# Chapter 1

# Introduction

One of the common types of dementia is Alzheimer’s disease (AD), in which aging is one of the most critical factors. Studies show that 5.8 million people in the US have AD, which means approximately 10% of the population older than 65 years old are suffering from this disease. Currently, invasive medical procedures such as Position Emission Tomography (PET) and Magnetic Resonance Imaging (MRI) are employed for AD diagnosis. On the other hand, recent studies suggested that AD can be diagnosed and monitored by analysing neuronal loss from the retina, which is related to the decrease of the retinal nerve fiber layer thickness. Optical Coherence Tomography (OCT) is a non-invasive tool that produces cross-sectional images of the retinal structure. Therefore, it is hypothesized that OCT imaging providing the structure of the retinal layers that can be a useful biomarker in AD patients to evaluate and follow these patients in combination with the next generation of automatic machine learning allowing a detailed and yet efficient approach to quantify and identify the AD disease.

During the past decade, different CNN models e.g. LeNet, GoogleNet and DenseNet are utilized in different problems such as image segmentation, image recognition, speech recognition. The increasing demand in real-life applications for developing new architectures with the highest efficiency and accuracy, which is very challenging and has been subjective (man-made architectures) given infinite possibilities.

Moreover, finding the best hyper-parameters is significantly costly using human experts; therefore, there are many research efforts in the area of hyper-parameter tuning such as random search and Bayesian optimization. However, it is still very challenging to find the best hyper-parameter set and model architecture for the predictive task at hand.

Recently, Neural Architecture Search (NAS), which is classified as a subfield of Auto ML, plays a significant role in improving the network architecture to achieve higher accuracy without manual trial and error. NAS process has been categorized into three steps: search space, search strategy, and performance estimation strategy. At the search strategy stage, NAS uses a recurrent neural network to sample different child architectures and train candidate architecture convergence to obtain their accuracies on the validation set. This accuracy is used (as a reward or punishment) to update the controller parameters; hence the controller will generate better architectures with higher accuracy over time.

## Literature Survey

Segmentation of the retina’s boundary layers in OCT images is a challenging task since it is a complex image that presents a large amount of noise and deformations caused by pathologies (drusen, drusenoid PED and geographic atrophy) that change the characteristics of the edges of the layers. Several studies have been published demonstrating methodologies for segmenting the edges of retinal layers.

In a method is presented for segmentation of the total retina, which lies between the internal limiting membrane (ILM) and the retinal pigment epithelium (RPE), and the RPEDC, which lies between the RPE and the Bruch’s membrane. To validate the method, 220 B-scans extracted from 20 volumes are used. All patients had age-related macular degeneration (AMD) in the intermediate stage and with druses, and some with geographic atrophy. The method is based on graph cutting. As a pre-processing step, noise reduction is performed through a rectangular mean filter. From the resulting image, a graph is constructed with the pixel values as weights. The search for each edge is performed sequentially. Considering all layers, the algorithm showed an average difference of 0.95 pixels, a value even smaller than the difference in segmentation between two specialists. Together with the one developed by these works refer to the development of the DOCTRAP software (Duke OCT Retinal Analysis Program) designed for use in OCT segmentation research.

In a framework is presented that combines convolutional neural networks (CNN) with graph search algorithms to segment nine retinal edges in OCT images. The method was validated with 60 volumes (2915 B-scans) of twenty eyes from people with non-exudative AMD. The CNN is trained with the characteristics of the layers’ edges to estimate the positioning of the edges of the eight layers. Then, these values are put through a graph search algorithm for the final definition of the boundaries. The results found were compared to the results obtained by the DOCTRAP and OCT Explorer segmentation software. These results were superior to those obtained by the OCT Explorer software but still inferior to those obtained by the DOCTRAP software.

