Retinal OCT Fluid Detection And Segmentation

Report of CS450 Major Project
Submitted in partial fulfillment of the requirements for the degree of

BACHELOR OF TECHNOLOGY
in
COMPUTER SCIENCE AND ENGINEERING
by

Akshay Dhayal (181CO105)
Chemat Wangail (181CO114)
Saswat Kumar Nayak (181CO147)

under the guidance of

Dr. Jeny Rajan

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA
SURATHKAL, MANGALORE - 575025

April, 2022

DECLARATION

We hereby *declare* that the Project Work Report entitled "Retinal OCT Fluid Detection And Segmentation", which is being submitted to the **National Institute of Technology Karnataka, Surathkal**, in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science in the Department of Computer Science and Engineering, is a *bonafide report of the work carried out by us.* The material contained in this Report has not been submitted at any University or Institution for the award of any degree.

Registration Number, Name & Signature of the Student(s)

(1) 181CO105 - Akshay Dhayal -

(2) 181CO114 - Chemat Wangail -

(3) 181CO147 - Saswat Kumar Nayak -

Department of Computer Science and Engineering
Place: NITK, Surathkal
Date: 29/04/2022

CERTIFICATE

This is to *certify* that the Project Work Report entitled "Retinal OCT Fluid Detection And Segmentation", submitted by

Sl. No., Registration Number & Name of the Student(s)

(1) 181CO105 - Akshay Dhayal

(2) 181CO114 - Chemat Wangail

(3) 181CO147 - Saswat Kumar Nayak

as the record of the work carried out by them, is *accepted as the B.Tech. Project Work report submission* in partial fulfillment of the requirement for the award of degree of **Bachelor of Technology** in Computer Science and Engineering in the Department of Computer Science and Engineering.

Guide
Dr. Jeny Rajan

Dept of Computer Science and Engineering
NITK Surathkal, Mangalore

ABSTRACT

Optical coherence tomography (OCT) may produce micrometer-resolution 3D pictures of retinal structures as a non-invasive imaging method. As a result, it's often employed in the detection of retinal illnesses of fluids involving edema in the retinal layers and underneath them. This imaging technique is beneficial for detecting disorders linked with fluid accumulations, such as diabetic macular edema (DME) or age-related macular degeneration (AMD). A fully convolutional neural network is built using raw pictures and layers that have been segmented using a graph-cut technique.

The flowing pixels were recognised and labelled by the network. To discover the segmented fluid zones, two deep learning model i.e Standard U-net Segmentation and DeepLabv3 Model are being applied.To create the segmentation model , this task requires a dataset named RETOUCH dataset. This dataset has a lot of patients Retinal OCT with fluid problems and with labeled fluid segmented images.

Keywords— Retinal fluid, OCT, Fully convolutional network,U-Net, DeepLabV3

Contents

Chapter 1

INTRODUCTION

1.1 OVERVIEW

Human eyes are important organs that can sense light and provide colour vision in binocular pairs. The eyes are susceptible to a few retinal fluid disorders as people age and external factors influence them. Intraretinal fluid (IRF), subretinal fluid (SRF), and pigment epithelial detachment are the most common components of retinal fluid. These are also known as age-related macular degeneration (AMD) and retinal vein occlusion (RVO). Medical images that are considered to be useful for distinguishing retinal pathology can be used to detect and segment the retinal fluid occupation area.

Optical coherence tomography (OCT) is a widely used medical imaging technology that has advanced rapidly over the last three decades. Huang et al were the first to propose such an imaging technique in 1991. The OCT technique detects back scattered near-infrared light and reconstructs the depth profile of the biological tissue sample using the basic principle of the low coherent light interferometer. Because of its high resolution, the OCT modality is a good choice for cross-sectional imaging of the retina. The fluid area can be visualised with different reflectivity measures from the surrounding tissues using images of the retinal structure obtained by the OCT modality. The processing and analysis of OCT images may aid in distinguishing between different retinal fluid conditions and assessing the progression of retina pathology. Figure below depicts one typical OCT image segmentation. .0

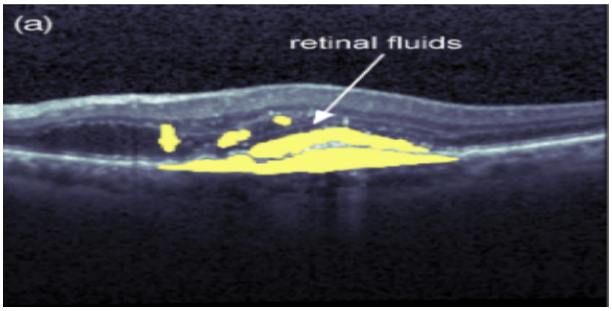

Figure 1.1: Retinal fluids containing OCT image

Computer vision and deep learning algorithms have greatly improved and changed in recent years. Deep learning network topologies, in compared to classic neural network approaches, often include many more hidden layers with tremendous scalability to automatically extract morphological characteristics from raw picture data. Deep learning algorithms have had a substantial influence on retinal fluid segmentation using OCT images since 2016. Convolutional neural networks (CNN), fully convolutional networks (FCN), U-shape networks (U-Net), and hybrid computational approaches are the most prominent deep learning frameworks for segmenting retinal fluids. Deep learning's main technique is to recognise the contours of retinal fluids and then solve a classification issue using the semantic context retrieved from OCT pictures.

Figure below shows the number of academic publications on deep learning for retinal fluid segmentation published between 2017 and 2020. According to Web of Science statistics, the number of relevant articles on retinal fluid segmentation subjects has steadily increased each year, from one in 2016 to 32 in 2020. OCT, retinal fluid, segmentation, IRF, SRF, PED, AMD, and macular edema are among the publications' keywords. The current success of deep learning architectures for retinal fluid segmentation necessitates a review of the recent progress of retinal fluid segmentation-related works. This study examines the segmentation of three retinal fluids in OCT pictures, namely the IRF, SRF, and PED, using various deep learning architectures.

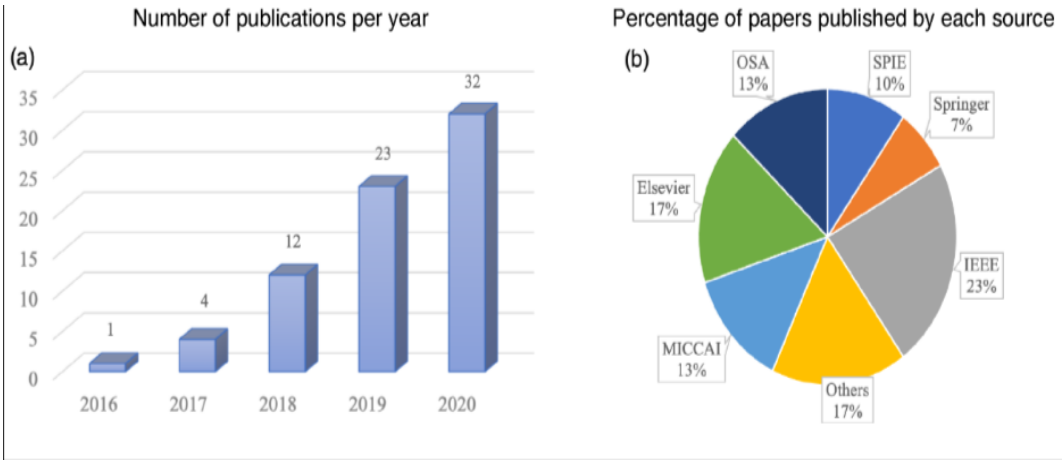

Figure 1.2: Statistics of the papers and their sources on the deep learning methods for retinal fluid segmentation published in recent years.

