTATIONAL ENGINEERING AND NETWORKING**
**(DATA SCIENCE)**




**Center for Computational Engineering and Networking (Data Science)**

**AMRITA SCHOOL OF ENGINEERING**

**AMRITA VISHWA VIDYAPEETHAM**

COIMBATORE - 641 112 (INDIA)

**JUNE - 2021**

AMRITA SCHOOL OF ENGINEERING

# AMRITA VISHWA VIDYAPEETHAM

COIMBATORE - 641 112

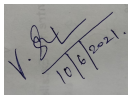
AMRITA VISHWA VIDYAPEETHAM

श्रद्धावान् लभते ज्ञानम्

DEEMED TO BE UNIVERSITY UNDER SECTION 3 OF THE UGC ACT, 1956

## BONAFIDE CERTIFICATE

This is to certify that the thesis entitled **"PARKINSON'S DISEASE CLASSIFICATION USING MR IMAGES SYNTHESIZED BY GENERATIVE MODELS"** submitted by **Yamini Madan (Register Number- CB.EN.P2CEN19016),** for the award of the **Degree of Master of Technology** in the **"COMPUTATIONAL ENGINEERING AND NETWORKING (DATA SCIENCE)"** is a bonafide record of the work carried out by her under our guidance and supervision at Amrita School of Engineering, Coimbatore.



**Dr. Sowmya V**
Project Guide

**Dr. Gopalakrishnan EA**
Project Co-Guide

*Submitted for the university examination held on*

**INTERNAL EXAMINER** **EXTERNAL EXAMINER**

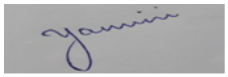
## AMRITA SCHOOL OF ENGINEERING

# AMRITA VISHWA VIDYAPEETHAM

COIMBATORE - 641 112

## DECLARATION

I, **YAMINI MADAN (CB.EN.P2CEN19016),** hereby declare that this thesis entitled **"PARKINSON'S DISEASE CLASSIFICATION USING MR IMAGES SYNTHESIZED BY GENERATIVE MODELS"**, is the record of the original work done by me under the guidance of **Dr. Sowmya V**, Assistant Professor (Sr.Gr). and **Dr. Gopalakrishnan EA**, Assistant Professor (Sr.Gr), Centre for Computational Engineering and Networking (Data Science), Amrita School of Engineering, Coimbatore To the best of my knowledge this work has not formed the basis for the award of any degree/diploma/ associate ship/fellowship/or a similar award to any candidate in any University.



**Place:Coimbatore**

**Date:** **Signature of the Student**

COUNTERSIGNED

Dr. K.P.Soman
Professor and Head
Center for Computational Engineering and Networking (Data Science)

# Contents

# Acknowledgement

I would like to express my gratitude to all who have helped me directly or indirectly in this project. I would like to offer pranam to the lotus feet of MATA AMRITANANDAMAYI DEVI who showered her blessings all throughout the project duration. I would like to thank Dr. Soman KP, Chairperson, Centre for Computational Engineering and Networking (Data Science), Coimbatore, for providing me with all the necessary facilities needed to carry out this project. I would like to express my gratitude to my guide Dr. Sowmya V, Assistant Professor, Department of CEN, Dr. Gopalakrishnan EA, Assistant Professor, Department of CEN and Ms. Iswarya KV, Research Scholar, Department of CEN for their valuable inputs, guidance, encouragement and generous assistance throughout the development of the project. I am also thankful to my panel members, Dr. Geetha P, Dr. Neethu Mohan, Dr. G Jyotish Lal and Mr. B Premjith for their encouragement and valuable suggestions. I would also like to thank all the Teaching and Non-Teaching Staff of CEN Department for providing me with all the support whenever needed.

I pay my respects and love to my beloved parents and all other family members and friends who have helped me throughout the project duration.

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| MRI | Magentic Resonance Imaging |
| PD | Parkinson's Disease |
| NC | Normal Cohorts |
| GAN | Generative Adversarial Networks |
| SSIM | Structure Similarity Index |
| PSNR | Peak Signal to ratio |
| DL | Deep Learning |
| ML | Machine Learning |
| CNN | Convolutional Neural Network |

# Abstract

Neuroimaging imaging techniques such as Magnetic Resonance Imaging (MRI) have improved our understanding of the working of the brain and aided in the early diagnosis of brain disorders such as Parkinson's Disease (PD). These accurate structural representations in the form of MRI, coupled with the advancements in computational resources and Machine Learning (ML) techniques, have allowed for easy identification of biomarkers for automatic diagnosis. Practically, a major limiting factor for ML based work is the amount of data available. The aim of this work is to identify the slices of interest in identifying PD and resolve the class imbalance problem in the T2 MRI modality of PPMI dataset as well. While the data in the larger group (PD) can be selected in proportion to the smaller group (Normal Cohorts - NC), we have approached the issue with data augmenting strategies that will aid in the improved classification of the images by a Deep Neural Network (DNN) in an unbiased manner, which will be verified for model generalizability using cross validation. We propose the usage of generative models in identifying and augmenting the slices of interest of NC class data to meet the PD class. A subject-level classification is performed on the test data with DNN trained on both real data as well as the synthetic data. Additionally, the best performing transfer learned deep learning mode is compared with Capsule Network and evaluated using standard metrics in identifying the model that performs better in PD classification.

# Chapter 1

# Introduction

Parkinson's Disease, a progressive neurodegenerative disorder occurs due to loss of dopamine producing neurons in the mid brain region. The loss of neurons occurs due to accumulation of iron in the brain cells which is captured well by T2 MRI scans [1,2]. Visible motor symptoms are characterized by slowed movement (bradykinesia), stiffness in limbs, slurred speech and restricted movement [3]. Being a progressive disorder, the clinical motor-symptoms is broken into 5 stages where the condition deteriorates over time. Neuroimaging techniques like Magnetic Resonance Imaging (MRI) scan, Positron Emission Tomography (PET), Computed Tomography (CT) and Single Photon Emission Computed Tomography (SPECT) serve to be a potential biomarker by capturing the morphological changes in the brain. These techniques prove to be useful in early diagnosis as the visible symptoms of the disease start manifesting evidently only after 60% of the dopaminergic neurons are atrophied [4]. SPECT and PET scans are effective in detecting the presence of disease from an MRI only when the disease has advanced to 80% of critical stage [5]. MRI is the most preferred as it does not involve harmful radiations and is also effective in detecting Parkinson's Disease. Deep Learning

(DL) models and Artificial Intelligence (AI) can be leveraged to learn and identify the features of an image to assist medical personnel in efficient medication to patients and in performing surgeries as well.

**Subject level Classification**

Classifying imaging results at the subject level is the requirement for disease diagnosis. The AI algorithm in use can classify a subject as affected by disease or undiseased (Normal Cohort). This will also help in easy interpretation of the results and does not require much technical expertise.

**Region of Interest**

In deep learning, the model's performance improves based on the availability of a large dataset for training. However, the quality of images and features to be learnt is the basic necessity for a good model performance. Since there are multiple slices in an MRI, certain slices could be less informative than the others. Extracting and identifying the slices which correspond to the mid-brain is of importance and the Region of Interest in this study.

**Model generalizability**

When we test the model performance on 1 split of test data, it gives an understanding of the model performance on that particular test split. This is influenced by factors such as classes of image in the split, quality of images, etc. On using cross-validation, the model is validated on k folds of the data [6, 7] which gives a broader insight into the model's performance over unseen data as well.

Nevertheless, one of the major problems faced for reliable performance of a Deep Learn-

ing model is lack of sufficient data to build a fair and realistic model to be deployed in real time critical scenarios. This can be overcome by data augmentation where modified images of the existing data is added to increase the number of samples of the classes having insufficient images. This acts as a regularizer thereby prevents overfitting of the DL model. While simple data augmentation techniques involve a simple rotate and crop of the images, models trained on simple augmented data might not very effective in handling new kinds of pattern in data. Generative modelling that is gaining a lot of momentum in recent years, produces realistic images whose distributions are similar to the original image distribution. This allows creation of new variations in patterns of the features learnt expanding the diversity of the images generated. In our proposed work, we aim to use generative modelling to overcome data insufficiency, which is explained in the further sections.

# Chapter 2

# Literature Survey

One of the earliest works in identifying Parkinson's Disease using Convolutional Neural Networks (CNN) was done by [5]. Their work utilized the AlexNet architecture to identify PD from selected slices of T2 MR images. However, there is no clear demarcation on the range of slices used in the work to detect PD from T2 MRI. [8] have used simple data augmentation techniques such as flip and rotate, to overcome class imbalance between PD and NC class. However, the augmentation process is done before splitting the data into train and test sets that will result in data leakage. Further, an ensemble deep learning model is used to classify Parkinson's Disease from T2 MR images and results are discussed at the slice level. In [1], the model is evaluated at the subject level for T1 images using a 3D CNN to identify the presence of Parkinson's Disease, and model generalizability is verified using k fold cross-validation. Classification is done at the subject level on T1 MRI images. There has been no initiative in evaluating deep learning architectures at the subject level on T2 MRI images with the view of verifying the model's generalizability to the best of our knowledge.

GANs generate superior images in terms of quality and possess the potential to

generalize new patterns while maintaining the structural information [9]. The authors have used GAN in improving the T1 MRI dataset to identify PD. GAN is used to maximize the results for PD classification. LeNet 5 classifier performance is compared on training with and without synthetic images. PSNR is the metric used in evaluating quality of generated images. The classification metrics of the model trained with data augmentation shows a 3% improvement in accuracy when compared to the model trained without data augmentation. Another application of using GAN in medicine is in the early detection of diabetes through Diabetic retinopathy. Diabetic retinopathy is a complication that develops due to diabetes which hinders eyesight. Being a leading cause of blindness, it occurs in those suffering from diabetes for a prolonged period. To generate realistic samples of the diseased retinal image samples, synthetic images of the proliferative class are produced using Generative Adversarial Networks (GANs). Results indicate that the generated images are as competent as the real images of the proliferative class [10]. Simple data augmentation techniques using cropping and rotation lack the generalization abilities as the model is not exposed to a variety of different patterns in images. This is overcome by using GANs to generate synthetic data which have a good generalization capability while capturing the data distribution [11]. DCGAN is found to more stable during training, in comparison to the vanilla GAN architecture. It is used to generate images for 3 stages of Alzheimer's Disease from PET scans. Metrics used to assess image quality is mean PSNR and mean SSIM. Performance of DCGAN and WGAN image generation with respect to size of image generated for different MRI modalities along the sagittal axis [12]. It is observed that

DCGAN generated better images for a smaller dimension (64*64) and WGAN can generate better images for bigger dimensions (128*128) while maintaining the sequence specific texture and avoiding mode collapse which occurs in DCGAN. Nevertheless, artefacts remain in the synthetic images. Visual Turing Test is used to analyse the image generation of both the generative architectures.

Variational Autoencoder (VAE) is another generative model using latent vectors to represent the essential information of the input images which will be used for reconstruction of images. Speech impairment is one of the characteristic biomarkers of PD and an Autoencoder Neural Network trained to learn the features and differentiate between the speech recording of a PD subject from an NC subject is described in the work by [13]. In an attempt to identify the optimal size of a latent vector dimension, [14] have used Variational Autoencoder to determine the effect of latent variable size for the reconstruction of new MNIST images. A latent vector of size 10 captures all the possible variation in styles of the digits which is not so in the case of smaller latent vector size. [15] used Deep Variational Autoencoders as a feature extractor to identify ADHD from fMRI. The latent variables generated are used in estimating Functional Brain Networks. A robust Variational Autoencoder model is used by [16] to represent the complex, high level features using latent variables used to identify lesion from brain MRI images Capsule Networks are known to perform better than traditional neural networks as they capture the spatial and temporal relationship in features of the data. The following 2 works based on non-medical [17, 18] and medical data [19] indicate better classification of images due to pre-processing. Hyperparameters tuned are batch

size, number of epochs, learning rate and filter size, channels of the convolutional layer. Input images are scaled down to 28*28 to reduce number of computational parameters. It does not seem to cause loss of prominent features. From the literature the powerful ability of GANs and VAE to generate realistic synthetic images is exhibited. As discussed, they are widely applied in medical imaging to enhance the performance of deep learning model. We aim to extend and experiment its application on generating images to improve detection of PD as well.