Presents a method for segmenting the edges of retinal layers in OCT images. Three public databases were used, two with a total of 210 B-scans of patients without a pathology and the other with 220 B-scans of patients with intermediate AMD. Two models based on the combination of Convolution Neural Network (CNN) and Long Short Term Memory (LSTM) are created, 1 to segment eight retina layers and another to segment three layers. CNN is used to extract a region of interest and detect the layers’ edges, while LSTM is used to track the layer boundary. This model is trained with a mix of normal and AMD cases. The results show an average absolute error (EMA) in pixels of 1.30±0.48 μm, less than the error of marking the bases of 1.79 ± 0.76 μm.

In a new linearly parameterized conditional random field model (LP-CRF) is proposed to segment the retinal layers’ edges OCT images. Two public databases were used, one with 107 B-scans from patients without pathology and the other with 220 B-scans from patients with intermediate AMD. The proposed LP-CRF comprises two convolution layers to capture each region’s appearance and layer boundary, the relative weights of the previous form, and an additional term based on the similarity of appearance of the adjacent boundary points. All types of energy are learned in conjunction with a Support Vector Machine. The proposed method segments all the retina’s limits in a single step, and for the images without pathology, eight edges are extracted and with AMD three edges. The mean absolute error reached was 1.52±0.29 μm pixels for the segmentation of 8 boundaries in the Normal data set and 1.9±0.65 μm pixels for three edges in the combined AMD and Normal data set.

Proposed a method for segmenting the edges of retinal layers in OCT images. They used two image databases, one for pediatric patients without a history of pathologies and intermediate AMD patients and pathologies. The method used a trained recurrent neural network (RNN) as a fragment-based classifier to segment seven layers boundaries in the first base and three edges in the second base. The results indicated that the RNN architecture is a viable alternative to CNN for image classification tasks that exhibit a clear sequential structure. Compared to CNN, RNN showed a slightly smaller absolute mean error.

A method for segmenting the edges of the ILM, RPE, and BM layers is presented in The method was tested on two public databases with normal patients and with intermediate AMD, one provided by [[11](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0251591#pone.0251591.ref016)] with 384 volumes, the other account with 19 volumes made available by [[1](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0251591#pone.0251591.ref007)]. The deep neural network architecture, capsule network, was used to do the initial segmentation of the edges of three layers of the retina, where it classifies fragments extracted from OCT images in the four classes (ILM, RPE, BM, and fundus). To refine the segmentation, the author made use of the graph cut technique, which is applied to the neural network probability maps. As a result, the method reached, respectively, for the two bases, 0.75 and 2.04 average of the mean absolute error for the ILM, 0.93 and 1.97 for the RPE, 1.09 and 2.05 for BM.

In a new method is presented to segment ILM, RPE, and BM in OCT images. Three image bases were used to train and validate the method, with healthy patients’ images and intermediate AMD. The authors use an ensemble method called deep Forest, which uses several Random Forests to generate the classification of patches extracted from OCT images. As pre-processing, normalization of pixel intensities is applied. The patches are divided into four classes, internal limiting membrane, retinal pigment epithelium and druse complex (RPEDC), Bruch’s membrane, and fundus. After classification, a graph-based technique developed by the author is applied to refine segmentation. As a result, the method achieved average errors of 0.81, 1.35, and 1.23 pixels for the three bases, respectively.

Developed a method for segmentation of retinal layers (ILM, RPE, and BM) in OCT images based on the wave algorithm, a mathematical model of the equation of the potential fluid energy in fluid mechanics. The method uses the base provided by and uses the absolute mean error to evaluate the results. The method obtained an EMA of less than 1.5 pixels in all evaluations.

## Motivation

Researchers trying to provide medical diagnosis or aid using AI techniques are constantly bombarded with new datasets for the same problem. This issue greatly hinders their generalizability, hence the team decided to find the best hyperparameter configuration for Optical coherence tomography (OCT) and diabetic macular edema (DME) dataset in general and also analyze the magnitude of change in hyperparameters from one dataset to another of the same purpose. This will help other researchers optimize and train their models efficiently.

## Problem Statement

OCT Image Segmentation using U-Nets.

## Objectives of project

To achieve the best model suited for implementing a deep learning model which classifies the images of Optical coherence tomography (OCT) and diabetic macular edema (DME) dataset and tuning hyperparameters to attain better results.