1.2 MOTIVATION

A good amount of recent work has tried to develop an efficient model to detect and Segment the Retinal OCT Fluids. We want to make easier for human being to detect Retinal related diseases early. The Goal is to create a deep learning model which can accurately detect age-related macular degeneration (AMD) disease. It should also be capable of segmenting the different types of retinal fluids like IRF, SRF, PED which are mainly responsible for the cause of Macular Edeme disease. In this Report, we have presented results for automatic Retinal fluids segmentation with the standard U-net architecture and DeepLABV3 Architecture.

Chapter 2

LITERATURE REVIEW

2.1 BACKGROUND AND RELATED WORKS

In this age of technology and medical advancement, it is important for the devices and doctors to be able to operate in unison for curing the ailment. Along with expertise of doctor, there must be a reliable and efficient device for the doctor to maximize his/her potential. Along this road, there is a specific disease which causes people to lose there sight. These are called **Macular edema**. Macular Edema is the building of fluid in the macula, present in the centre of retina. The problem is presence of different types of fluids in macula causing blindness. Macular Edema can either be age related or due to diabetes.

In [3]study, the author proposes two different deep learning model for segmentation of fluid in the OCT scans, deep learning model being **U-Net** and **DEEPLABV3**. Goal of the study is to segment various fluids such as 'Intra Retinal Fluid'(IRF), 'Sub Retinal Fluid'(SRF) and 'Pigment Epithelial Detachment'(PED)

According to the authors of [1] paper, ML has started a new way in the fight against macular Edema. They have only utilized simple algorithms so far, but there's a lot of room for improvement. The use of CNN illustrates that there are many opportunities that can be taken advantage of. Backbone network that have been pre-trained, such as ResNet and Xception, could be employed for feature extraction. A comparative analysis among various segmentation models on RETOUCH dataset is conducted.

In the research paper [10], the author designs segmentation model based on U-Net architecture, using layer segmentation as pre-processing algorithm and morphologic operations as post-processing algorithm. The drawback regarding the author's model was it only segmented IRF(Intra Retinal Fluid) and with a low F1-score. The reason being for such a low performance from the model was absence of larger high quality data set, as model was trained only on OCT images of 30 patients.

With respect to [10], the author of [9] implemented an improvised version of retinal segmentation, using modified 'U-Net adversarial' architecture. The results were better by a significant margin of 3%. Subsequently, the advised model in [9] was joined in the "RETOUCH CHALLENGE".

In the documentation of [1] research paper, an observation at the generic pipeline to perform semantic segmentation on various background work such as mentioned in [4], [8] and [5] was mentioned. The pipeline consists of:

1. B-Scan extraction(OCT images read)
2. Pre-Processing section(data augmentation, resize)
3. Deep learning architecture(FCN, U-Net, DeepLab)
4. Post Processing(Random Forest Classifier, Median Filter)
5. Segmented image extraction

With respect to the segmentation model proposed by author of the paper [5],our proposed model of DEEPLABV3 is compared with for performance evaluation. In the base-line model, there are certain pipeline components which are not utilized in proposed system, such as **absence of post-processing component** and **utilization of U-Net architecture**.

2.2 OUTCOME OF LITERATURE REVIEW

After going through the above mentioned papers we observed that data processing is playing an important role because it makes data suitable for training purposes. Preprocessed data has helped the model to get the feature to segment fluids of different type. It has also helped to achieve good accuracy. Machine learning algorithms and deep learning algorithms play a key role in semantic segmentation.These algorithms are getting high accuracy upto 98 percent. Although the fight against macular edema can not be won,the damage can be reduced by using effective algorithms. Segmenting types of fluids in retina is and will always be a crucial component to detect functionality and degradation of sight.

2.3 PROBLEM STATEMENT

Accurate segmentation of fluids in retinal region, where various classifications of fluid are accumulated, using deep neural network architectures

2.4 OBJECTIVES

- To make models based on deep neural networks to detecting and segmenting the fluid in B-scans.
- Comparison of experimental results between numerous models on different datasets and different extraction techniques also to find out which model gives the best results.
- To perform a data analysis on various datasets and use it to make some informed decisions.
- To understand the applicability of CNN-based deep models for fluid segmentation in OCT images.
- To perform a comparitive analysis of proposed DEEPLABV3 deep learning model with the base-line model mention in citation[5].

2.5 GOALS

Below mentioned are the goals which are expected to be achieved via the two proposed deep learning architecture:

- Capability of the model to accurately segment regions for different types of fluid present in various OCT images from each OCT obtaining vendor individually.
- Obtaining a better segmentation model from related work on basis of utilized performance metrics.

Chapter 3

PROPOSED METHODOLOGY

3.1 U-NET ARCHITECTURE

As previously stated, the U-net network is divided into two sections: The first is the contracting route, which employs standard CNN architecture. The contracting route is made up of two 3 3 convolutions followed by a ReLU activation unit and a max-pooling layer in each block. This pattern is repeated several times. The U-innovation net's comes in the second phase, dubbed the expanding route, in which each step uses 2 2 up-convolution to upsample the feature map. The feature map from the contracting path's corresponding layer is then cropped and concatenated onto the upsampled feature map.

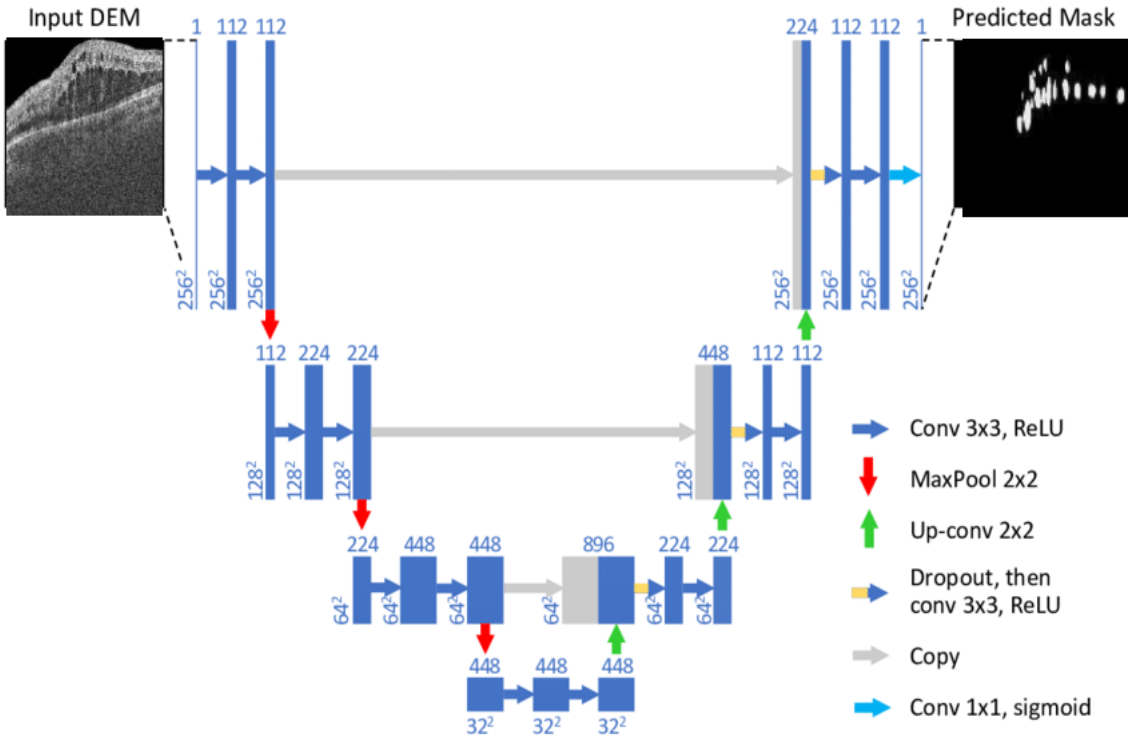

Figure 3.1: Basic U-net architecture with arrows representing the different operations

