Medical Image Segmentation in Multi-domain Using Deep Learning Methods

By

Shaown Paul (011191224)
Sazzad Hasan (011191221)
Md Eram Bin Tanbir (011191182)
Jayed Hasan (011191051)
Rakibul Islam (011191220)
Rahul Mallik (011191242)

Submitted in partial fulfilment of the requirements
of the degree of Bachelor of Science in Computer Science and Engineering

February 11, 2023

Department of Computer Science and Engineering
United International University

Abstract

Previously, physicians had to analyze medical reports based on their previous knowledge, to diagnose diseases or abnormalities of a patient. And sometimes, this manual analysis leads to inconsistency, and also needs time to do that. Recently, deep learning has been used in medical image segmentation and made automatic disease diagnosis possible from those medical reports. In medical image segmentation, U-Net and some of its variants have made significant improvements. In this, paper, we proposed a novel architecture based on a variant of U-Net, called U2-Net, which is a nested deep U-Net architecture to capture more contextual information from different scales. We also used squeeze and excitation Residual Parallel Convolution Block Attention Module (RPCBAM), Dense Convolution Block Attention Module (DCBAM), and a novel Attention Fusion Block (AFB) with residual CBAM to capture more spatial attention. For the evaluation of our model, we fit our models into several datasets: Drive, ISBI-2012, CHASEDB-1, HRF(diabetic), and Wound. For the drive dataset, model-1, model-2, and model-3 gave the Dice coefficient of 87%, 84.45 %, and 84.70% respectively.

Acknowledgements

This work would have not been possible without the input and support of many people over the last three months. We would like to express our gratitude to everyone who contributed to it in some way or another. First, we would like to thank our Supervisor MD Ryhan Ahmed, Assistant Professor at United International University for his constant advice and guidance. His suggestions challenged us to think outside of the box, his feedback improve our problem-solving skills and his encouragement provided us strength during times of disappointment. Last but not the least, we owe our family including our parents for their unconditional love and immense emotional support.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

In his chapter, we will highlight the project overview, motivation, objectives, methodology, and outcomes of our project.

1.1 Project Overview

In recent days medical image is very essential and has been used globally in the field of biomedical engineering [1]. Medical imaging plays a vital role in clinical diagnosis, mainly in treatment planning, surgery, and prognosis evaluation. Medical images are generally used for gathering life-saving information by non-invasive peering at human organs [2]. Nowadays medical imaging tools are getting modernized day by day. So, we can get medical imaging information from Magnetic Resonance Imaging(MRI), Computed Tomography(CT), Positron emission tomography(PET), digital pathology, and microscopy. These images can give us an anatomical view of organs. But there is a drawback. It is not so easy for radiologists to identify organs or lesions from these medical images. For example, the images we have got from CT and MRI both are very important in clinical diagnosis but these are in the 3D format we want to segment these images in 2D format. If the radiologist's hand marked all the images, then it will take up to 15 minutes per image [3]. That's why we need to develop an automated method that can identify, segment, and quantify lesion tissues.

But there are many challenges to segmenting medical images. First of all, there is a lack of image samples for a particular disease. Secondly, the same intensity and variable position shape and region of interest(ROI) make it difficult to segmentation methods. Finally, many factors in image capturing, such as sampling artifacts, Spatial aliasing, and noise factors can cause the boundary of structures unclear and disconnected. To resolve these issues many scientists have proposed different methods of segmentation. But deep learning methods have achieved outstanding performance in the sector of medical image segmentation.

The deep learning method not only learns the non-linear mapping between input data and estimated results but also learns the hidden features [4]. Among all deep learning methods of convolution, neural networks (CNNs) are the most successful image segmentation method [5]. In these CNN methods, U-shaped networks have made excellent results, that's why they became popular architectures for medical image segmentation. A recent study has proven that U-shaped architecture gained the best performance on 6 publicly available segmentation tasks and it can automatically go with any dataset [6]. Besides the U-shaped networks have shown the best performance in the segmentation of medical images and as a consequence, it becomes a benchmark in medical image segmentation within a very short time.