## 

## Chapter 2

## Requirement Analysis

## System Model

Diagram

Description generated with very high confidence

**Figure 1: System Model**

The input to the network structure is a set of OCT slice images corresponding to the retinal fundus image. The green line in this image indicates the scanning position of the corresponding OCT slice with a total of 19 scanning lines. The given set of a patient's OCT images is inputted into the trained network; the network outputs the results after classifying each pixel.

## Functional Requirements

* User shall be able to segment the images of Optical coherence tomography (OCT) and diabetic macular edema (DME) Dataset using a Unet model.
* Users shall be able to train only select layers using transfer learning.
* Users shall be able to optimize the model through any selected hyperparameter.

## Non-functional Requirements

* Users should segment the images with 80% of accuracy.

## Software and Hardware Requirements Specification

* The system must have an I3 processor and above, with a minimum of 4GB RAM.
* The operating system either Windows 7 and above or Ubuntu 16.04 and above

# Chapter 3

# System Design

## 3.1. Overview of SRS

A software requirements specification is the basis for your entire project. It lays the framework that every team involved in development will follow.

Software requirements specification establishes the basis for an agreement between customers and contractors or suppliers on how the software product should function (in a market-driven project, these roles may be played by the marketing and development divisions). Software requirements specification is a rigorous assessment of requirements before the more specific system design stages, and its goal is to reduce later redesign. It should also provide a realistic basis for estimating product costs, risks, and schedules. Used appropriately, software requirements specifications can help prevent software project failure

The software requirements specification document lists sufficient and necessary requirements for the project development. To derive the requirements, the developer needs to have a clear and thorough understanding of the products under development. This is achieved through detailed and continuous communications with the project team and customer throughout the software development process.

## Diagram Description generated with very high confidence3.2. system architecture

**Figure 2: System Architecture**

In the proposed system the team has considered a pipe and filter architecture. Pipe and Filter is a simple architectural style that connects a number of components that process a stream of data, each connected to the next component in the processing pipeline via a Pipe.

The pipe and filter architecture consist of one or more data sources. The data source is connected to data filters via pipes. The above figure shows the Architecture of system. The unet model is trained on OCT and DME dataset. The dataset images are preprocessed and input to the unet model for training. After the training the images are segmented.

## Data design

A data-flow diagram (DFD) is a way of representing a flow of a data of a process or a system. The data-flow diagram is part of the structured-analysis modelling tool. The DFD also provides information about the outputs and inputs of each entity and the process itself. A data-flow diagram has no control flow, there are no decision rules and no loops.