Then there are two 3x3 convolutions in a row, followed by ReLU activation. To

decrease the feature map to the appropriate number of channels and produce the segmented image, an additional 1x1 convolution is done at the end. Cropping is required because pixel characteristics near the edges contain the least amount of contextual information and must be deleted. This creates a 'U-shaped' network and, more significantly, propagates contextual information along with it, allowing it to separate items in a given area using context from a larger overlapping area. The full U-net architecture is depicted in Figure.

3.2 DATASET

The RETOUCH dataset comes from the MICCAI 2017 retinal OCT fluid challenge[2], in which the OCT images were labelled with three types of retinal fluid: IRF, SRF, and PED. The macular edema caused by AMD was diagnosed in half of the patients, whereas the edema caused by RVO was found in the other half. Because the competition's testing data isn't open to the public, the OCT image data accessible to all researchers serves as the training set. There are a total of 70 OCT volumes in the training data. Cirrus (Model: 5000), Triton (Model: T-1000/T-2000), and Spectralis OCT systems, in particular, acquired 24, 22, and 24 volumes, which have been categorised as IRF, SRF, PED, and normal, respectively. Within each volume, there are 128 (512 1024 pixels), 128 (T-2000: 512 885 pixels, T-1000: 512 650 pixels), and 49 (512 496 pixels) B-scan images acquired with Cirrus (Carl Zeiss Meditec Inc., Jena, Germany), Triton (Topcon Corporation, Tokyo, Japan), and Spectralis (Heidelberg Engineering Inc., Heidelberg, Germany). At least one fluid is seen in each of these B-scan pictures. The annotations and volume of this dataset were received from the Medical University of Vienna (MUV) in Austria, Erasmus University Medical Center (ERASMUS) in the Netherlands, and Radboud University Medical Center (RUNMC) in the Netherlands. Human graders from the MUV and RUNMC clinical facilities manually wrote comments on the B-scan plane. An ophthalmology resident supervised four MUV graders, who were instructed by two retinal experts. A retinal expert was in charge of two RUNMC graders. The RETOUCH dataset was used in the majority of the relevant research discussed in this review study.

3.3 DATA PREPROCESSING METHODS

Before utilizing the compiled data for training, we need to utilize certain processing techniques to further refine it. This is required as it aids us in reducing the size of the actual data and dramatically reducing the Model training time.

3.3.1 GETTING IMAGES PATH

First we create a CSV File that stores all the images name to be used later. After then, we create 2 arrays image path and label path to store all input images and output images path in them.

3.3.2 RESIZE AND NORMALIZE IMAGES

The images in Input and label folder have different resolutions, so we need to first resize them in 1 resolution. First we read both input and label images and normalize them so that they lie between 0 and 1.Then they are resized to dimension (128 * 128). By using higher or better resolution, we can get more accurate results but it will dramatically increase the time taken to train the model.

3.4 CONSTRUCTING U-NET ARCHITECTURE

We can now continue to develop the U-Net architecture after importing the essential libraries. You may accomplish this in a single class by declaring all of the parameters and values in the correct sequence and continuing the process until you reach the end, or you can do it in a few iterative phases. I'll adopt the former way since it's easier for most people to grasp U-Net's architecture with the aid of a single block. As illustrated in the architectural diagram, we'll use three iterative blocks: the convolution operation block, the encoder block, and the decoder block. We can easily construct the U-Net architecture using these three elements. Let's go over each of these function code blocks one by one and see how they work.

Algorithm 1 U-Net Function

Input :img_size

Contraction Path :

- 2 times 3x3 Convolution followed by ReLu

- Max pooling layer to down-sample image

- Repeat above until bottleneck

Expansion Path :

- 2x2 up-Convolution to up-sample feature map

- feature maps from contraction cropped and concatenated

- 2 times 3x3 Convolution followed by ReLu

- Convolution 1x1 followed by Sigmoid function

If you try to recreate the full U-Net design in a single layer, you may discover that the total structure is fairly large due to a large number of distinct blocks to process. We can easily design the U-Net architecture in a few lines of code by breaking our respective functions into three independent code blocks: convolution operation, encoder structure, and decoder structure. The input layer, which will include the appropriate forms of our input picture, will be used.

Following this step, we'll gather all of the primary and skip outputs and send them on to the next block. We'll build the next block and the rest of the decoder architecture until we get to the output. According to our intended output, the output will have the needed dimensions. We have one output node with the sigmoid activation function in this situation. We'll utilize the functional API modeling system to develop our final model, which we'll then hand over to the user so they can use it to do any job using the U-Net architecture.

3.5 DEEPLABV3 ARCHITECTURE

The fundamental concept behind architecture of DEEPLABV3[7] is division of the system into two segments, or to components: The first segment is referred to as

'**Encoder**' and the second segment as '**Decoder**'. Unlike the earlier mentioned U-net architecture, DEEPLABV3 is in some sort a **hierarchical** system.
The architecture is pretty straight forward. The image is processed in following manner:

- Image enters the encoder, which consists of a dense/deep convolutional neural network(CNN) such as **Xception, ResNet, VGG**, etc. This backbone network performs the task of feature extraction such as 'edges' and 'corners'.
- Atrous convolution is employed in the last few blocks of the backbone to regulate the size of the feature map.
- An ASPP network is added to the retrieved characteristics from the backbone to categorize each pixel according to its classifications.
- To acquire the picture's actual size, which will be the final segmented mask for the image, the output from the ASPP network is sent through a 1 x 1 convolution.
- In the decoder section, to the low - level feature map obtained from the backbone network a 1 x 1 convolution is performed and output of this is concatenated with output of encoder which is upsampled by 4.
- The decoder then performs a 3 x 3 convolution followed by upsampling by 4 to obtain segmented image.

The two figures very well illustrate the working schema of DEEPLABV3 encoder architecture and DEEPLABV3 encoder-decoder architecture respectively.