Therefore, we propose to generate and compare the quality of synthetic T2 MRI of Normal Cohorts, which is imbalanced when compared to number of PD subjects using 3 widely used generative models – DCGAN, WGAN and VAE.

The potential of using Capsule Network has improved drastically over time and we propose to use the architecture to identify its prospective as a classifier in PD detection. Eventually we plan to compare the performance of transfer learned DenseNet201 architecture with Capsule Network. The test results are evaluated at the subject level which will be of assistance in the real time scenario.

To the extent of our knowledge, there has been no published work to generate MRI images for T2 modality using the mentioned generative models or using Capsule Network as a classifier architecture. Also, classification of test data at the subject level using DNNs have not been addressed in the literature so far.

## 2.1 Problem statement

- Identifying Region of Interest (RoI) pertaining to the detection of Parkinson's disease
- Verify the performance of CNN model in terms of model generalizability to classify the disease
- Classification of predicted results at the subject level
- Overcoming class imbalance using generative techniques
- Compare DL model performance with and without data augmentation

## 2.2 Objectives

- Experimenting with the Region of Interest to find the optimum slice range containing important feature necessary in identifying PD
- Using Stratified Cross Validation to verify CNN model generalizability
- Overcoming class imbalance by implementing data augmentation of Normal Cohort images using generative techniques on ROI
- Compare the synthetic image generative capability of generative models – WGAN,DCGAN and Variational Autoencoders.
- Compare performance of test data classification by DL models with and without data augmentation

# Chapter 3

# Background

## 3.1 MR Images

An MRI in general contains multiple slices (Fig.3.2) which capture deformities or abnormalities in various sections of the body – in our case, the brain layer wise. Depending on the scanner used the number of slices varies. The 3 primary imaging planes of an MRI images are Axial, Sagittal and Coronal as shown in Fig.3.1. The magnetic fields produced by the scanner is manipulated to produce distinct types of image – T1 weighted MRI and T2 weighted MRI. The bright or white parts on a T1-w MRI represents proteins and lipids while the same in a T2-w MRI represents water con-

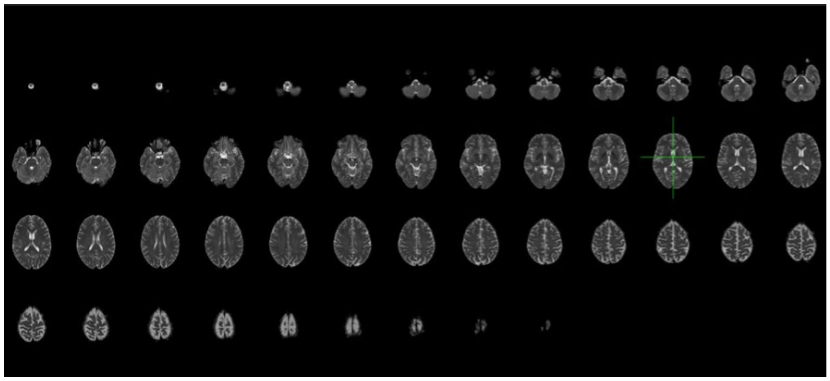


Figure 3.1: T2 MRI of a subject consisting of 48 slices

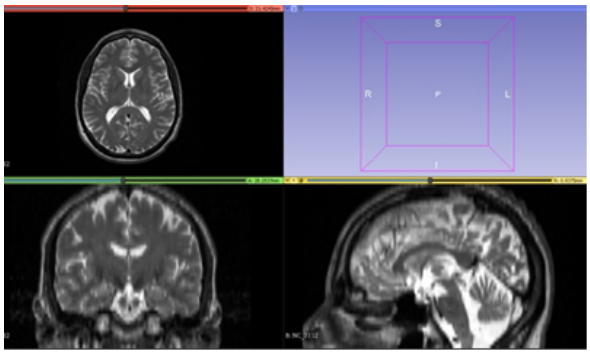


Figure 3.2: 3 planes of MRI a) Axial b) Sagittal c) Coronal (clockwise direction)

tent depicting conditions like Inflammation, Tumour, Haemorrhage, Infection (which is represented by the dark parts in a T1-w MRI).

## 3.2 Generative Modelling

Generative models have a probability distribution associated with its latent input variable. They learn the posterior distribution P(Y—X) via Bayesian rule where X is the input variable and Y is the variable representing reconstruction. Inputs to the GAN generator and VAE decoder have distributions. VAE models explicitly learn likelihood distribution P(X—Y) through loss function. GAN does not explicitly learn the likelihood distribution, instead learns it through generator fooling the discriminator. GAN generators aim at minimizing the difference between P(X) and P(Z—Y). Z is the fake or synthetic images that are generated by the model [23]. The latent variable in GAN is predefined which is the noise as input to generator, while in VAE it is learnt while training by minimizing of loss function which is the sum of kl divergence loss and reconstruction loss. In VAE, the encoder model is the discriminator and decoder model are the generator.

# Chapter 4

# Proposed Work

The overall methodology followed in this work is described which is broadly classified into 3 parts:

1. Pre-processing techniques are followed to bring the data to desirable format.

2. Identifying ROI of T2 MRI which will be used in training using 3 CNN architectures. Results are classified and compared at at the subject level.

3. Images of NC are synthesized using the identified ROI using with 3 generative architectures to overcome class imbalance. Performance of the best classifier model with without data augmentation is compared using standard metrics. The same is performed for Capsule Network and the performance of the CNN and Capsule Network is compared to identify the model that performs better over the test data.

## 4.1 Dataset Description

Parkinson's Progressive Marker Initiative (PPMI) dataset [20] which is an open database initiative for PD detection from signals and images is used in the proposed work. T2 weighted MRI images of 100 subjects are used in the 2nd part of the work - to identify

Table 4.1: Data Description

| | |
|---|---|
| **Modality** | **Magnetic Resonance Imaging (MRI)** |
| **Research Group** | **Normal Cohort and Parkinson's Disease** |
| **Visit** | **Baseline** |
| **Acquisition Plane** | **Axial** |
| **Acquisition Type** | **2D** |
| **Field Strength** | **3T** |
| **Slice Thickness** | **3mm** |
| **Scanner Manufacturing** | **Siemens Trio Tim scanner** |
| **Pixel Spacing** | **X=0.9,Y=0.9** |
| **Weighting** | **T2** |

ROI in detecting PD, and 240 subjects in the 3rd part of the work - using generative models to produce synthetic images from the ROI. Scans are taken along axial axis using Trio Tim scanner. The chosen sample data of T2 MRI image modality has 48 slices for each subject, The MRI images which is split in 80:20 ratio for training and testing, respectively. Of 100 subjects, 40 PD and 40 NC subjects are used for training with the remaining for testing. Of 240 subjects, 116 PD subjects and 84 NC subjects for training and 38 subjects' images with equal proportion of PD and NC for testing. A detailed description of the data is given in Table 4.1.

### 4.1.1 Data Preprocessing

Image pre-processing consisting of 3 parts – Bias Field Correction (BFC), Skull stripping and Intensity normalization is initially done for all the images (see Fig.4.1). BFC removes the non-uniformities of intensity in the MRI [21]. Skull stripping extracts the brain tissue by eliminating region belonging to the cranium to reduce overload in feature extraction [22]. Intensity normalisation is done to bring in the range of image values

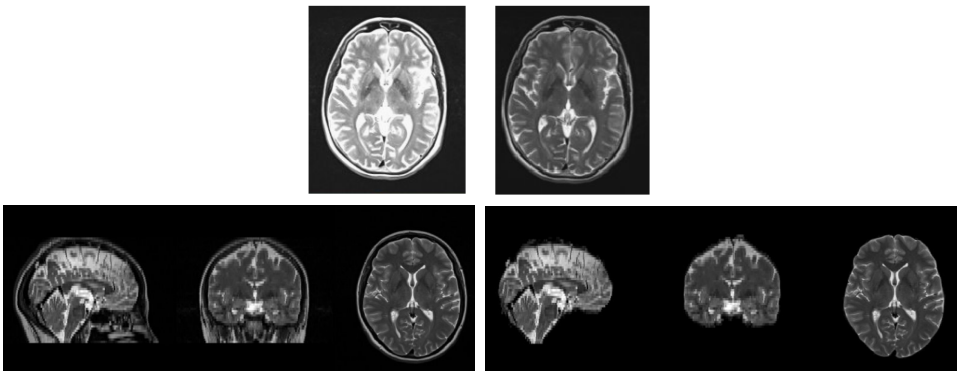


Figure 4.1: Image preprocessing a) Bias Field Correction b) Skull stripping

between 0 and 255 which helps the model. The images are stacked to form a 3-channel image to ensure compatibility with Keras deep learning models. The processed images are split into training and testing before data augmentation to avoid data leakage.

### 4.1.2 Evaluation Metrics

Metrics used to evaluate DL classifier are Accuracy, Precision, Specificity (True Negative Rate), Sensitivity (True Positive Rate) and F1 score. Specificity determines the number of predicted NC subjects correctly predicted. Higher specificity indicates lesser number of mispredictions of a negative subject to be diseased. Sensitivity or Recall determines the number of positive subjects correctly predicted. Higher sensitivity indicates lesser chances of missing to detect the disease when it is present. Precision determines the number of positive predictions that are actually true. It determines the number of subjects wrongly classified as the negative class. F1 score gives the harmonic mean which is a weighted arithmetic mean of precision and recall.

The metrics used to evaluate the quality of synthetic images generated are mean Structural Similarity Index Measure (SSIM) and mean peak Signal to Noise (PSNR)

ratio [9, 11]. SSIM measures the similarity of a synthetic image with the original data structure. Luminance, Contrast and Structure are considered while determining the same. Mean PSNR measures the lossy reconstruction of an image indicates presence of noise while comparing the image pixels with input data.

$$SSIM = \frac{(2\mu_x\mu_y + C_1) + (2\sigma_{xy} + C_2)}{(\mu_x^2\mu_y^2 + C_1) + (\mu_x^2\mu_y^2 + C_2)} \tag{4.1}$$

,where x is the synthetic image and y is the original image. $\mu_x$ and $\mu_y$ are the average of x and y respectively. $\mu_x^2$ is the variance of x, $\mu_y^2$ is the variance of y, xy is the covariance of x and y. C1 and C2 are the constant values which is a stabiliser term.

$$PSNR = 20\ log_{10}\frac{(max(y))^2}{\frac{1}{n}\Sigma(y_i - (y_i))^2} \tag{4.2}$$

,where n is the number of pixels in the image and i is the iteration index from 1 till n. The numerator is maximum pixel value of the original image and denominator represents the mean squared error of original and synthetic image distribution.

# Chapter 5

# Experiments and Results

## 5.1 Classifier Architecture

The overall methodology followed in this subsection's work is shown in Fig.5.1. Transfer learning of deep learning architectures - VGG19, DenseNet121 and Densenet201 are experimented with and their performance is recorded to find the best architecture for the chosen dataset. All layers of DenseNet201 and DenseNet121 are trained, while for VGG-19 only the last 3 layers of the architecture are trained to get a comparable performance. Over the best performing deep learning architecture, we have experimented on the range of minimum number of slices that contains the necessary features for the model to learn to identify PD. The penultimate layer is removed (include top = "False"). Fully connected layers and dropout layers are added as shown in Fig.5.2. Dropout, a regularisation parameter is added to prevent overfitting. The number of hidden layers and neurons in each layer has been obtained on an experimental basis.