### Level 0 dataflow diagram

Diagram

Description generated with very high confidence

**Figure 3: Level 0 DFD**

### Detailed DFD

Diagram

Description generated with very high confidence

**Figure 4: Detailed DFD**

# Chapter 4

# Implementation

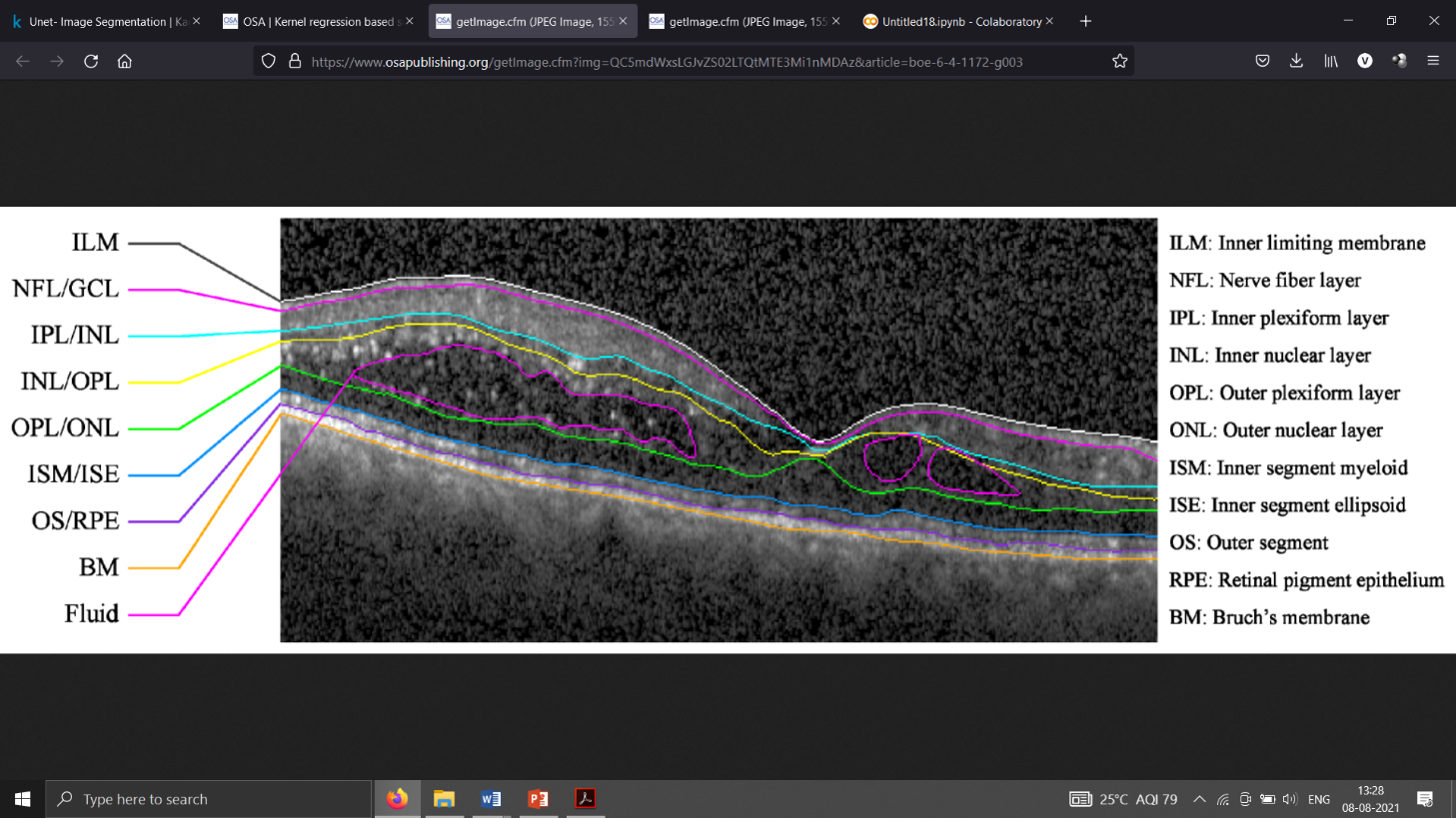
## 4.1. dataset Description

### 4.1.1. DME training data set

To learn our DME classifier, we obtained training data separate from our validation data set. We used the Duke Enterprise Data Unified Content Explorer search engine to retrospectively identify patients within the Duke Eye Center Medical Retina practice with a billing code for DME (ICD-9 362.07) associated with their visit. An ophthalmologist then identified six patients imaged in clinic using the standard Spectralis (Heidelberg Engineering, Heidelberg, Germany) 61-line volume scan protocol with severe DME pathology and varying image quality. Averaging of the B-scans was determined by the photographer, and ranged from 9 to 21 raw images per averaged B-scan. The volumetric scans were 𝑄=61 B-scans × 𝑁=768 A-scans with an axial resolution 𝑟𝑖=3.87 µm/pixel, lateral resolution (𝑟𝑗) ranging from 11.07 – 11.59 µm/pixel, and azimuthal resolution (𝑟𝑘) ranging from 118 – 128 µm/pixel.

### 4.1.2. Target classes

To generate the target classes for classifier training, they manually segmented fluid-filled regions and semi-automatically segmented all eight retinal layer boundaries following the definitions in Figure 5. This was done for 12 B-scans within the training data set (two from each volume). The B-scans selected consisted of six images near the fovea (B-scan 31 for all volumes) and six peripheral images (B-scans 1, 6, 11, 16, 21, and 26, one for each of the six volumes). Then used the manual segmentations to assign the true class for each pixel, with a total of eight possible classes defined in [Table 1](https://www.osapublishing.org/boe/fulltext.cfm?uri=boe-6-4-1172&id=312754#t001).