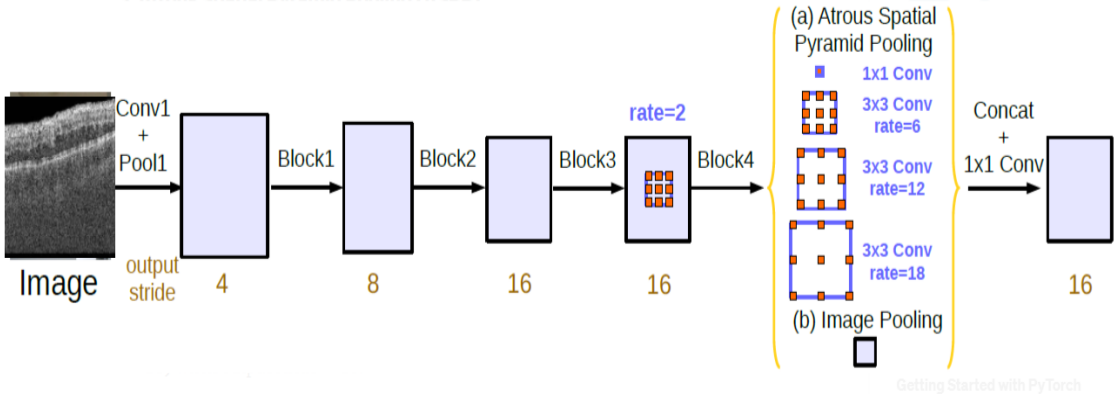

Figure 3.1: DEEPLABV3 encoder architecture

The Deep Convolution Neural Network specified in the DEEPLABV3 Architecture is **Modified Xception**.

The detailed components about Xception backbone is discussed in later sections.

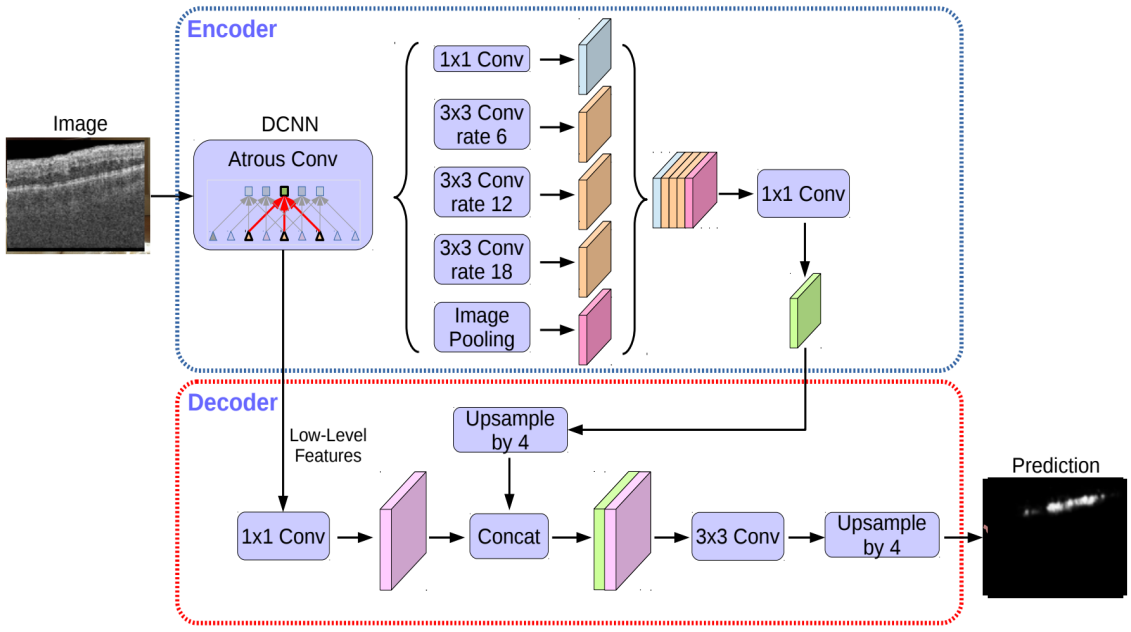

Figure 3.2: DEEPLABV3 encoder-decoder architecture

3.6 DATA PREPROCESSING METHODS

As discussed earlier, the **RETOUCH** dataset containing OCT images have to be pre-processed before training with through a convolutional neural network for segmentation.[1] Following preprocessing procedures were implemented:

3.6.1 GETTING IMAGES PATH

As mentioned earlier, a CSV file was created to obtain the names of the images. Through this CSV file, 2 arrays were initiated to derive image paths of the training data and lable data.

3.6.2 RESIZE AND NORMALIZE IMAGES

After obtaining the file paths to image data and label data, iteration over the arrays were performed to normalize the data between the scale of [0, 1] and dimensionality reduction to get a resized image of (128*128).

3.6.3 DATA AUGMENTATION

From an OCT image vendor perspective, the limited amount of data degrades learning of the model. Hence **data augmentation** was used to acquire additional images by performing 'random_shift' or 'flipping the axis' or even 'random slight rotation'. Each operation was performed by using the keras module which was then trained in real time.

3.7 DEEPLABV3 CONSTRUCTION

Construction of DEEPLABV3's architecture is pretty straight forward. In this section of the report, there will be a explanation regarding creating the **ASPP** module, which explicitly performs 'Atrous Spatial Pyramid Pooling' and creating **DEEPLABV3** module, which instantiate a DeepLabV3 model when it is called.
Within the DEEPLABV3 module, there is declaration of backbone network(Xception is this scenario) for feature map extraction.

3.7.1 ASPP MODULE

According to the architecture, the feature map obtained from backbone network, which now acts as an input tensor for Atrous Spatial Pyramidal Pooling(ASPP). The use of ASPP is justified since it has been discovered that as the sample rate increases, the number of valid filter weights decreases.
In ASPP, dialated convolution takes place. This means convolution on the feature map is carried out but defining a space between values in the kernel. The 'space' is described as 'dilation rate'.
Following are the processes that happen parallelly in **ASPP** function:

- A 1x1 convolution
- 3x3 convolution with dilation rate 6
- 3x3 convolution with dilation rate 12
- 3x3 convolution with dilation rate 18
- Image pooling

Output from these layers are concatenated followed by a 1x1 convolution which then enters the decoder component.

Algorithm 2 ASPP function

Input: 'feature_map' from **Xception** DCNN
Parallel processess:
- 1x1 Convolution

- 3x3 Atrous convolution at rate = 6

- 3x3 Atrous convolution at rate = 12

- 3x3 Atrous convolution at rate = 18

- Image Pooling

Image stack ← **Concatenate from each layer**

1x1 convolution ← *Imagestack*

3.7.2 Xception BACKBONE NETWORK

Xception is a 71-layer deep convolutional neural network. The flow of the CNN is categorized into 3 parts:

- Entry Flow
- Middle Flow
- Exit Flow

Entry Flow

In this flow, a series of normal convolution and depth separable convolution is carried out. Multiple skip connection are also added to obtain feature map much efficiently. In this modified architecture of **Xception**, the 'image pooling' layers are replaced with **depth_separable_convolution** layers.