**Subject level classification**

The predicted result for a slice using the sigmoid activation from the final layer of

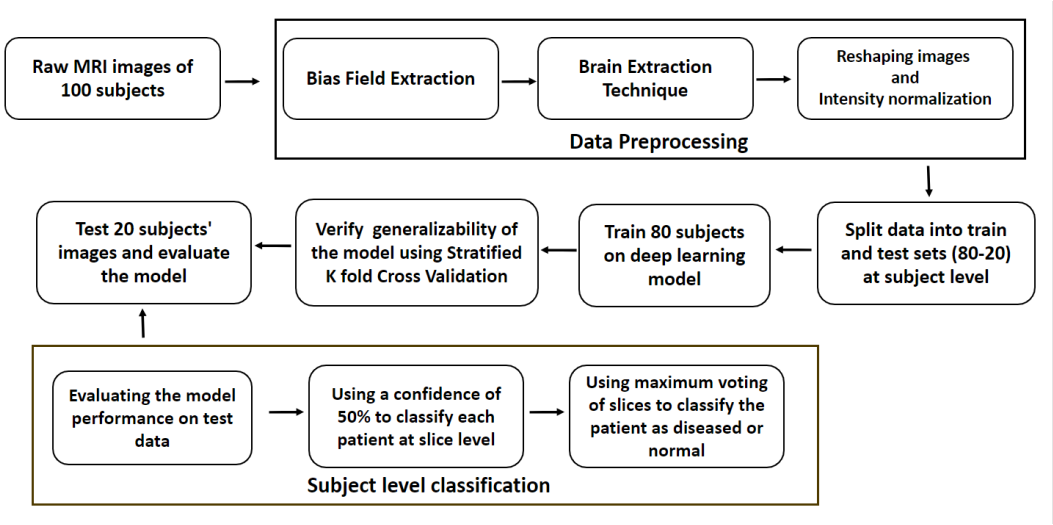



Figure 5.1: Methodology followed in classifying subjects using CNN models

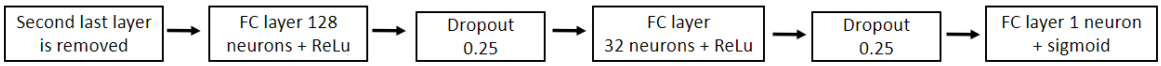



Figure 5.2: Modifications done to the deep learning architectures used

the deep learning architecture is a probability value ranging between 0 and 1. With a confidence value of 50%, the predicted probability for every slice is classified into either of the classes (0 or 1) which is represented as a vector 'pred'. To generate a subject level decision on the patient, maximum voting is applied which can be denoted as

$$y = argmax(pred[i]) \tag{5.1}$$

where the index i is the number of slices per subject and y is the resultant scalar value which is either 0 or 1 indicating the presence of disease in a subject. For the test data, weights trained on the training data are used on the same slices as in training slice set. Fig.5.3 represents the same in a structured approach of a flow diagram.

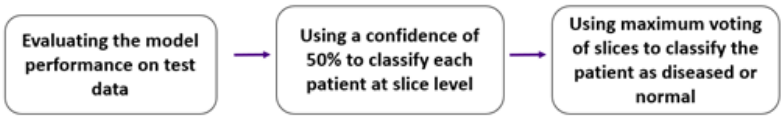



Figure 5.3: Logic followed in classifying results at the subject level

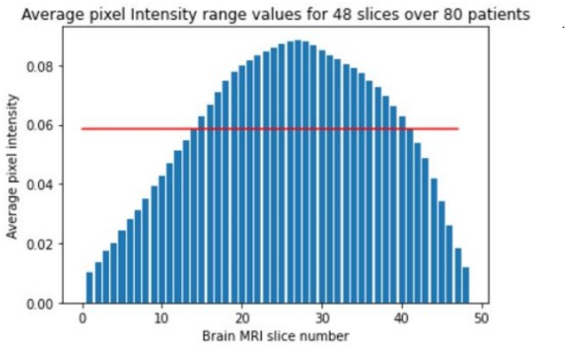



Figure 5.4: Average pixel intensity for 48 slices of 80 subjects

**Region of Interest**

From Fig.3.2, it is clear that the first few slices of the MRI contain mostly black pixels and lack any information pertaining to the brain tissue. Slices towards the end do not have any information about the midbrain which has the features indicating PD. Hence, we hypothesize that slices in the middle range are of importance in this study.

On calculating the average pixel intensity of every slice of the 80 training subjects as shown in Fig. 5.7, it is noted that slice range between 15-42 has pixels of intensity above the mean average pixel intensity value represented by the red line, which is considered as a threshold. The slice range which satisfies this threshold is expected to be the Region of Interest. To validate this, we have compared the performance of DenseNet201 on all the 48 slices and with just the selected ROI slices. On experimentation, it is observed that the slice range of 20-40 gave better results (see Table 5.1). Therefore instead of


Figure 5.5: Confusion matrix representing classification of test results using model trained on: a) 48 slices of 80 patients b) 20 slices of 80 patients

Table 5.1: Comparing performance of DenseNet201 on tuning number of slices

| **Number of slices per subject** | **Model performance over cross validation** | | | | **Test results** | |
|---|---|---|---|---|---|---|
| | **Specificity** | **Sensitivity** | **F1 avg** | **F1 std dev** | **Accuracy** | **F1 score** |
| 48 | **0.89** | 0.66 | 0.708 | 0.026 | 55% | 0.69 |
| **20** | 0.88 | **0.81** | **0.852** | **0.011** | **70%** | **0.7** |

using all 48 slices, only 20 slices of each subject is used. This leads to better detection of PD as a lot of unwanted information gets removed.

The model trained on 48 slices has high Specificity as it learns the features of an NC better and has low Specificity since 5 PD subjects were misclassified as a Normal Cohort. On the other hand, using 20 slices improves the performance. There is notable increase in the values of F1 average score and a reduced variation in the standard deviation. The Sensitivity and Specificity values are balanced, meaning the model is able to differentiate between NC and PD.

**Model generalizability**

To determine the generalizability of the model performance over unseen data, 5-fold stratified cross validation is used resulting in each fold having 20% of the data in

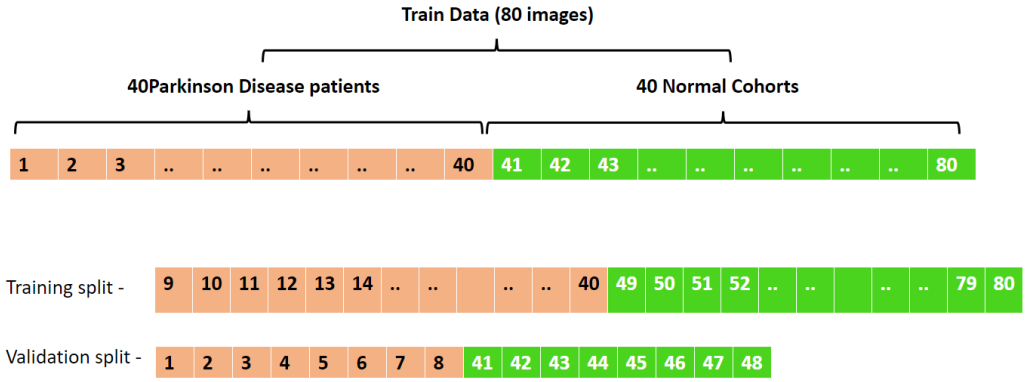



Figure 5.6: Stratified cross validation shown over 1 sample fold

validation and remaining data for training. By using Stratified validation, the NC and PD patients in every fold is equally distributed (see Fig.5.6) and hence give us a good idea on the performance of the model on detecting both classes. The 5-fold split is implemented with a fixed seed number so that the performance can be compared for the 3 different architectures. The metric used to evaluate the performance over each fold is F1 score. Hyperparameters are tuned in a way so as to maximize the F1 average value across the folds and minimise the F1 standard deviation. By doing this the generalizability of the model will be achieved.

The tuned hyperparameter values for each of the model as specified in Table 5.2. Different optimizers like Stochastic Gradient, Adam, AdaBoost and Adagrad have been tried, and Adam was identified as performing the best in backpropagating the loss during training. Binary CrossEntropy loss function is used as we are dealing with a binary classification problem. Addition of a dropout layer enables in training the model for more epochs by which the right features are learnt. This prevents the model from

overfitting. Also, a smaller learning rate is required for the model to be able to identify the difference in brain features of an image having PD and NC. Larger learning rates tends to make the model to be biased towards one particular class. $e^{-4}$ is the learning rate for DenseNet201 and $e^{-5}$ is the learning rate for the other 2 architectures. While evaluating the models on the test data - 20 slices of 20 subjects were tested using the weights learned by the model during training,it is observed that all the 3 models perform as well as the corresponding architecture's crossvalidation results during training

To identify the suitable architecture for the given dataset, 20 slices of 80 subjects have been trained on 3 deep learning architectures and their performances compared. Over the 5 folds during cross validation, we can observe that Specificity of VGG-19 is higher than its Sensitivity which means that it identifies Normal Cohorts better than PD patients. DenseNet201 maintains optimal trade-off between identifying diseased and non-diseased images, i.e Sensitivity and Specificity are in a similar range. And the F1 average score of DenseNet201 is high and standard deviation is the least amongst the architectures. The metric values of cross validation during training are specified in Table 5.3. DenseNet201 is the architecture that generalises data the best amongst the 3 architectures while DenseNet121 performs the least. While evaluating the models on the test data - 20 slices of 20 subjects were tested using the weights learned by the model during training,it is observedthat all the 3 models perform as well as the corresponding architecture's cross-validation results during training (Table 5.4).

Table 5.2: Hyperparameter tuning for the three deep learning architectures

| **Architecture** | **Learning Rate** | **Dropout** | **Number of epochs** |
|---|---|---|---|
| VGG19 | $e^{-5}$ | 0.5 | 10 |
| DenseNet121 | $e^{-5}$ | 0.4 | 8 |
| DenseNet201 | $e^{-4}$ | 0.2 | 5 |

Table 5.3: Comparing model generalizability for training data on ROI slices

| Architecture | **Model performance over 5-fold CV** | | | |
|---|---|---|---|---|
| | **Specificity** | **Sensitivity** | **F1 avg** | **F1 std** |
| VGG-19 | **0.88** | 0.74 | 0.84 | 0.06 |
| DenseNet121 | 0.68 | 0.73 | 0.76 | 0.08 |
| DenseNet201 | 0.81 | **0.88** | **0.85** | **0.01** |

Table 5.4: Comparing model performance over test data

| **Architecture** | **Model performance over test data** | | | |
|---|---|---|---|---|
| | **Specificity** | **Sensitivity** | **Accuracy** | **F1 score** |
| VGG-19 | 0.80 | 0.60 | 70% | 0.73 |
| DenseNet121 | 0.60 | 0.90 | 70% | 0.75 |
| DenseNet201 | 0.70 | 0.70 | 70% | 0.70 |

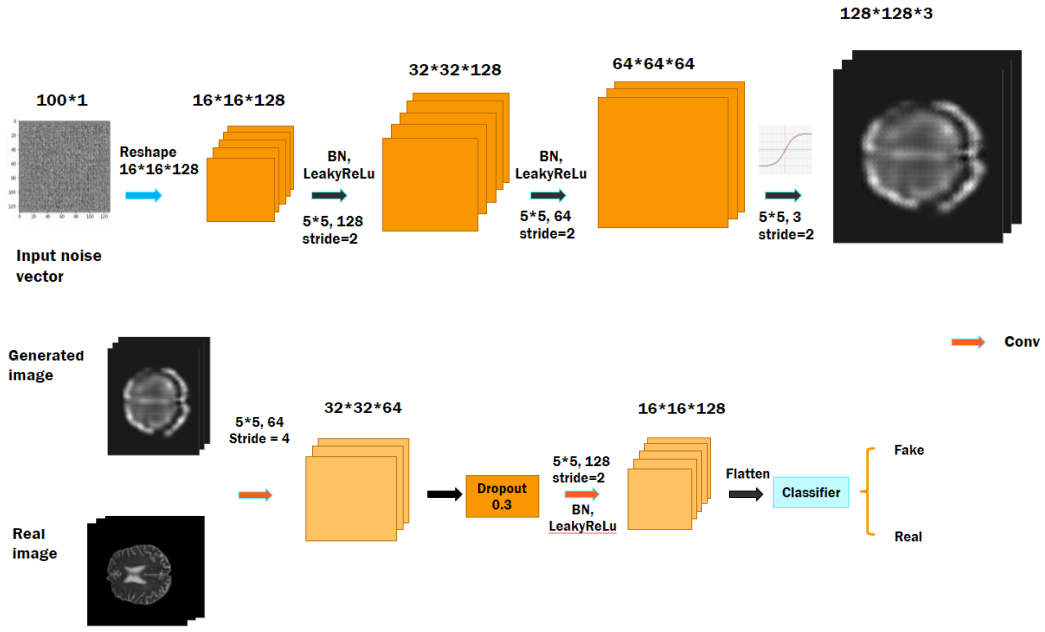



Figure 5.7: Adversarial networks a) Encoder architecture b) Decoder architecture

## 5.2 Generative Modelling

In this section,the techniques used to generated synthetic MR images from 3 models is described in detail.