**Figure 5: Target retinal layer boundaries and fluid to segment on an SD-OCT**

**A screenshot of a computer

Description generated with high confidence B-scan of an eye with DME**

**Table 1: DME classes**

**4.1.3. DME validation data set**

Validation data set is obtained by identifying ten patients with DME that were not included in the training data set. The method for selecting these data sets is described in Section 5.1, with the difference that the images had to be of adequate quality (i.e. layer and fluid boundaries needed to be visible). The image acquisition parameters were consistent with the training data set, and lateral and azimuthal resolutions ranged from 10.94 – 11.98 µm/pixel and 118 – 128 µm/pixel, respectively. We made the entire data set available online, including the training and validation data sets and their corresponding automatic and manual segmentation results.

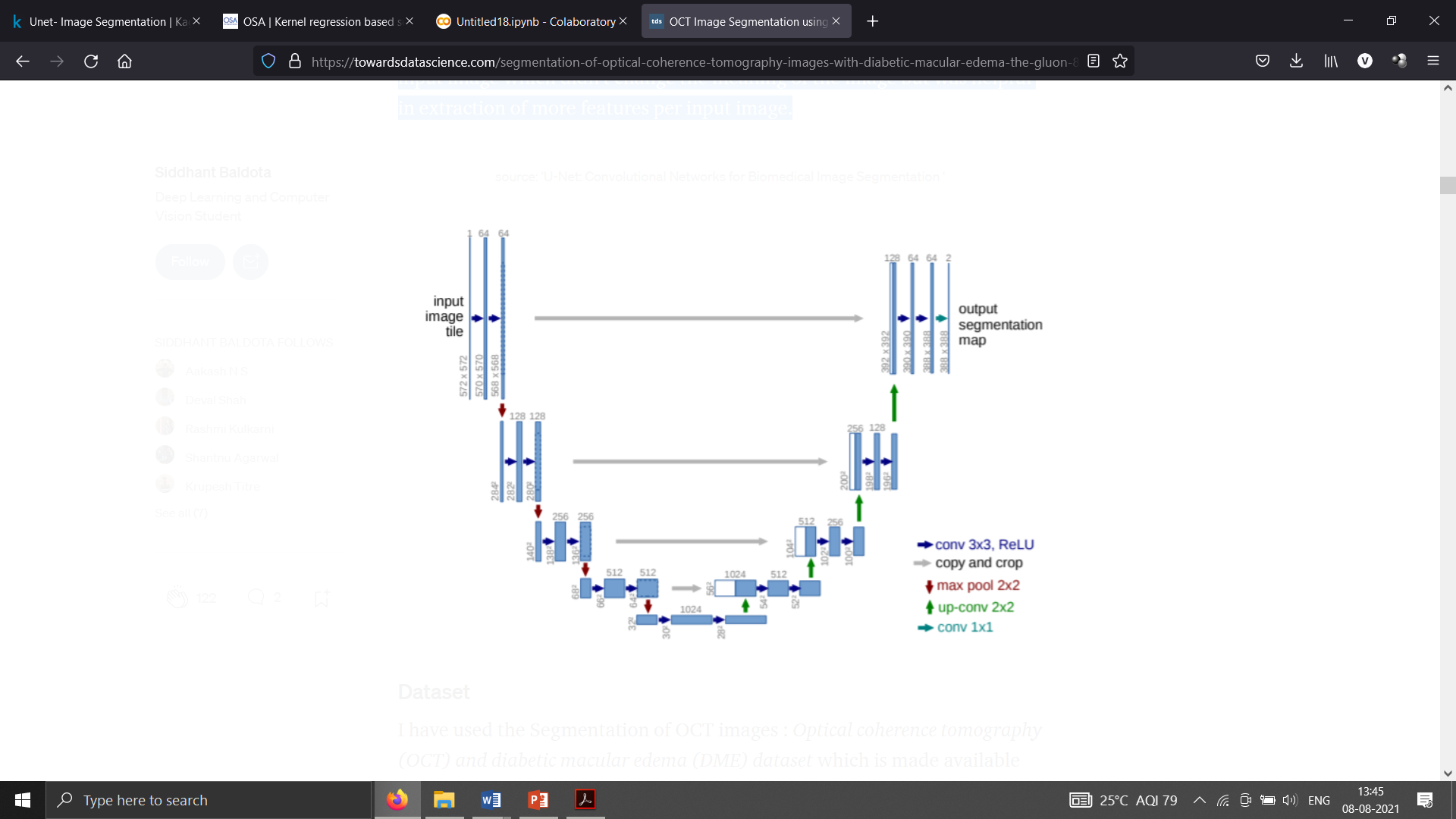
## 4.2. image pre-processing

Prior to training the classifier, we perform image pre-processing steps for all B-scans within a volume. We first remove Spectral is registration boundaries by replacing them with the mirror image of the valid image regions. To standardize the images, we resize the image to a lateral and axial resolution of 𝑟𝑗=13.4µm/pixel and 𝑟𝑖=6.7µm/pixel, respectively, and then linearly normalize the image to range from 0 to 1.

## 4.3. Unet

On the 18th of May, 2015, Olaf Ronneberger, Philipp Fischer, and Thomas Brox released a paper on Convoluted Neural Networks (CNNs) for Biomedical Image Segmentation. Traditionally CNNs were used for image classification and object detection. This approach wasn’t suitable for though as the desired output was that of localization where a class label was to be assigned to each pixel. Since biomedical tasks didn’t have a huge amount of positive samples in comparison to a image classification problem or an object detection problem. Ciresan et al. attempted training a network, using the sliding window setup to predict the class label of each pixel by using a region of interest or patch around that pixel as the input features. This method was though novel was quite sedate and resulted in