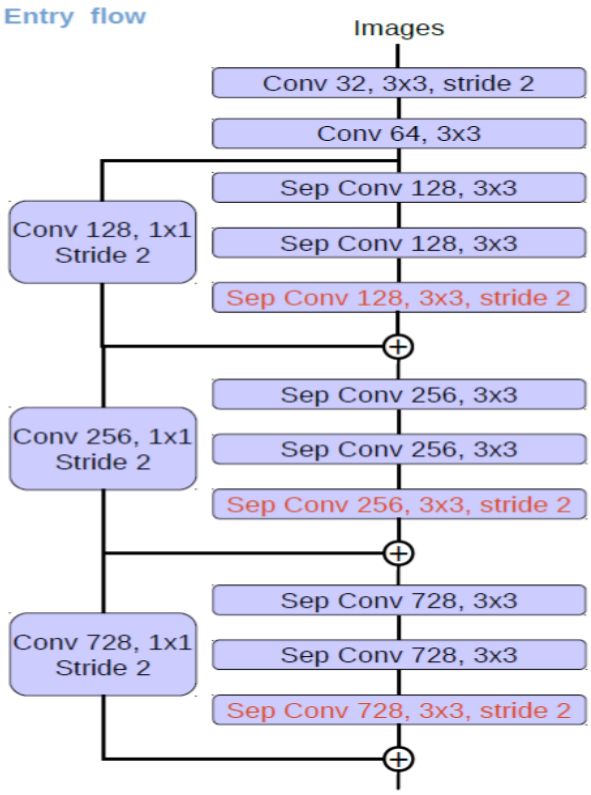

Figure 3.1: Xception Entry Flow

Middle Flow

In this flow, separable convolution is repeated 16 times to increase channels of the input feature map.

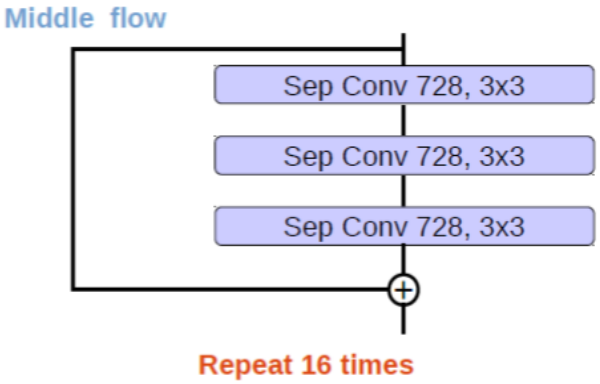

Figure 3.2: Xception Middle Flow

Exit Flow

In this flow, skip connection on residual block containing max pooling and separable convolution is implemented to increase the number of channels in feature map.

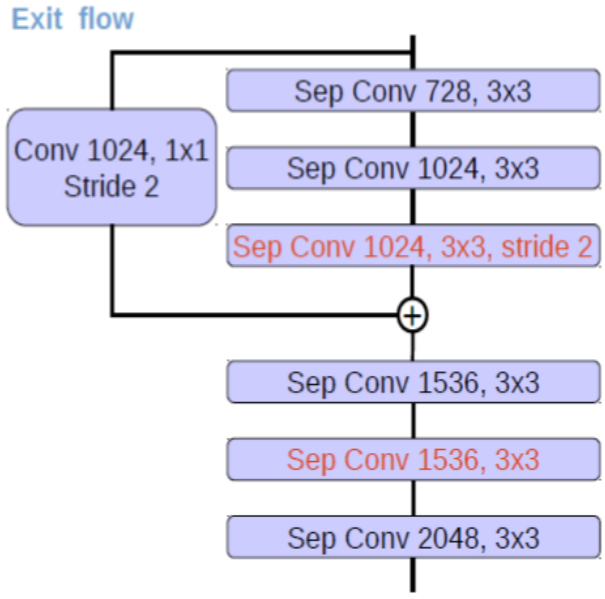

Figure 3.3: Xception End FLow

'Xception' backbone helps in extraction of feature that are low level and high level as well.

3.7.3 DEEPLABV3 MODULE

This module is basically a function which calls the 'Xception' function and 'ASPP' function to compute the encoder part of the system.[6]
The output image generated from this module is an image with retinal fluid being segmented.

Following section explains the working of DEEPLABV3 module as a pseudo code representation.

Algorithm 3 DEEPLABV3 function

Input: img_size, weights

if $weights \neq PASVOC$ **then**
    Raise Value error

else
    Call function **Xception**
Call function **ASPP**
Feed output of **Xception** into **ASPP**
1x1 Convolution$\leftarrow Output \leftarrow ASPP$
$1x1Convolution(Upsampled) + Decoder\_output = Segmented\_Image$

3.8 PERFORMANCE METRICS

Various scores and accuracy of the models and techniques chosen are mentioned below. To compare and evaluate our model, we use accuracy (A), precision (P), recall (R), and Dice-Coefficient (F1) as evaluation metrics. These can be combined in a confusion matrix.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.1}$$

$$Precision = \frac{TP}{TP + FP} \tag{3.2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.3}$$

$$Dice - Coefficient(F1) = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \tag{3.4}$$

where TP = True Positive,

TN = True Negative,

FP = False Positive,

FN = False Negative

True negatives (TN) and true positives (TP) represent the number of times the model predicts correctly. Similarly, false negatives (FN) and false positives (FP) represent the number of times the model predicts incorrectly. By computing the number of correct and incorrect predictions in such a matrix, we can observe how classification models are confused when making predictions. The goal of creating robust models is to keep the number of false positives and negatives low.

3.9 PROPOSED METHODOLOGY PIPELINE

For the previously mentioned DeepLabV3 methodology, the pipeline executed is:

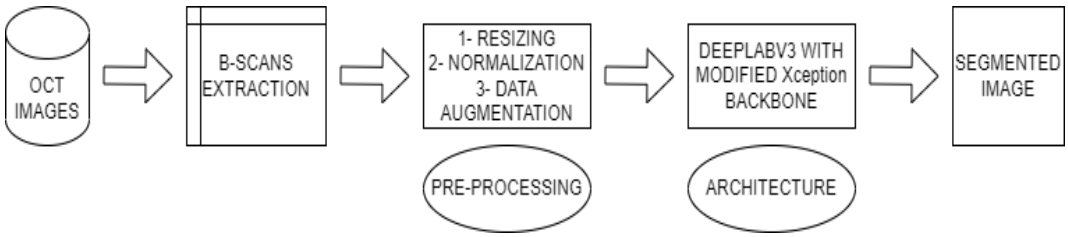

Figure 3.1: Pictorial representation of pipeline followed for semantic segmentation

Chapter 4

RESULT AND ANALYSIS

In the proposed model of U-Net is trained on 7000 samples, validated on around 700 samples(0.1 samples). It took few hours to run epochs on RETOUCH dataset using Convolution layers and to show the results.

4.1 TRAINED U-NET MODEL

In next step, we will compile and train the model to see its performance on the data. We are also using a checkpoint to save the model so that we can make predictions in the future. I interrupted the training procedure after 100 epochs as I was quite satisfied with the result obtained. You can choose to run it for more epochs.

The Above U-Net based model was trained over the 3 cyst vendors datsets like Cirrus, Topcon, Spectralis and given are their respected Performance metrics obtained from model are shown below in the figure.