1.2 Motivation

Medical images have become an important reference for doctors to understand and analyze the disease, which helps diagnose disease and estimate treatment. Image segmentation is considered the most important medical imaging modality as it extracts the region of interest (ROI) through a semi-automated or automated process. It divides an image into areas based on a specific description, such as Body organ/tissue segmentation in medical applications for boundary detection, locating, segmenting, and quantifying the lesion tissues, tumor detection/segmentation, and mass detection. Our main objective is to implement deep learning-based U-shaped architecture to segment the blood vessels from the retinal images and polyps from the colonoscopy images. Because our proposed algorithm can help us to detect polyp size in the colon and the flow of blood vessels inside the retina as the doctor and clinician takes quick decisions for proper treatment.

1.3 Objectives

In our paper, we mainly try to find a deep learning method where we can segment the blood vessels from the retinal images and polyps from the colonoscopy images for identifying detailed information about these areas by isolating only the necessary information. We first analyze medical images which we can get from different sources. But it is very difficult to identify the problem in blood vessels inside areas of the retina from retinal images. Besides, identifying colorectal polyps and tumors inside the colon from colonoscopy images is also the toughest task. Because the blood vessels and polyp size are diverse. There are different challenges to the segmentation of the medical image. Therefore, we explored several methods for medical image segmentation proposed by various scientists. So, we see that deep learning-based methods give excellent results in the field of medical image segmentation. Among all deep learning methods, the most successful image analysis methods are convolution neural networks (CNNs). In these CNN methods, U-shaped nets have had excellent performance and thus have become the popular technology for medical image segmentation. For this reason, we developed a CNN-based U-shaped architecture that segments blood vessels and polyps from retinal and colonoscopy images.

1.4 Our Contribution

In our project, we dealt with medical images. Nowadays we see there are many diseases such that Colorectal cancer, Colonic polyps, Diabetic Retinopathy, Skin Lesions, etc. Medical image segmentation helps doctors/clinicians quickly identify diseases. The ability to isolate only relevant areas through medical image segmentation enables a more accurate interpretation of anatomical data. image segmentation can remove any unwanted noises from a medical image. We have applied our model to multiple domains such as the retinal dataset, colonoscopy polyp dataset, and skin lesion dataset. In our model, we applied a nested U-shaped architecture which is an idea we got from a paper named U2-Net[7]. The authors of that paper used U2-Net architecture for salient object detection. Without reducing the resolution of an image, it extracts good feature maps. Though the computational cost isn't increasing. To achieve this they used Residul U-block (RSU) block which can extract features without downgrading the feature map resolution. They got better performance without using any pre-trained model. we are the first to use U2-Net for medical image segmentation with some major changes. We have proposed two different models by fine-tuning them in U2-Net architecture. In our first proposed model, we used Squeeze and Excitation(SE)[8]block in the skip connection of the outer encoder part of U2-Net architecture. SE is a module used in Convolutional Neural Networks (CNN) to increase performance with a slight amount of computational cost. Squeeze and Excitation module only attention to channel-wise information. Inside this module, there has a squeeze module, excitation module, and scale module. By using the SE block the performance of the metrics is increased tremendously. In this method, we get an 87% dice coefficient in DRIVE,98% in ISB-2012,90.19% in CHASEDB-1,85.71% in HRF, and 88.22% in wound datasets. In our proposed model 2, We have used Residual Parallel CBAM(RPCBAM) and AFB module. RPCBAM can boost representational power. The specialty of this paper is it has less overhead and easily copes with any kind of convolutional architecture [9]. We used parallel CBAM in the outer Unet's encoder part. We also used the AFB module to skip connections in the outer U-Net. In this model, we got a dice coefficient of 84.45% in DRIVE and 88 % in CHASEDB-1,81% HRF datasets. Also, we proposed another model where we proposed a new version of CBAM and use the concept of the Densely connected network and make **DCBAM**. We also use the squeeze and excitation concept in the skip path for passing the semantic feature on every of decoder side. This model also performed very well like our proposed model-2. In this model, we got a dice coefficient of 84.70% in DRIVE,96.73 % in ISBI-2012, and 88.34 % in CHASEDB-1,86.73% HRF datasets.