high amounts of redundancies because the patches overlapped. Moreover there was a trade-off between patch size and localization accuracy. Greater the patch-size, higher was the context but this resulted in the increase of max pooling layers which diminished the spatial information resulting in lower localization accuracy. If the patch-size was reduced, then the context reduced. This resulted in the creation of U-Nets, which are pseudo-fully connected networks.

The primary proposition was a providing a normal convolving network with consecutive layers sequentially. The max-pooling layers were substituted by up sampling layers, resulting in increase of pixel density and ultimately, resolution. These higher resolution features from the convolving path were concatenated with the up sampled output. Successive convolutional layers were added as the expanding path which resulted in an increase in the number of feature channels to enable the network pass on spatial information to the high resolution layers. This model doesn’t contain any fully connected layers and uses the most important part of each feature extractor (convolution). The overall structure was U-shaped and hence the name U-net. In case of larger images, the missing context was produced by flipping the input image and by performing different augmentations on the input image which didn’t change the meaning of the image but was helpful in extraction of more features per input image.



**Figure 6: U-Net: Convolutional Networks for Biomedical Image Segmentation**

The lowest resolution was taken as 32x32 pixels. Each sequential block of convolutional layers was max pooled and stacked on each other. To form the base of the U-Net, a plateau block consisting of 3 convolutional layers was created for data consistency which had the equal number of input and output channels. They were later up-sampled by transposing the convolutional layers. In turn, these layers were concatenated with the corresponding depth convolutional block. The output of the concatenation was passed on to the ascending limb of the U-Net or the expansive blocks. At each depth level, spatial information was continuously passed using concatenation.