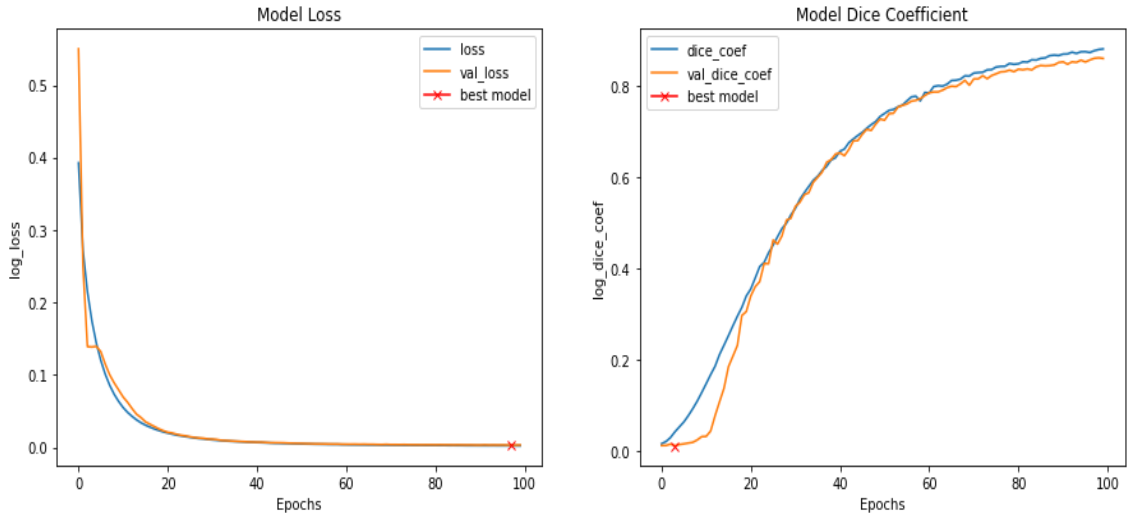

Figure 4.1: U-Net Model Loss and Dice Coefficient

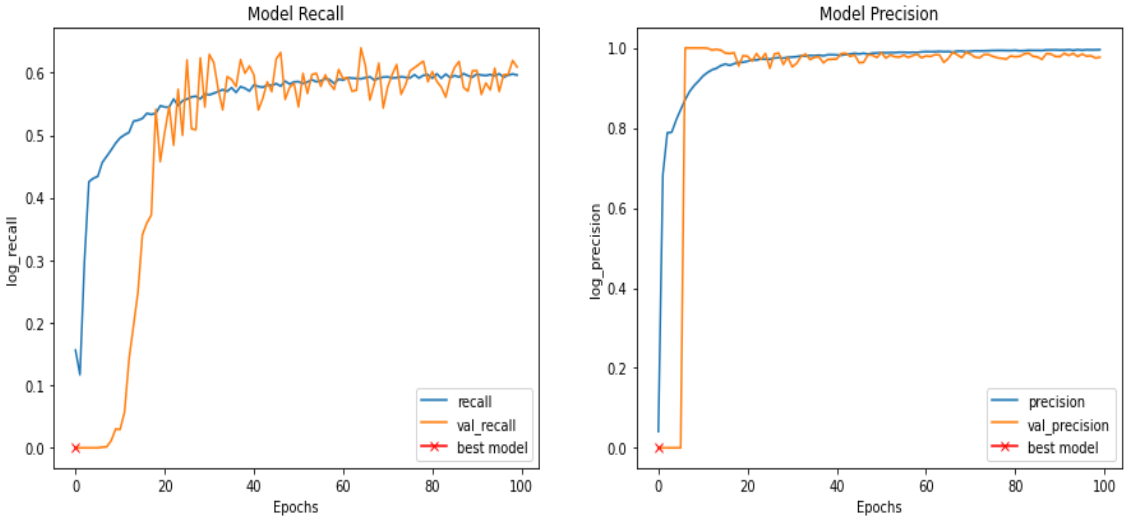

Figure 4.2: U-Net Model Recall and Precision

Following is the table which specifies loss, precision, recall and F1-score calculated over each OCT vendor.

.0

Cyst Vendors	Loss	Precision(%)	Recall(%)	F1-score(%)
Cirrus	0.003	98	70	86
Topcon	0.006	96	73	83
Spectralis	0.012	95	67	81

Table 4: Different Cyst Vendors Performance metrics

Also, the predicted segmentation on the test set is compared with actual segmented image set.

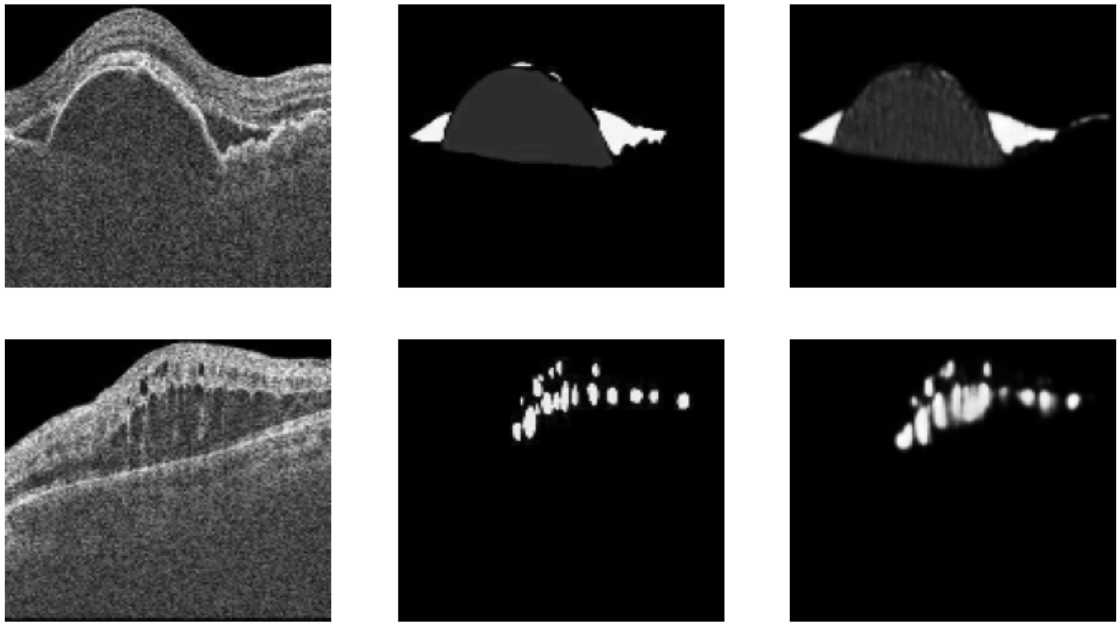

Figure 4.3: Input Image : Target segmentation : Predicted segmentation

4.1.1 TRAINED DEEPLABV3 MODEL

In the proposed model, the DeepLabV3 model is trained, tested and validated over the following size of dataset and parameters:

- Training set = 2700 images
- Validation set = 600 images
- Testing set = 200 images
- Optimizer = keras.optimizers.Adam
- Learning_rate = 0.001
- Epochs = 100

According to the **RETOUCH** data, OCT images are acquired from 3 vendors:

- Cirrus
- Spectralis

- Topcon

These parameters are used to train the model on each OCT image vendors separately. The model is then used to predict fluid segmentation(**IRF, SRF, PED**) on testing OCT images.

For training the model, **Early stopping was used** at patience value of '10' to ensure that in each epoch the log_loss value is decreased. If such stagnant nature is observed from the model then system stops training model and saves the corresponding 'weights' obtained for segmentation.

Following are the graphs which specify **model loss**, **F1-score**, **Recall** and **Precision** over each epoch during training the model.