A noise vector of size 100*1 sampled from normal distribution is the input to the generator of DCGAN which generates images of shape 128*128*3. The discriminator receives real and synthetic images as input and based on the predictions made, the loss function is continually updated. Batch normalisation and Leaky ReLU help in preventing overfitting. Binary cross entropy is the loss function used with label smoothing assigned a value of 0.1 to allow some margin of probabilistic learning of the error **??**. Architecture of DCGAN used is shown in Fig.5.7. The learning rate used is 0.001 and the model is trained for 500 epochs. The quality of images generated over different 5

Table 5.5: Metrics assessing quality of generated images by DCGAN

| **Batch size** | **Mean SSIM** | **Mean PSNR** |
|---|---|---|
| 8 | 64.38 | 66.50 |
| 16 | 64.70 | 67.16 |
| 32 | 65.32 | 67.21 |
| **64** | **65.48** | **67.31** |
| 128 | 61.20 | 62.34 |

Table 5.6: Metrics assessing quality of generated images by WGAN

| **Batch size** | **Mean SSIM** | **Mean PSNR** |
|---|---|---|
| 8 | 60.77 | 66.25 |
| **16** | **61.30** | **66.24** |
| 32 | 60.95 | 66.52 |
| 64 | 59.86 | 66.7 |

different batch sizes ranging between 8 and 128 is experimented. It is noted that as the batch size increases the quality of images generated by DCGAN improves. However, beyond batch size 64 the quality starts decreasing. The same is reflected in Table 5.5. A boxplot is constructed to show the variance in SSIM values (Fig. 5.8). Sample images generated with the best batch size for DCGAN is visualised at a few epochs in Fig. 5.9 which show the gradual improvement in reproducing features of the brain MRI. Time taken for each epoch on an average is 2.5 seconds.

DCGAN tends to have mode collapse i.e., the diversity in patterns of synthetic images is limited. The same is reflected in the fluctuating loss curve of DCGAN in Fig.5.10. This is overcome by replacing the loss function with Wasserstein loss or Earth Mover distance (EMD). It minimises the difference in distribution between original and synthetic images. To use EMD, the norm of the gradient must be a maximum of 1 at every point. This can be accomplished using gradient penalty (GP). Therefore, in

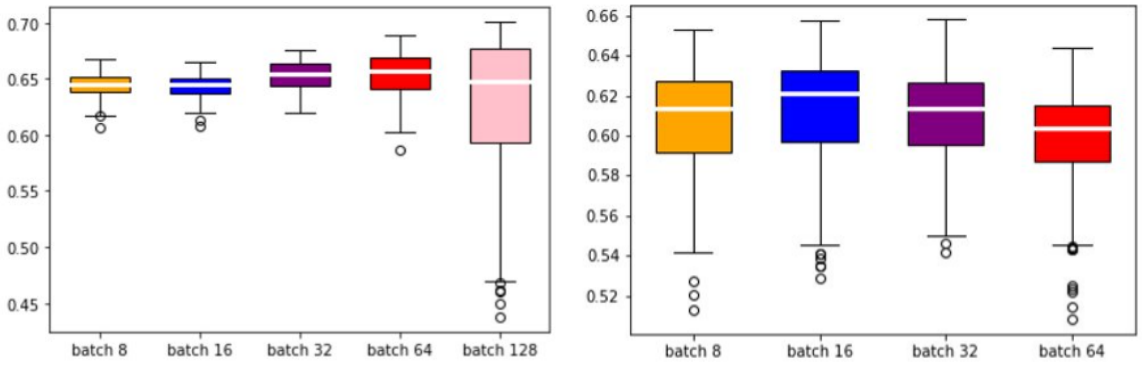



Figure 5.8: Box plot representing the SSIM spread for synthetic images generated over different batch sizes for DCGAN and WGAN respectively

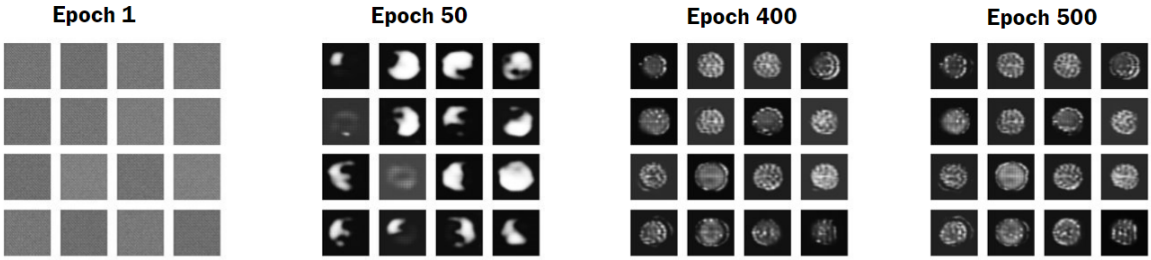



Figure 5.9: Sample images generated for a batch size 64 by DCGAN

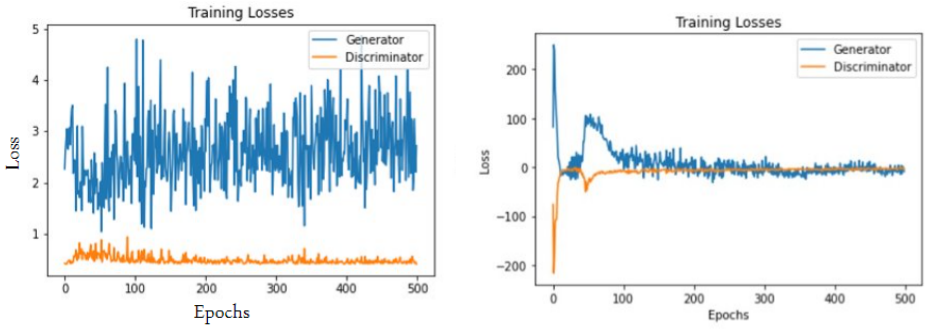



Figure 5.10: Caption

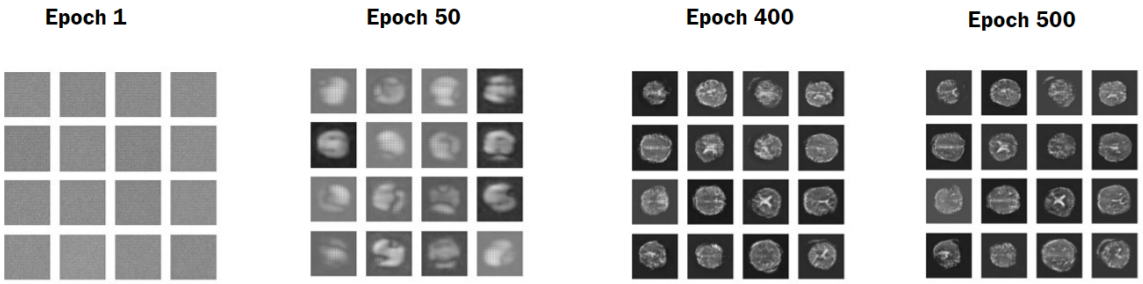



Figure 5.11: Sample images generated for a batch size 16 by WGAN

our work, we use WGAN-GP. Batch sizes are varied as in DCGAN and batch size 16 is found to give best results. Comparison of the image quality metrics over different batch sizes are represented in Table 5.6 and visually depicted in Fig.5.8. There is a significant improvement in images produced by WGAN-GP when compared to DCGAN (Fig. 5.11). However, the converse is reflected in the image quality metrics. The mean values of SNR and SSIM of WGAN is lesser than DCGAN. The training is stable and there are lesser fluctuations in the loss curve as seen in Fig. 5.10. The drawback of this architecture is that the average time taken for each training epoch is 8.7 seconds.

The adversarial architecture is prone to artefacts like checkboard effects which cause a grainy appearance of the synthetic images. We have tried rectifying this by using Near-

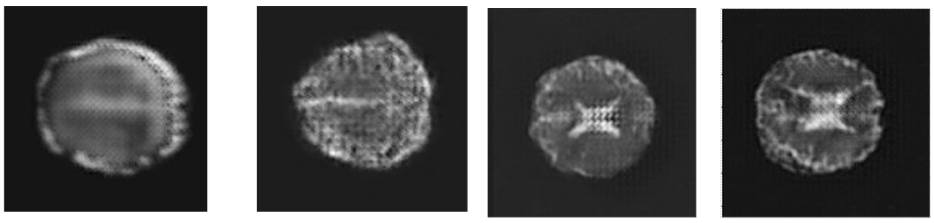


Figure 5.12: Replacing Transposed Convolution with Nearest Neighbour Convolution a) DCGAN b) WGAN

est Neighbour (NN) Convolution instead of Transposed Convolution in the generator ref25. The difference in image quality while using NN convolution is seen from Fig. 5.12. The metrics improve by 2% when replacing Transposed Convolution with Nearest Neighbour Convolution.

*Variational Autoencoder* represents the input data using the latent vector in a probabilistic manner. The encoder outputs a latent vector where each variable represents a probability distribution. The input data is encoded in a compressed format using the mean and variance of the data. The decoder samples values from the latent vector probability distribution and reconstructs the data while being able to maintain the pattern diversity of images. This is ensured by the loss function which comprises of reconstruction loss and Kullback Leiber loss. Kl loss maximises the similarity in distribution between real data and synthetic data. The encoder and decoder architecture used in this work is shown in Fig.5.13. The image is resized into a shape of 128*128*1 to reduce computational costs.

Latent vector dimension is experimented with to find the suitable size to generate NC images. The comparison of sample images produced for dimension 3, 35 and 500 is shown in Fig.**??**. Visually, the quality of images and the pattern decreases with increase

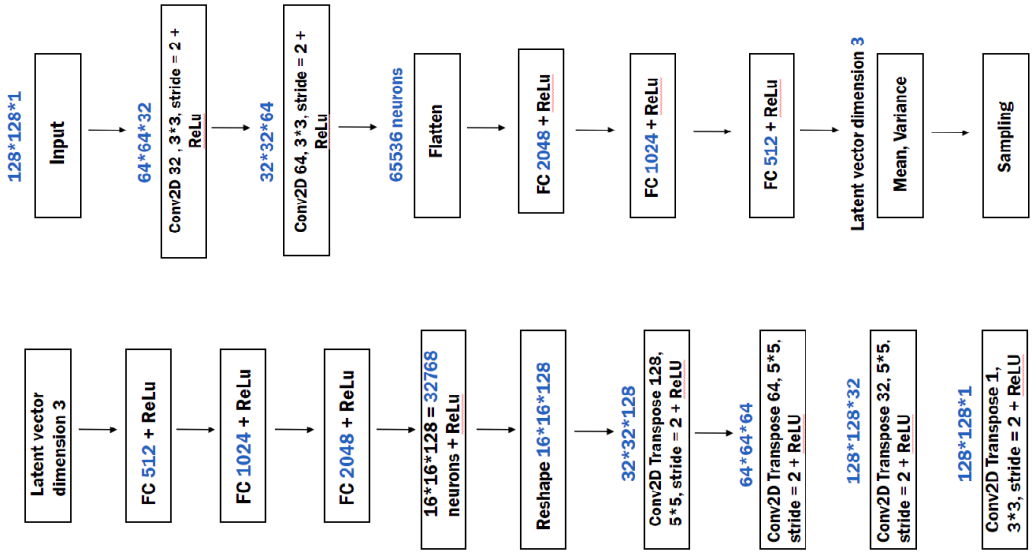



Figure 5.13: Encoder and Decoder architecture of Variational Autoencoder

Table 5.7: Metrics assessing quality of generated images by VAE

| Latent space dim | Mean SSIM | Mean PSNR |
|---|---|---|
| 3 | **68.2** | **68.62** |
| 35 | 66.3 | 67.67 |
| 500 | 66.8 | 67.31 |