**4.4 Source code:**

**4.4.1 Importing Required Libraries:**



**4.4.2 Pre-processing dataset:**

**4.4.3 Training Model using Unet architecture:**



# Chapter 5

# Results and conclusion

**5.1.Results**

The network performance of the results of training and test datasets is evaluated. For the training dataset, the segmentation effect and network convergence are observed according to the accuracy rate of the dataset and loss trend during the training process. For the test dataset, the segmentation accuracy, F1 score, and Kappa coefficient are computed according to the prediction results of this dataset in order to analyze the network generalization ability. The definitions of each evaluation metric are as follows:

1. Accuracy: This reflects the network ability to determine the entire dataset. The overall accuracy and the accuracy of each class are analyzed.
2. F1 score: The harmonic mean of precision and recall, also known as the Dice coefficient.
3. Kappa coefficient, Kappa is a statistic that measures inter-rater agreement for qualitative (categorical) items. It is generally assumed to be a more robust measure than a simple percentage agreement calculation because it considers the possibility of the agreement occurring by chance.

Here are some results I obtained after using my trained model.

Graphical user interface

Description generated with very high confidence

**Figure 7: Results plotted in colab**

## 5.2. conclusion

We proposed a boundary aware U-Net for retinal layers segmentation by detecting accurate boundary. Based on encoder-decoder architecture, we design a dual tasks framework with low-level outputs for boundary detection and high-level outputs for layers segmentation. Specifically, we first use the multi-scale input strategy to enrich the spatial information in the deep features of encoder. For low-level features from encoder, we design an edge aware (EA) module in skip connection to extract the pure edge features. Then, a U-structure feature enhanced (UFE) module is designed in all skip connections to enlarge the features receptive fields from the encoder. Besides, a canny edge fusion (CEF) module is introduced to architecture, which can fuse the priory edge information from segmentation task to boundary detection branch for a better predication.

U-Nets give better results than the sliding window segmentation model any day. They also work on minimal data.

## References

**[1].** Hitzenberger CK, Trost P, Lo PW, Zhou Q. Three-dimensional imaging of the human

retina by high-speed optical coherence tomography. Opt Express. 2003;11(21):2753–

2761. [[PubMed](https://www.ncbi.nlm.nih.gov/pubmed/19471390)] [[Google Scholar](https://scholar.google.com/scholar_lookup?journal=Opt+Express&title=Three-dimensional+imaging+of+the+human+retina+by+high-speed+optical+coherence+tomography&author=CK+Hitzenberger&author=P+Trost&author=PW+Lo&author=Q+Zhou&volume=11&issue=21&publication_year=2003&pages=2753-2761&pmid=19471390&)]

**[2].** Kashani AH, Keane PA, Dustin L, Walsh AC, Sadda SR. Quantitative subanalysis of

cystoid spaces and outer nuclear layer using optical coherence tomography in age-

related macular degeneration. Invest Ophthalmol Vis Sci. 2009; 50(7):3366–3373.

[[PMC free article](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2908154/)] [[PubMed](https://www.ncbi.nlm.nih.gov/pubmed/19168893)] [[Google Scholar](https://scholar.google.com/scholar_lookup?journal=Invest+Ophthalmol+Vis+Sci&title=Quantitative+subanalysis+of+cystoid+spaces+and+outer+nuclear+layer+using+optical+coherence+tomography+in+age-related+macular+degeneration&author=AH+Kashani&author=PA+Keane&author=L+Dustin&author=AC+Walsh&author=SR+Sadda&volume=50&issue=7&publication_year=2009&pages=3366-3373&pmid=19168893&)]

**[3].** Zheng YL, Sahni J, Campa C, Stangos AN, Raj A, Harding SP. Computerized

assessment of intraretinal and subretinal fluid regions in spectral-domain optical

coherence tomography images of the retina. Am J Ophthalmol.2013; 155(2):277–

286. [[PubMed](https://www.ncbi.nlm.nih.gov/pubmed/23111180)] [[Google Scholar](https://scholar.google.com/scholar_lookup?journal=Am+J+Ophthalmol&title=Computerized+assessment+of+intraretinal+and+subretinal+fluid+regions+in+spectral-domain+optical+coherence+tomography+images+of+the+retina&author=YL+Zheng&author=J+Sahni&author=C+Campa&author=AN+Stangos&author=A+Raj&volume=155&issue=2&publication_year=2013&pages=277-286&pmid=23111180&)]