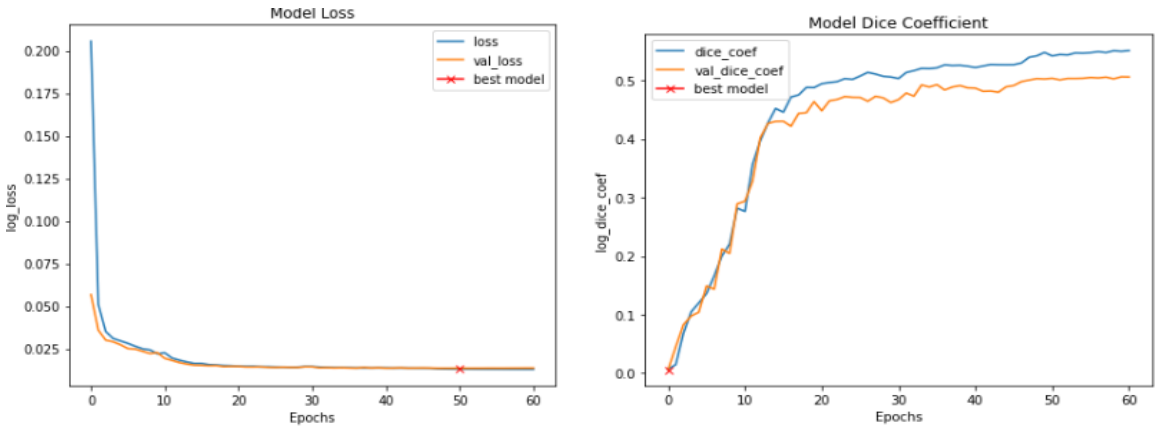

Figure 4.1: DeepLabV3 Model Loss and Dice Coefficient

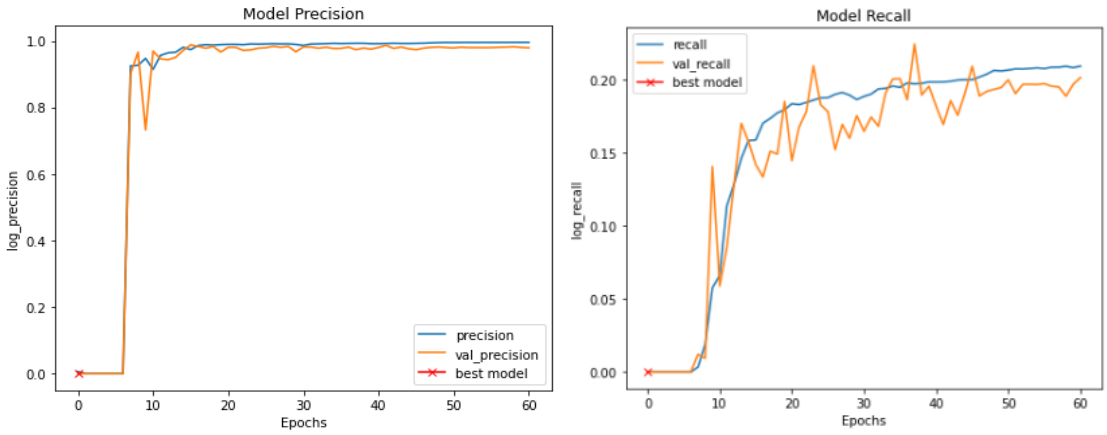

Figure 4.2: DeepLabV3 Model Precision and Recall

This table describes the loss, precision, recall and F1-score obtained on segmentation of IRF, SRF and PED on the various OCT scans obtained from different machines/vendors.

Cyst Vendors	Loss	Precision(%)	Recall(%)	F1-score(%)
Cirrus	0.004	94	65	77
Topcon	0.005	95	72	82
Spectralis	0.032	88	74	80

Table 4: Different OCT Vendors Performance metrics

These images are some predictions made by the DeepLabV3 model.

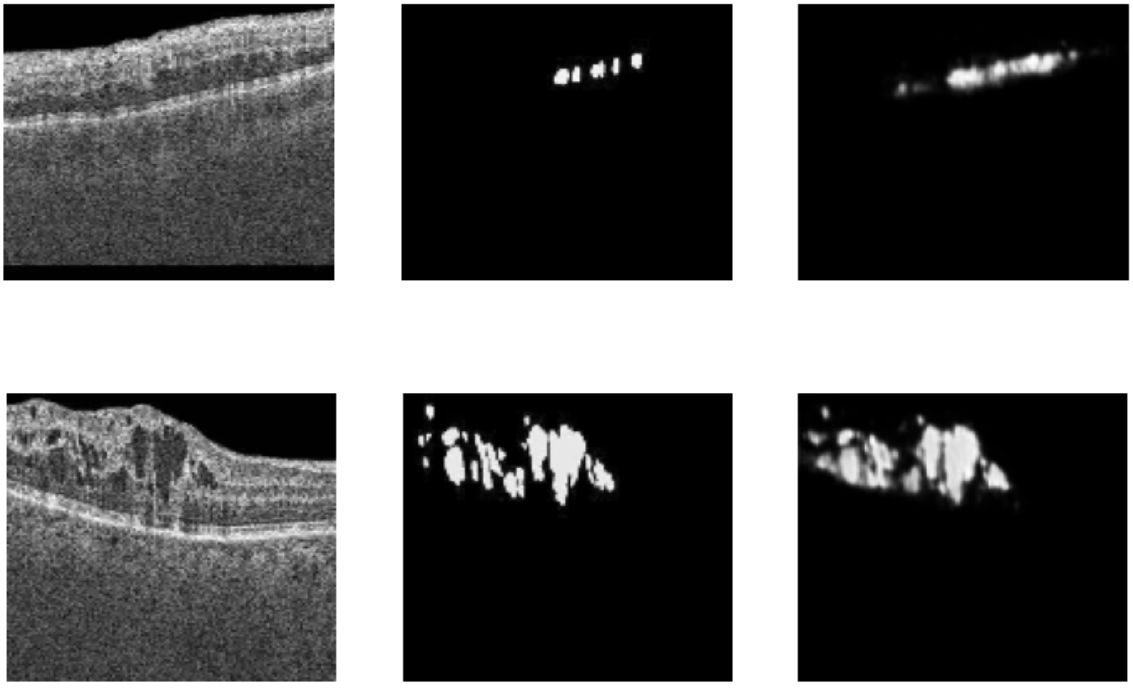

Figure 4.3: Ground truth : Target segmentation : Predicted segmentation

4.2 COMPARATIVE ANALYSIS

In the above sections of the report, we discussed generic pipeline of the proposed systems, architecture of the DEEPLABV3 and U-Net deep learning model and evaluated their learning on basis of various performance metrics such as dice coefficient, precision,etc.

But, there should be a comparative analysis between the proposed model and the base-line model to acquire knowledge about which model accurately segments the IRF, SRF and PED fluid in OCT scans, and hence in real life help doctors to efficiently operate on Macular Edema disease.

The base-line model used to compare DEEPLABV3 proposed model with is explained in detail in the citation[5]. The table describes a structured comparison between proposed DEEPLABV3 model and base-line model with respect to many aspects such as **pre-processing algorithm, deep learning architecture used, fluid to be segmented**, etc.