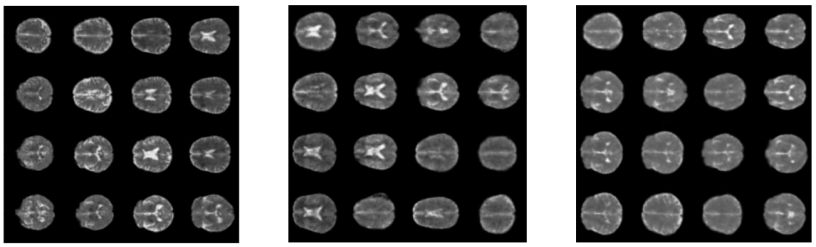


Figure 5.14: Sample images generated by Variational Autoencoder for dimensions 3,35 and 500

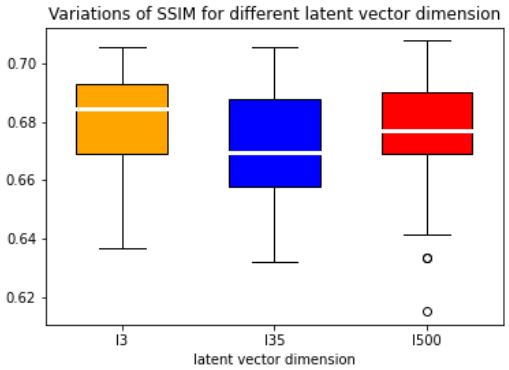



Figure 5.15: Box plot representing the SSIM spread for synthetic images generated over different batch sizes for VAE

in vector size (Fig.5.14), which is reflected in the loss curve and kl loss curve as well (Fig.5.16). Choice of the suitable latent vector dimension is then ascertained through the mean SSIM and PSNR metric values (Table 5.7).

The tuned hyperparameters for this model architecture are a batch size of 32 and the standard Adam optimizer with a learning rate of 0.001 is used. The model is trained for 200 epochs and average time taken for each epoch on average is 5 seconds.

### 5.2.1 Classification using deep learning models

DenseNet201 architecture pretrained on ImageNet weights, is used in our work as the model generalizability is verified and it is sensitive in detecting both NC and PD classes

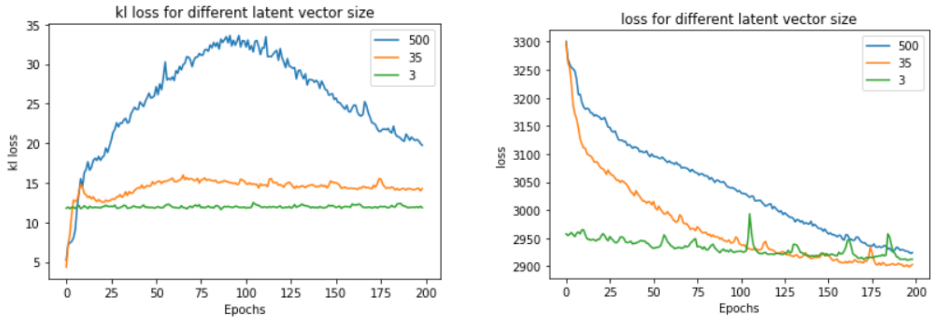



Figure 5.16: Total loss curve and kl loss curve for Variational Autoencoder over the 3 latent vector dimensions experimented

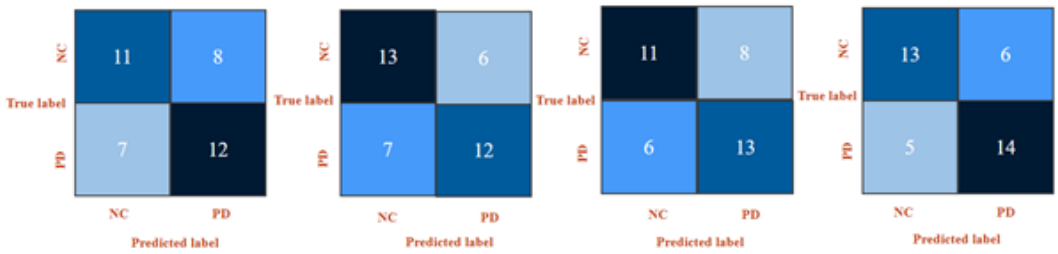



Figure 5.17: Confusion matrix generated by DenseNet201 model trained a) without data augmentation ; synthetic images from b) VAE c) DCGAN d) WGAN

equally well. All the layers of the models are trained and the final layers are tuned to suit our classification problem. A batch size of 32 is used with an Adam optimizer having a learning rate of 0.0001. The weights of the model with the best accuracy over validation set is saved and used on the test data set. The performance of the classifier with and without data augmentation of synthetic images from the 3 generative models is compared and analysed to comprehend the performance. Confusion matrix in Fig. 5.17 represents the number of subjects correctly classified and the misclassifications as well. Precision and F1 score is calculated for each of the class from the confusion matrix to get a deeper understanding of the model's diligence. Being a binary class problem, specificity of PD class is the sensitivity of NC class and vice versa, so we have

Table 5.8: Comparing DenseNet201 classifier model performance with and without augmentation

| | **Specificity** | **Sensitivity** | **Precision** | | **F1 score** | | **Accuracy** |
|---|---|---|---|---|---|---|---|
| | | | **PD** | **NC** | **PD** | **NC** | |
| **Real** | 0.58 | 0.63 | 0.6 | 0.61 | 0.61 | 0.59 | 0.60 |
| **VAE** | 0.68 | 0.63 | 0.66 | 0.65 | 0.64 | 0.66 | 0.66 |
| **DCGAN** | 0.57 | 0.68 | 0.62 | 0.64 | 0.65 | 0.62 | 0.63 |
| **WGAN** | **0.68** | 0.74 | **0.7** | **0.72** | **0.71** | **0.7** | **0.71** |

mentioned the Specificity and Sensitivity only for the PD class in Table 5.8. Considering PD as the positive class, VAE and WGAN have least number of False Positives (FP)and highest number of True Negatives (TNs). They have similar performance in terms of detecting NC class. DCGAN has an improved performance when compared to model without data augmentation, in detecting PD class alone. WGAN has highest number of True Positives (TPs) and least number of FNs. Next comparable performance is with images generated by VAE. Images generated by DCGAN are identified more as PD class (FNs). Performance over test data with model trained after augmentation of all the 3 generative models show a considerable improvement in metrics. Model trained with augmentation of WGAN and real images predicts the negative class with a specificity of 0.68 which is better than specificity of augmentation with DCGAN. It has a higher sensitivity than specificity indicating it identifies PD class better than NC class. Precision for NC class is slightly higher than PD class for both models trained with and without data augmentation. WGAN has the highest precision for both PD and NC classes is highest followed by VAE. F1 score is highest for PD class followed by F1 score for NC by VAE.

Table 5.9: Architectural details of Capsule Network

| **Layer** | **Output shape** |
|---|---|
| **Input Layer** | (10,224,224,3) |
| **Conv2D** | (10,222,222,256) |
| **PrimaryCap_Conv2D** | (10,107,107,256) |
| **PrimaryCap_reshape** | (10,366368,8) |
| **Primarycap_squash** | (10,366368,8) |
| **DigitCaps (Capsule Layer)** | (10,2,16) |
| **Input layer** | (None,2) |
| **Mask** | (10,32) |
| **Capsnet** | (10,2) |
| **Decoder** | (224,224,3) |

The model trained with augmented Normal Cohort images generated by WGAN gives the best results in classification as the FPs and FNs are minimal which is important in medical field. However, VAE does have a good scope, as it is capable of generating new images with lesser artefacts unlike in GANs. Artefacts in synthetic images hinders the subtle learning of features learned by model during training.

Architecture tuned for Capsule Network is shown in Table 5.9. Learning rate is varied between 0.1 and 0.0001. Number of epochs is varied between 10 and 30. Batch size is varied from between 8 and 32. The parameters chosen based on best performance is 0.001 for Learning rate and a smaller batch size of 10. The number of filters is varied from 64 to 256 and kernel size is varied between 3 and 9. It is seen that 128 filters with a size of 3*3 gives the best result. 32 Capsules used are of dimension 8 each, with squash activation function connected to the convolutional layer. The model is trained for 30 epochs and weights of the best model is saved based on improvement in validation accuracy which is used to evaluate model performance over the test data. Training

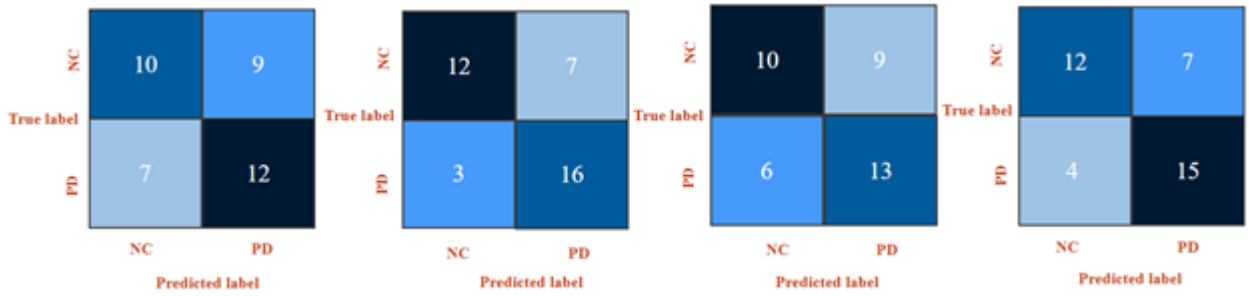



Figure 5.18: Confusion matrix generated by Capsule Network model trained a) without data augmentation; synthetic images from b) VAE c) DCGAN d) WGAN

images from VAE took the longest time; images from the adversarial models took almost the same time to train. 200 synthetic images from each of the generative model is augmented and the CapsNet model trained. Beyond 200 images, the performance of the model saturates The confusion matrix generated by using CapsNet model trained with real and synthetic images on test set is shown in Fig. 5.18. Model trained on synthetic images from VAE have least number of FPs and FNs and highest number of TPs and TNs on the test data. Testing of model augmented with WGAN synthetic images yields the same performance over NC predictions. Number of TPs and FNs of augmentation with WGAN follow closely to augmentation with VAE. All the generative models identify the TPs which is the PD class with least mispredictions than the NC class. WGAN augmented models minimise the gap of correctly predicted NC and PD subjects. DCGAN has a slight improvement in classification of PD subjects, but performance in classification of NC subjects remains the same as in a model trained without data augmentation. Specificity of VAE and WGAN is equal and highest (5.10). Sensitivity is highest for VAE followed by WGAN. WGAN has the highest precision

Table 5.10: Metrics to evaluate Capsule Network architecture with and without data augmentation

| | **Specificity** | **Sensitivity** | **Precision** | | **F1 score** | | **Accuracy** |
|---|---|---|---|---|---|---|---|
| | | | **PD** | **NC** | **PD** | **NC** | |
| **Real** | 0.53 | 0.63 | 0.57 | 0.59 | 0.6 | 0.56 | 0.58 |
| **VAE** | 0.63 | **0.84** | 0.7 | **0.8** | 0.76 | **0.7** | 0.74 |
| **DCGAN** | 0.53 | 0.69 | 0.6 | 0.62 | 0.64 | 0.57 | 0.61 |
| **WGAN** | **0.63** | 0.79 | **0.75** | 0.68 | **0.77** | 0.65 | 0.71 |

amongst the generative models as lesser number of PD subjects are misclassified as NC subjects. Precision of detecting PD subjects is higher than detecting NC subjects. On the other hand, VAE has the highest precision in detecting NC subjects and precision of VAE for NC subjects is more than that of PD subjects. A similar trend follows for the F1 score. VAE has maximum accuracy closely followed by WGAN. Model trained with synthetic images from VAE gives maximum correct classifications (TPs and TNs) and least misclassification (FPs and FNs).