**[4].** Chen Q, Leng T, Zheng LL, Kutzscher L, Ma J, de Sisternes L, Rubin DL.

Automated drusen segmentation and quantification in SD-OCT images. Med Image

Anal.2013; 17(8):1058–1072. [[PMC free article](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3795829/)] [[PubMed](https://www.ncbi.nlm.nih.gov/pubmed/23880375)] [[Google Scholar](https://scholar.google.com/scholar_lookup?journal=Med+Image+Anal&title=Automated+drusen+segmentation+and+quantification+in+SD-OCT+images&author=Q+Chen&author=T+Leng&author=LL+Zheng&author=L+Kutzscher&author=J+Ma&volume=17&issue=8&publication_year=2013&pages=1058-1072&pmid=23880375&)]

**[5].** Chen XJ, Niemeijer M, Zhang L, Lee K, Abramoff MD, Sonka M. Three-

Dimensional segmentation of fluid-associated abnormalities in retinal OCT:

probability constrained graph-search-graph-cut. IEEE Trans Med Imaging.

2012;31(8):1521–1531. [[PMC free article](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3659794/)] [[PubMed](https://www.ncbi.nlm.nih.gov/pubmed/22453610)] [[Google Scholar](https://scholar.google.com/scholar_lookup?journal=IEEE+Trans+Med+Imaging&title=Three-dimensional+segmentation+of+fluid-associated+abnormalities+in+retinal+OCT:+probability+constrained+graph-search-graph-cut&author=XJ+Chen&author=M+Niemeijer&author=L+Zhang&author=K+Lee&author=MD+Abramoff&volume=31&issue=8&publication_year=2012&pages=1521-1531&pmid=22453610&)]

**[6].** Zhang L, Zhu WF, Shi F, Chen HY, Chen XJ. Automated Segmentation Of

Intraretinal Cystoid Macular Edema For Retinal 3D OCT Images With Macular Hole;

2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI); [Accessed

On23July,2015].Available at: <https://ieeexplore.ieee.org/abstract/document/7164160>.

[[Google Scholar](https://scholar.google.com/scholar_lookup?journal=Automated+Segmentation+Of+Intraretinal+Cystoid+Macular+Edema+For+Retinal+3D+OCT+Images+With+Macular+Hole&author=L+Zhang&author=WF+Zhu&author=F+Shi&author=HY+Chen&author=XJ+Chen&)]

**[7].** Slokom N, Zghal I, Trabelsi H. Segmentation of Cyctoids Macular Edema in Optical

Cohenrence Tomography; 2016 2nd International Conference on Advanced

Technologies for Signal and Image Processing (ATSIP); [Accessed on 28 July,

2016]. Available at: <https://ieeexplore.ieee.org/abstract/document/7523096>.

**[8].** Fernandez DC. Delineating fluid-filled region boundaries in optical coherence

tomography images of the retina. IEEE Trans Med Imaging.2005; 24(8):929–945.

**[9].** Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic

segmentation; The IEEE Conference on Computer Vision and Pattern Recognition

(CVPR); [Accessed on 20 May, 2016]. Available at:

<https://arxiv.org/pdf/1605.06211.pdf>. [[Google Scholar](https://scholar.google.com/scholar_lookup?journal=Fully+convolutional+networks+for+semantic+segmentation&author=J+Long&author=E+Shelhamer&author=T+Darrell&)]

**[10].** Bai FL, Marques MJ, Gibson SJ. Cystoid macular edema segmentation of Optical

Coherence Tomography images using fully convolutional neural networks and fully

connected CRFs; Computer Vision and Pattern Recognition; [Accessed on 15 Sep.,

2017]. Available at: <https://arxiv.org/abs/1709.05324>. [[Google Scholar](https://scholar.google.com/scholar_lookup?journal=Cystoid+macular+edema+segmentation+of+Optical+Coherence+Tomography+images+using+fully+convolutional+neural+networks+and+fully+connected+CRFs&author=FL+Bai&author=MJ+Marques&author=SJ+Gibson&)]