Pipeline	**Pre-processing**	**Model**
Base-Line	Data Aug, Layer Seg	U-Net
Proposed	Resize, Data Aug, normalize	DeepLabV3

Table 4: Pre-processing and model used in base-line and proposed system

Pipeline	**Post-processing**	**Metric**	**Result**
Base-Line	Random Forest	F1-Score	0.82(Cir), 0.75(Top), 0.74(Spec)
Proposed	-	Precision, F1-Score, Recall	0.77(Cir), 0.82(Top), 0.80(Spec)

Table 4: Post-processing and comparison of F1-score for each vendor used in base-line and proposed system

On comparing the proposed DEEPLABV3 model with base-line model(which uses U-Net), the F1-score is the common metric used for evaluating performance of each

pipeline. It is visible that proposed system provides lower F1-score than base-line model on dataset from Cirrus cause the majority of disease indicated by OCT image in Cirrus is Age-related Macular Edema.

Hence, the proposed model performs effectively than base-line if the disease mentioned in OCT images are Diabetic Macular Edema.

Chapter 5

CHALLENGES FACED AND IMPROVEMENTS

5.1 HIGH VARIABILITY OF RETINAL FLUID

The main problem when segmenting severely diseased scans is the high variability in the shape and location of fluid filled regions. Retinal fluid can form in any layer of the retina, but it can also span many layers or be layered on top of other fluid-filled areas. This makes modelling all of the possible interactions between fluids and retinal layers challenging.

5.2 TIME REQUIRED FOR PREPROCESSING

We collected a variety of datasets but each dataset had a different format and combining them required much time and effort. Furthermore, it was not feasible to train the models on raw data due to the sheer size of the datasets. To deal with this, we used image libraries to extract compact meaningful information from the data.

5.3 LACK OF STANDARDIZED MACHINES

The absence of a consistent acquisition technique between multiple OCT device is a fundamental problem for AI-based OCT Image analysis. Images of various sizes, contrast levels, and textures are created without standardization, and thus are not generalizable to a single Model. Image quality varies depending on capture periods and signal-to-noise ratios between devices.

5.4 LIMITATION OF HARDWARE RESOURCES

Another major difficulty is computational restrictions. Deep learning is plagued by GPUs with insufficient dynamic random access memory (DRAM). Inadequate pro-

cessing capacity can limit minibatch size, convolution depth, or number of epochs required to train the model.

5.5 REQUIREMENT OF LARGE DATSETS

The need on huge volumes of manually labelled ground truth data in order to effectively understand the considerable variability seen in OCT images with severe macular edoema is a constraint of our machine-learning technique.

Chapter 6

CONCLUSION AND FUTURE WORK

CNN, FCN, U-Net, and hybrid approaches were used to create the vast majority of deep learning architectures. The retinal fluid segmentation findings may be influenced by the fact that OCT pictures in the same dataset may come from multiple suppliers. As a result, future research should concentrate on domain generalisations in order to minimise segmentation's dataset reliance. The sophisticated deep learning algorithms provide a number of advantages for the reliable and effective identification of retinal fluid lesions, which can help with ocular illness diagnosis decision-making in a variety of clinical settings. Furthermore, automated retinal fluid segmentation might minimise medical specialists' burden by removing the subjective impact of manual detection.

In the future, with the advancement of sophisticated deep learning technology, retinal fluid segmentation methods based on deep learning architectures will gradually replace the manual segmentation process in clinical practise. On three separate datasets related with different types of illnesses, the suggested technique performed well, with a DSC of above 80 %. The segmentation method performed badly on several volumes because to the small amount of training samples in the available data sets. We predict higher fluid segmentation and detection accuracy as more data is collected in the future.

Bibliography

[1] Khaled Alsaih, Mohd Zuki Yusoff, Tong Boon Tang, Ibrahima Faye, and Fabrice Mériaudeau. Deep learning architectures analysis for age-related macular degeneration segmentation on optical coherence tomography scans. *Computer methods and programs in biomedicine*, 195:105566, 2020.

[2] Hrvoje Bogunović, Freerk Venhuizen, Sophie Klimscha, Stefanos Apostolopoulos, Alireza Bab-Hadiashar, Ulas Bagci, Mirza Faisal Beg, Loza Bekalo, Qiang Chen, Carlos Ciller, et al. Retouch: the retinal oct fluid detection and segmentation benchmark and challenge. *IEEE transactions on medical imaging*, 38(8):1858–1874, 2019.

[3] Hanzhi Chen, Yun-Fei Hsu, Zeju Qiu, and Co-Mentor M Sc Laure Vuaille. Domain adaptation for annotation-efficient image segmentation. 2021.

[4] Sung Ho Kang, Hyoung Suk Park, Jaeseong Jang, and Kiwan Jeon. Deep neural networks for the detection and segmentation of the retinal fluid in oct images. *MICCAI Retinal OCT Fluid Challenge (RETOUCH)*, 2017.

[5] Donghuan Lu, Morgan Heisler, Sieun Lee, Gavin Ding, Marinko V Sarunic, and Mirza Faisal Beg. Retinal fluid segmentation and detection in optical coherence tomography images using fully convolutional neural network. *arXiv preprint arXiv:1710.04778*, 2017.

[6] Donghuan Lu, Morgan Heisler, Sieun Lee, Gavin Weiguang Ding, Eduardo Navajas, Marinko V Sarunic, and Mirza Faisal Beg. Deep-learning based multiclass retinal fluid segmentation and detection in optical coherence tomography images using a fully convolutional neural network. *Medical image analysis*, 54:100–110, 2019.

[7] Gabriella Moraes, Dun Jack Fu, Marc Wilson, Hagar Khalid, Siegfried K Wagner, Edward Korot, Daniel Ferraz, Livia Faes, Christopher J Kelly, Terry Spitz, et al. Quantitative analysis of oct for neovascular age-related macular degeneration using deep learning. *Ophthalmology*, 128(5):693–705, 2021.

[8] Abdolreza Rashno, Dara D Koozekanani, and Keshab K Parhi. Detection and segmentation of various types of fluids with graph shortest path and deep learning

approaches. In *Proc. MICCAI Retinal OCT Fluid Challenge (RETOUCH)*, pages 54–62, 2017.

[9] Ruwan Tennakoon, Amirali Khodadadian Gostar, Reza Hoseinnezhad, and Alireza Bab-Hadiashar. Retinal fluid segmentation and classification in oct images using adversarial loss based cnn. *Proc. MICCAI Retinal OCT Fluid Challenge (RETOUCH)*, pages 30–37, 2017.

[10] Freerk G Venhuizena, M Grinsvena, Carel B Hoyngb, Thomas Theelenb, BV Ginnekena, and Clara I Sancheza. Vendor independent cyst segmentation in retinal sd-oct volumes using a combination of multiple scale convolutional neural networks. *Medical Image Computing and Computer Assisted Intervention-Challenge on Retinal Cyst Segmentation*, 2015.

Appendix A

BIO-DATA

Name: Akshay Dhayal
Address: National Institute of Technology Karnataka, Surathkal
Email: akshaydhayal.181co105@nitk.edu.in
Contact No.: +91-9980134333

Name: Chemat Wangail
Address: National Institute of Technology Karnataka, Surathkal
Email: chemat.181co114@nitk.edu.in
Contact No.: +91-9103208603

Name: Saswat Kumar Nayak
Address: National Institute of Technology Karnataka, Surathkal
Email: saswatkumar.181co147@nitk.edu.in
Contact No.: +91-6303994301