# Chapter 6

# Conclusion

Initially, the data has been split into train and test set at the subject level which avoids data leakage. The slices that are most essential in identifying the presence. of Parkinson's Disease is our Region of Interest, which are found to be in the slice range of 20-40. DenseNet201 is concluded to be the model that generalises best amongst the 3 architectures as indicated by the metrics. Training a model on 20 slices instead of 48 slices shows improved performance in detecting Parkinson's Disease. The test results are classified at the patient level so that the easily interpretable model can be of assistance to the medical personnel. There is a significant increase in performance as indicated by the metrics when using models trained with augmentation over test data. The slices of interest of Normal Cohort subjects are generated using 3 generative models and their performance over different batch sizes are compared for GANs and latent vector dimension for VAE. Based on mean SSIM and PSNR, it is observed that DCGAN generates the best quality images for a batch size of 64, beyond which it deteriorates. WGAN generates images of better quality without mode collapse for a batch size of 16 and pattern diversity in the images are captured well. VAE has the best metrics for a latent

vector dimension of three and, the pattern diversity and image clarity reduces with increase in latent dimension Comparing the image quality metrics for the adversarial models, the magnitude of values for DCGAN is higher than WGAN. Quality metrics of VAE is higher than the adversarial models. However, when augmenting the images and training the deep learning model, WGAN augmented DenseNet architecture has a superlative performance closely followed by model trained with VAE augmentation. For CapsNet architecture, the VAE augmented model performs better over the test results. In both cases, model with synthetic images from DCGAN exhibits a mediocre performance. This is possibly due to the lack of pattern variation in the images generated by the model. Accuracy using CapsNet architecture is 3% more than the accuracy using DenseNet201 as classifier. From the results generated, it is noticeable that VAE which has not been explored much with respect to PD classification has significant potential as a generator model. While using DenseNet201 as a classifier, although augmentation with WGAN gives the best performance in terms of all the metrics, VAE is a close contender where overfitting does not occur in the DenseNet model. As future work, the generative ability of the models can be extended to the PD class for experimentation. It might be required as the intricate features that define PD has to be generated accurately. Different modalities of MRI can be experimented [26] in the perspective of generating synthetic images of high quality. Capsule GAN [27] and VAEGAN [29] are the other architectures which have a high potential produce better and realistic images. CapsuleGANs combines the adversarial abilities of GANs with spatial mapping of features and VAEGANs is expected to generate images without artefacts and able

to generate new features in accordance with the original distribution.

# References


1. Pyatigorskaya N, Gallea C, Garcia-Lorenzo D, Vidailhet M, Lehericy S. A review of the use of magnetic resonance imaging in Parkinson's disease. Therapeutic Advances in Neurological Disorders. July 2014:206-220. doi:10.1177/1756285613511507

2. Sumeet Shinde, Shweta Prasad, Yash Saboo, Rishabh Kaushick, Jitender Saini, Pramod Kumar Pal, Madhura Ingalhalikar, Predictive markers for Parkinson's disease using deep neural nets on neuromelanin sensitive MRI, NeuroImage: Clinical, Volume 22, 2019

3. Pinter, B., Diem-Zangerl, A., Wenning, G.K., Scherfler, C..(2015), Mortality in Parkinson's disease: A 38-year follow-up study. Mov Disord., 30: 266-269.

4. Cheng, H.C., Ulane, C.M. and Burke, R.E., 2010. Clinical progression in Parkinson disease and the neurobiology of axons. Annals of neurology, 67(6), pp.715-725.

5. Sivaranjini, S., Sujatha, C.M.: Deep learning based diagnosis of Parkinson's Disease using convolutional neural network. Multimedia Tools and Applications 79(21-22), 15467–15479 (mar 2019).

6. Wen, J., Thibeau-Sutre, E., Diaz-Melo, M., Samper-Gonz´alez, J., Routier, A.,

Bottani, S., Dormont, D., Durrleman, S., Burgos, N., Colliot, O.: Convolutional neural networks for classification of alzheimer's disease: Overview and reproducible evaluation. Medical Image Analysis 63, 101694 (jul 2020).

7. Chakraborty, S., Aich, S., Kim, H.C.: Detection of parkinson's disease from 3t t1 weighted mri scans using 3d convolutional neural network. Diagnostics 10(6), 402 (2020)

8. Kiranbabu Rajanbabu, Iswarya Kannoth Veetil,V. Sowmya, E.A. Gopalakrishnan, K.P. Soman : Ensemble of Deep Transfer Learning models for Parkinson's Disease Classification (Accepted in fourth International Conference SOFT COMPUTING AND SIGNAL PROCESSING (ICSCSP-2021))

9. Kaur, S., Aggarwal, H. and Rani, R., 2021. MR Image Synthesis Using Generative Adversarial Networks for Parkinson's Disease Classification. In Proceedings of International Conference on Artificial Intelligence and Applications (pp. 317-327). Springer, Singapore.

10. Balasubramanian, R., Sowmya, V..: Analysis of adversarial based augmentation for diabetic retinopathy disease grading. In: 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE (jul 2020)

11. Islam, J., Zhang, Y. GAN-based synthetic brain PET image generation. Brain Inf. 7, 3 (2020).

12. Han, C., Hayashi, H., Rundo, L., Araki, R., Shimoda, W., Muramatsu, S., Furukawa, Y., Mauri, G. and Nakayama, H., 2018, April. GAN-based synthetic brain MR image generation. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (pp. 734-738). IEEE.

13. Kose U., Deperlioglu O., Alzubi J., Patrut B. (2021) Diagnosing Parkinson by Using Deep Autoencoder Neural Network. In: Deep Learning for Medical Decision Support Systems. Studies in Computational Intelligence, vol 909. Springer, Singapore.

14. Q. Xu, Y. Yang, Z. Wu and L. Zhang, "Different latent variables learning in variational autoencoder," 2017 4th International Conference on Information, Cybernetics and Computational Social Systems (ICCSS), Dalian, 2017, pp. 508-511 17.

15. N. Qiang, Q. Dong, Y. Sun, B. Ge and T. Liu, "Deep Variational Autoencoder for Modeling Functional Brain Networks and ADHD Identification," 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), 2020, pp. 554-557

16. Akrami H, A. Joshi A, J. Li, S. Aydore and R. M. Leahy, "Brain Lesion Detection Using A Robust Variational Autoencoder and Transfer Learning," 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), 2020, pp. 786-790

17. Kurup, R.V., Sowmya, V. and Soman, K.P., 2019a, January. Effect of data pre-processing on brain tumor classification using capsulenet. In International

Conference on Intelligent Computing and Communication Technologies (pp. 110-119). Springer, Singapore.

18. Kurup, R.V., Anupama, M.A., Vinayakumar, R., Sowmya, V. and Soman, K.P., 2019b, September. Capsule network for plant disease and plant species classification. In International Conference On Computational Vision and Bio Inspired Computing (pp. 413-421). Springer, Cham.

19. Anupama, M.A., Sowmya, V. and Soman, K.P., 2019, April. Breast cancer classification using capsule network with preprocessed histology images. In 2019 International Conference on Communication and Signal Processing (ICCSP) (pp. 0143-0147). IEEE.

20. Marek K, Jennings D, Lasch S, Siderowf A... The Parkinson Progression Marker Initiative (PPMI). Progress in Neurobiology. 2011 Dec 1;95(4):629-635.

21. McCarthy, Paul. (2020, July 9). FSLeyes (Version 0.34.0). Zenodo. http://doi.org/10.5281/zenodo

22. Fedorov A, Beichel R, Kalpathy-Cramer J, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Magnetic Resonance Imaging. 2012;30(9):1323-1341. doi:10.1016/j.mri.2012.05.001

23. https://medium.com/@lishuo1/which-one-should-you-choose-gan-or-vae-part

24. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A. and Chen, X., 2016. Improved techniques for training gans. arXiv preprint arXiv:1606.03498.

25. Odena, A., Dumoulin, V. and Olah, C., 2016. Deconvolution and checkerboard artifacts. Distill, 1(10), p.e3.

26. Writer, Staff., (2018) 'Five Stages of Parkinson's', parkinsonsresource, 16 May, Available at: parkinsonsresource.org/wp-content/uploads/2012/01/The-FIVE-Stages-of-Parkinsons-Disease.pdf (Accessed: 13 March 2021).

27. Jaiswal, A., AbdAlmageed, W., Wu, Y. and Natarajan, P., 2018. Capsulegan: Generative adversarial capsule network. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops (pp. 0-0).

28. Reimão, S., Pita Lobo, P., Neutel, D., Correia Guedes, L., Coelho, M., Rosa, M.M., Ferreira, J., Abreu, D., Goncalves, N., Morgado, C. and Nunes, R.G., 2015. Substantia nigra neuromelanin magnetic resonance imaging in de novo Parkinson's disease patients. European journal of neurology, 22(3), pp.540-546.

29. Yu, X., Zhang, X., Cao, Y. and Xia, M., 2019. VAEGAN: A Collaborative Filtering Framework based on Adversarial Variational Autoencoders. In IJCAI (pp. 4206-4212).

# List of Publications based on this research work

1. Yamini Madan, Iswarya Kv, Sowmya V, Gopalakrishnan EA, Soman KP, *Deep Learning based approach for Parkinson's Disease Classification using Region of Interest*, 4th International Conference on Intelligent Sustainable Systems.

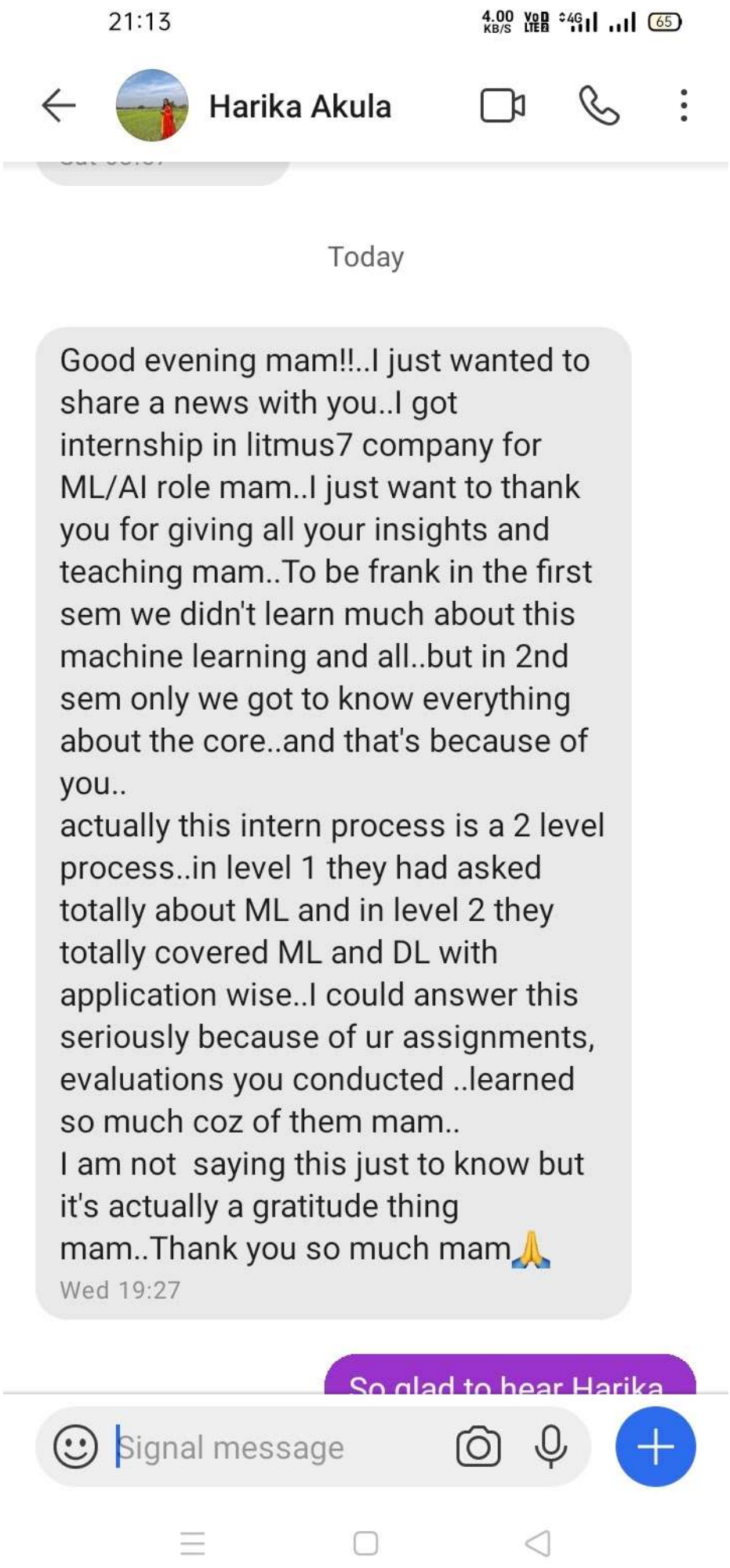
   Accepted and presented on Feb 26, 2021.

2. Yamini Madan, Iswarya Kv, Sowmya V, Gopalakrishnan EA, Soman KP, *Synthetic Data Augmentation of MRI using Generative Variational Autoencoder for Parkinson's Disease detection*, 9th International Conference on Frontiers of Intelligent Computing : Theory and Applications.

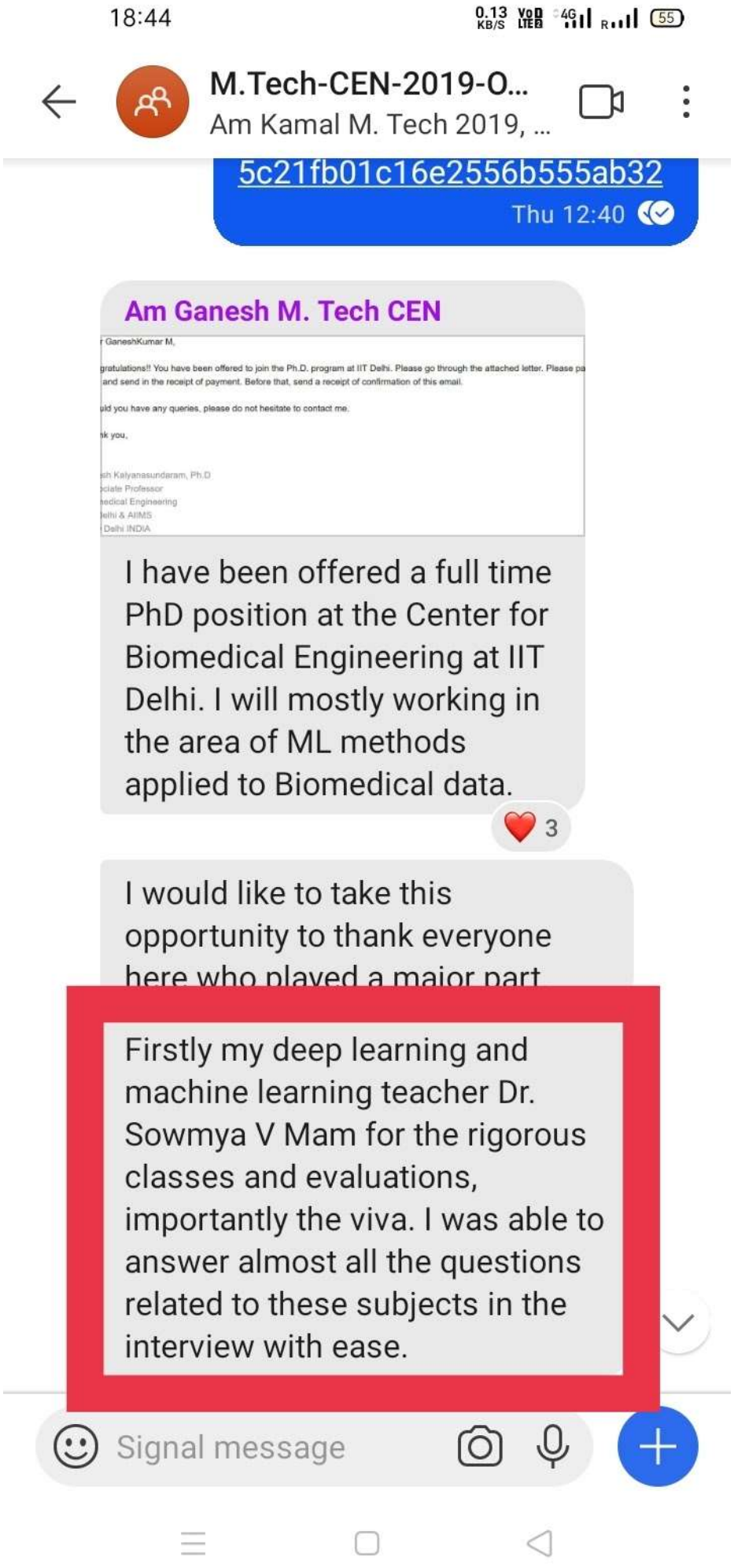
   Accepted and presentation to be held on 25 June 2021.

3. Yamini Madan, Iswarya Kv, Sowmya V, Gopalakrishnan EA, Soman KP, *Parkinson's Disease Classification using MRI images Synthesized by Generative Models*, 4th International Conference on Computational Intelligence and Data Engineering.

   Submitted.


21:13
Harika Akula
Today
Good evening mam!!..I just wanted to share a news with you..I got internship in litmus7 company for ML/AI role mam..I just want to thank you for giving all your insights and teaching mam..To be frank in the first sem we didn't learn much about this machine learning and all..but in 2nd sem only we got to know everything about the core..and that's because of you..
actually this intern process is a 2 level process..in level 1 they had asked totally about ML and in level 2 they totally covered ML and DL with application wise..I could answer this seriously because of ur assignments, evaluations you conducted ..learned so much coz of them mam..
I am not saying this just to know but it's actually a gratitude thing mam..Thank you so much mam🙏
Wed 19:27
So glad to hear Harika
Signal message


21:13
Am Pavan M.Tech
I am in the link 15:01
Today's class 15:01
People ca join as team 15:01
Okay maam 15:02
The next team can join 15:08
Okay ma'am 15:08
This message was deleted 17:03
Ma'am actually I want to say this,
Interview they asked me about my projects and I told them that, I recently started working on CNN they asked me many things in CNN, then the thing is the before day you gave us full review on CNN,
I explained everything in CNN to them like why you have thought us
17:03
*what you have explained us in class
17:04
Thank you ma'am 17:04
Type a message


18:44
M.Tech-CEN-2019-O...
Am Kamal M. Tech 2019, ...
5c21fb01c16e2556b555ab32
Thu 12:40
Am Ganesh M. Tech CEN
GaneshKumar M,
gratulations!! You have been offered to join the Ph.D. program at IIT Delhi. Please go through the attached letter. Please pa
and send in the receipt of payment. Before that, send a receipt of confirmation of this email.
uld you have any queries, please do not hesitate to contact me.
nk you,
sh Kalyanasundaram, Ph.D
ociate Professor
edical Engineering
elhi & AIIMS
Delhi INDIA
I have been offered a full time PhD position at the Center for Biomedical Engineering at IIT Delhi. I will mostly working in the area of ML methods applied to Biomedical data.
3
I would like to take this opportunity to thank everyone here who played a major part
Firstly my deep learning and machine learning teacher Dr. Sowmya V Mam for the rigorous classes and evaluations, importantly the viva. I was able to answer almost all the questions related to these subjects in the interview with ease.
Signal message

keepflying™
reimagine revenue hours

**Letter of Employment**

**Date:** 19th May 2021

**To:**
U. Vamsi Krishna
B2/2 Rites flats Ashok Vihar phase-3,
New Delhi-10052

Dear Vamsi Krishna,

This has reference to your position as Junior Data Scientist, CBMM Supply, Services and Solutions Pte. Ltd, Singapore. We, after discussions with you and assessing the company's needs and requirements, would like to appoint you effective June 1st 2021.

We are happy to offer you this position in our company on the following terms and conditions:

1. **Date of Appointment**

   Your appointment shall be effective from 1st June 2021, and your assignment will be based at Chennai. Remote work arrangements can further be agreed upon in subsequent discussions.

2. **Salary and remuneration**

   Your salary and remuneration will be as per the details herein enclosed under **Annexure A**.

3. **Other Work**

   This is a full-time employment with the company and you shall devote yourself exclusively to the business of the company.

**4. Transfer and Deputation**

You are liable to be transferred in such capacity as the Company may from time to time determine to any other location of the company in the world, without any change in terms and conditions of employment, at the sole discretion of the Board of Directors. In such a case, you will be governed by the rules and service conditions applicable to new assignments and locations.

**5. Confidential Information**

You will not at any time without the consent of the Board of Directors, disclose or divulge or make public except on legal obligations any information regarding the company's affairs or administration or research carried out whether the same may be confined to you or become known to you in the course of your service or otherwise.

**6. Conflict of Interest**

It is intended to avoid conflict between your interest as an employee, and the interest of the Company in dealing with suppliers, customers and all other organizations or individuals doing or seeking to do business with the Company. Further, if any Conflicts of Interest do arise in future, you will promptly report the same to the management immediately.

Noted below are a few examples of Conflict of Interest:

a. You or any dependent member of your family is not to have an interest in any organization, which has business dealings with the company, where there is an opportunity for preferential treatment to be given or received.

b. You or any dependent member of your family are not to buy/sell or lease any kind of property, facilities or equipment from or to the Company or any affiliate or to any Company, firm or individual that is or is seeking to become a contractor, supplier or customer, except with the knowledge and consent of the Management.

c. You are not to serve as an officer, director or in any other management capacity or as a consultant to another company or organization doing or seeking to do business with the Company or an affiliate except with the knowledge and consent of management of our Company.

d. You are not to use or release to a third party any data on decisions, plans, competitive bids or any other information concerning the Company, which might be prejudicial to the interest of our Company.

e. You or any dependent member of your family are not to accept commission, a share in profits or other payments, loans (other than with established banks or financial institutions), services, excessive entertainment and travel or gifts of more than nominal value from any individual or organization doing or seeking to do business with the Company.

**7. Non Compete**

You will not undertake similar assignments from competitors dealing with products similar to those of CBMM Supply, Services and Solutions Pte. Ltd, Singapore, during and for a period of six months after termination of the course of this contract.

**8. Protection of Interest**

If you conceive or develop any new or advanced methods of improving process/formulae/systems in relation to the operation of the Company, such developments will be fully communicated to the Company and shall remain sole right/property of the Company.

**9. Service Rules**

You will be required to go through the Service Rules of the Company on your joining and your appointment will be governed by the spirit and the letter of the existing Service Rules and any modified Service Rules as and when they are enforced from time to time.

**10. Leave & Holidays**

You will be entitled to 12 days leave in a Year, public holidays, Emergency Leave, benefits, and other allowances as applicable to your category of employees and location of posting, in accordance with the rules of the Company.

**11. Past Record**

In the event, during your tenure of service at any time, it is found or revealed that any declarations given or furnished by you to the Company prove to be false or if you are found to have willfully suppressed any material information, in such a case, you are liable for removal from service at any time without any notice or compensation in lieu of the notice period.

**12. Notice Period**

This contract of employment is terminable by either party giving 60 days notice to the Board of Directors. The company reserves the right to pay or recover salary in lieu of notice period. Further, the Company may at its discretion relieve you from such date as it may deem fit even before the expiry of the notice period without compensation for the remaining period and is not bound to given any reason thereof if such termination is for ethical/moral grounds.

**13. On Separation**

On termination of this contract, you will immediately handover before you are relieved all correspondence, specifications, formulae, books, documents, cost data, market data, literature, drawings, effects or records etc., belonging to the Company or relating to its business and shall not make or retain any copies of these items.

**14. Roles and Responsibilities**

**Your roles and responsibilities have been outlined in Annexure B.**

Please feel free to seek clarifications if any concerning this appointment after which kindly endorse and return one copy of this letter signifying your acceptance.

We trust that in the role of **Junior Data Scientist** of CBMM Supply, Services and Solutions Pte. Ltd you will guide and expand the business of the company in the SE Asian markets and find your association with the company both personally and professionally rewarding.

With all our best wishes,

Yours Sincerely,
**For CBMM Supply, Services and Solutions Pte. Ltd**

**Pranatharthi Haran Sriram**
**CEO and Managing Director.**

**ACCEPTANCE**

I accept this offer of appointment on the terms and conditions stated above.

**U. Vamsi Krishna**

Date: 1st May 2019

**Annexure A**

| | |
|---|---|
| Name | U. Vamsi Krishna |
| Designation | Junior Data Scientist |
| Location | Chennai |
| Date of Joining | 1st June 2021 |
| Salary Monthly | |
| Gross Monthly Salary | INR 75,000 |
| Total A | INR 75,000 |
| | |
| Total B (any other compensation) | Nil |
| Total Compensation (A+B) per month in INR | INR 75,000 |

Remarks: Any Income Tax shall be deductible if applicable.

**For CBMM Supply, Services and Solutions Pte Ltd**

**Pranatharthi Haran Sriram**
**CEO and Managing Director.**

**Annexure B**

**Job Description – Junior Data Scientist**

## Responsibilities

- Designing and building efficient, scalable, and resilient micro services that run on AWS or Azure. The champion developer is expected to understand the full scope of a feature, how it will be realized in our UI, used by our customers, and how our system will perform and scale.
- Committing tested, documented, and reviewed code on a frequent basis, ideally daily. Code reviews and automated testing are core to our quality approach.
- Collaborating effectively with various stakeholders such as Product Director, Front end UI/UX team, and data engineers.
- Deep architectural understanding of good SaaS deployment patterns, and the technical options available from the various Cloud providers such as AWS or Azure.

## Required Skills

- Professional experience using Python.
- BS in Computer Science or equivalent experience.
- Proven development experience of RESTful APIs and data analysis libraries such as Pandas, Koalas, etc.
- Experience with the automated build process (continuous integration) and source code version control tools (GITHub).
- Experience with Exploratory Data Analysis including wrangling, grooming, transformation, and analysis.
- Experience applying statistical methods to solve data problems
- Experience with Microservice Architecture technologies and/or implementations.
- Must have AWS Lambda/Azure Functions, Batch, Serverless, ECR experiences.
- Ability to work in an agile environment.
- Familiar with integrating with data sources such as MongoDB, PostgreSQL, or SQL Server.
- Strong analysis and problem-solving experience.

## Preferred Skills

- Experience using large-scale distributed frameworks such as Azure Synapse, Apache Spark or Tensorflow.
- Experience in data pipeline tools Apache Kafka, SSIS, Azure Data Factory, Data Bricks and Talend.
- Experience with NoSQL technologies such as MongoDB, and Cassandra.

**For CBMM Supply, Services and Solutions Pte. Ltd**

**Pranatharthi Haran Sriram**
**CEO and Managing Director.**

**Bhanu Kumar M**
**IT, Chennai**

Date: 28-Jul-2021

**OFFER OF APPOINTMENT**

Dear **Bhanu Kumar M**

Welcome to the **OLAM** team!!

We are pleased to offer you appointment as **"Engineer (IS2)"** in our Company. The terms and conditions of your appointment are given below:

1. You will be paid an annual Gross salary of **INR 600,000/-**cost to company basis, which includes all your perquisites and allowances.

2. The breakup of the salary is detailed in the Annexure I. Your average annual target incentive will be **INR 75,000/-.** Please note that the actual incentive amount will depend upon the Company, Business Unit and your Individual Performance and the actual payout may vary depending upon these parameters.

3. Your next salary revision will be in **January 2023** as part of the appraisal cycle for the FY **2022**.

4. You will be covered by the Company's Mediclaim Policy and Gratuity as per company rules.

5. You will be on probation for a period of **6** months from the date of joining viz.,**02-Aug-2021,** which may be extended by another **Six** months, based on performance.

6. This offer of appointment is valid only till the date of joining you have accepted and committed as above and it will automatically cease in the event of your not joining us by the said date.

7. This appointment is terminable from either side, by giving:
a) ONE months' notice in writing or salary in lieu of notice, during Probation period
b) THREE month's notice in writing or salary in lieu of notice, after the Confirmation of the services with the company.

8. During the course of employment with the company, you shall not enter the service of employment, consultancy, full or part time, of any other person or organization or yourself carry on or be interested in any business.

9. You shall not, either during or after leaving employment of the company, divulge, make known or communicate to any other person or persons, firm, company, concern or yourself make use of any secrets or information, which you may acquire, receive or obtain in relation to the affairs of the company, or any other matter, which comes to your knowledge in the course of, or by reasons of your appointment with the company, except with the consent in writing from the company.

10. You will be entitled for leave as follows:
**36 days per year** comprising-12 Sick Leave, 12 Casual Leave & 12 Paid Leave.

11. Your appointment, continuation and permanency will always be subject to your remaining physically and mentally fit and alert considering the nature of your duties. The Management has every right to get you medically examined or re-examined at any time by the registered Medical Practitioner, or Eye Specialist or a Civil Surgeon appointed by the Company whose findings will be final and binding upon you.

12. You will be governed by the Policies of the company as may be applicable to you from time to time.

13. Your initial place of posting will be in Chennai. You are liable to be transferred from one job to another job or from one department to another department or from one establishment to another establishment if required by the Management. You shall do such other work, which will be assigned to you by the Management from time to time. Any such changes in assignment or transfer will not automatically entitled to any additional remuneration, allowance, compensation, or other sum in respect thereof.

14. It is also expressly agreed to by and between us that the Company shall be entitled to loan or transfer your services, provisionally for any duration or permanently, wholly or partly to any Company which is or at the material time may be an associate, affiliate, successor, assigns or subsidiary or principal contractor to, or the latter having a controlling interest in the said company.

15. You will report to **Mr.Rajaram K, Manager** or his nominee.

16. Your individual compensation is strictly between yourself and the company. It has been determined based on numerous factors such as job role, skills — specific background and professional merit. This information and any changes made therein should be treated as personal and confidential.

17. Before preceding on an overseas assignment you will be required to give the company, a written undertaking for dedicated services to the client, completing the work/project assigned and timely return to resume work in India. The details of such assignments including reimbursement of necessary expenditure will be communicated to you before your proceeding on such assignments.

18. While serving the Company, you shall give and devote the whole of your work day exclusively to your duties with the Company and shall not engage yourself, directly or indirectly without prior consent in writing of the Company with or without remuneration in any trade, business, occupation, employment, service or calling which is similar to or the same as that carried out by the Company nor shall you undertake any activities which are contrary to or inconsistent either with your duties and obligations under this appointment or with the Company's interests.

19. It is your responsibility to notify the company of any changes in your personal information (like address, contact phone number, additional qualifications, marital status, change of nomination, passport details, etc.) within 3 working days.

20. Your designation may be changed at the discretion of the company depending on the work assigned to you.

21. Upon your resignation or retirement from the company or termination of your services, you are required to return all assets and property of the company such as documents, machines, data, files and books etc. (including but not limited to leased properties).

22. You will retire in the normal course from the services of the company at the end of the month in which you attain the age of superannuation, which is 60 years.

23. This is a position of continuous responsibility and does not entail payment of extra time or overtime.

24. You may be selected and sponsored by the Company for training assignments with company's associates or other institutions abroad. You will diligently and beneficially, take part in such training and assignment. In such event, you will continue to serve the company after such training, for a minimum period as may be stipulated.

25. All programs, system logins, manuals, literatures etc. developed by you while in company service will at all times be deemed to be the sole property of the company. Also the company will at all times have the sole proprietary right in any new system which you may develop while in company's service.

26. You shall obtain written permission from the Management Team for any studies giving full details of examination and duration. While following studies, the duties and accountabilities of your job will not be compromised and the demand arising out of work will prevail over.

27. On the day of joining, you are requested to be present at **9:30 am** for your on-boarding formalities at Olam Information Services Pvt. Ltd., 12th Floor, Zenith Building, Ascendas IT Park, Chennai.

28. You are expected to remain in duty throughout the business/working hours of the organization and be present in time for any meeting or get together scheduled by the company.

29. For the purpose of this clause, the expression "The Company" shall in addition to Olam Information Services Private Limited, mean and include any firm, person or Company subsidiary to or affiliated to with Olam Information Services Private Limited.

30. If any declaration given or furnished by you to the company in any document submitted for employment proves to be false or if you have willfully suppressed any material information, you will be liable to be terminated without notice.

31. We request you to produce a proof of age, relieving letter from the previous employers, educational certificates, 3 passport size photographs and proof of last drawn salary with a true copy of the same for our records.

32. You are advised to keep all original certificates and passport with you all the time to enable you to produce the same at short notice, if required for visa processing purposes.

33. You shall be governed from time to time by the laws of the land as applicable to an employee in the company's service.

34. As substantial amount of technical and other information will be obtained by you or will be available to you, you will appreciate that any information so obtained must not be communicated directly or indirectly to any person, firm or company. You will therefore be agreed to sign a Secrecy Agreement of Non-Disclosure / Confidentiality.

35. Non-Disclosure Agreement:

"Company" for all purposes shall mean Olam Information Services Private Limited, Chennai.
(a) You shall not, at any time during the continuance or after the termination of your employment hereunder, divulge either directly to any person, firm or Company or use for yourself or another any knowledge, information, formulae, processes, methods, compositions, ideas or documents, concerning the business and affairs of the company or any of its dealings, transactions or affairs which you may acquire the company or any of its dealings, transactions or affairs which you may acquire or have to your knowledge during the course of and incidental to your employment.

(b) You will not undertake business of similar nature with any other company during the period of your employment with this company.

(c) If, during the course of your employment with the Company, you are provided with any Company assets, you shall maintain the same in good working condition and you shall return the items to this company prior to you separating from the services of the company. Any dues to be paid to you on your ceasing to be in the employment is liable to be withheld by the Company if the said items so provided by the Company are not returned to the Company, apart from the Company's right to proceed against you as per the provisions of law.

(d) You shall surrender all the records, correspondence and such of the papers connected with the business in the eventuality of your ceasing to be in the employment of this Company.

(e) During your employment with the Company, you shall be subject to, and have to abide by, the rules and regulations stipulated by the Company. The Company may, at its discretion, modify, from time to time, the rules and regulations, as it deems fit, without notice.

We have pleasure in welcoming you and looking forward to mutually meaningful association.

Yours truly,
For **Olam Information Services Private Limited,**

**Arvind Raj B**
**Vice President – Human Resources**

(The appointment letter is enclosed in duplicate and you are requested to sign the duplicate copy of the appointment letter.)

---

I have Read, Understood and Accepted the terms and conditions of employment. As desired, I will join the company's services w.e.f. _____.

Signature:……..  Date:……

**ANNEXURE – I**

Name: **Bhanu Kumar M**

Grade: **IS2**

Date of Joining: **02-Aug-2021**

The breakup of the Salary is given below:

| Salary Component / Break-up | | Monthly Amount (INR) | Annual Amount (INR) |
|---|---|---|---|
| Fixed Pay | Basic | 20,000 | 240,000 |
| | HRA | 10,000 | 120,000 |
| | Special Allowance | 10,100 | 121,200 |
| | **TOTAL** | **40,100** | **481,200** |
| Reimbursements | Telephone | 3,000 | 36,000 |
| | LTA | 2,500 | 30,000 |
| | Food Coupons | 2,000 | 24,000 |
| | **TOTAL** | **7,500** | **90,000** |
| Retirals | PF | 2,400 | 28,800 |
| | **TOTAL** | **2,400** | **28,800** |
| **GROSS CTC** | | **50,000** | **600,000** |
| Performance Incentives | | | 75,000 |
| **TOTAL CTC** | | **50,000** | **675,000** |

**Arvind Raj B**
**Vice President – Human Resources**