1.5 Organization of the Report

In the first chapter, we discuss the introduction of our project. Then we discuss the objective then we discuss the motivation of our paper. Finally, we discuss the outcome of the paper that we get from this paper. In the second chapter, we discuss the project's background, literature review, related applications, and similar research then we analyze gaps and finally, we give the overview of the chapter. Then the next chapter we discuss the functional and non-functional requirements, diagrams, and user interface design. Project strategy and assigned duties for all members, project standards, and restrictions are discussed in this chapter. Besides cost analysis as well as technical problems are also discussed in chapter four. Finally, in chapter five we discuss the overall project summary.

Chapter 2

Background

In this chapter, we analyze the related studies, some similar types of applications, and a gap analysis of our project.

2.1 Preliminaries

Some basic knowledge of deep learning and biomedical image segmentation is required to understand our project. such as Neural network, CNN, Attention mechanism, Residual network, U-Net architecture, Downsampling, and Upsampling are some topic which is required know. Basically, deep learning is a category of machine learning that is a nonlinear processing unit of multiple layers that can extract features and transform [10]where the current layer's output is the input of the next layer. Deep learning thinks the way the human brain thinks. There are many deep learning-based models which can segment an image accurately [11]. Neural networks are a set of deep learning methods. It is inspired by human biology.A neural network works in the same way as the neurons of the human brain work. It is used to extract patterns and solve problems. A neural network model has the ability to change its hyperparameters and thus reduces the total error of the model. The convolutional neural network is a type of Artificial Neural Network. It is also called a deep neural network. CNN has multiple layers which are: the convolutional layer, pooling layer, and fully connected layer [12].CNN deals with big data which can be used for image segmentation and natural language processing [12].

In deep learning, the attention mechanism comes from the attention mechanism of the human brain [13]. When a human brain receives multiple pieces of information at a time, the brain does not process all information. it cares about just some important information and ignores less important information. The attention mechanism is first used in deep learning to better the distinguishing ability on the RNN [14]. After that, it is extensively used in image segmentation, speech recognition, machine translation, etc. Nowadays attention mechanism is used in medical image segmentation [15]. There is a problem in the neural networks called vanishing gradient that causes the gradient to become 0. And if a neural network contains a lot of hidden layers, then the error rate becomes higher.

To avoid this problem researchers at Microsoft Research first proposed a new technique called Residual Network [16] in 2015. This architecture introduced a new technique where they used skip connections. The main idea is that if any layer decreases the performance of the neural network, then this technique will skip this layer by regularization. U-Net architecture is an end-to-end convolutional neural network architecture [14]. It is called U-Net for its 'U' shape. It has two parts (encoder and decoder) that's why it is called symmetric segmentation. UNet architecture is good for cell segmentation [14]. It can classify and identify the area of diseases. In image segmentation downsampling decrease the resolution of feature maps [14]. It is used for feature extraction. pooling is a strategy that is used in downsampling. On the other hand upsampling increases the resolution of feature maps [14]. It enhances the original image. some upsampling methods are unpooling, deconvolution, and bilinear.

2.2 Literature Review

In this section, we show up the related applications and paper studies that we were studying and analyzing for our project.

2.2.1 Relevant Applications

We have found three application which is related to our study. The applications are BIOMEDISA,3D SLICER, and FIJI.

BIOMEDISA is an open-source online platform that can segment large volumetric images. Biomedical is a semi-automated segmentation that can segment biomedical images like MRI and CT scan images. It allows deep learning for the fully automated segmentation of a series of similar samples.

3D SLICER is a desktop software that is widely used for medical image research. It has the ability to segment 2D/3D/4D images simultaneously. There are many tools for the manual and automatic registration of images.

FIJI is an image processing software that can process and analyze images. It can process biomedical and other images simultaneously. It is free software that has some additional features which extract patterns from images.

2.2.2 Relevant Studies

In this section of this chapter, we will present the summary of the paper which we are studying for acquiring knowledge for our project. We will highlight a summary of some extraordinary papers that are related to our topic.

Xuebin et.al. [7] proposed a two-level U-shaped architecture that is made for salient object detection(SOD). It can segment the most visually attractive object in an image. This nested u2net architecture is made with two nested u net architectures. Without reducing the resolution of an image, it extracts good feature maps. Though the computational cost isn't increasing. To achieve this they used Residul U-block (RSU) block which can extract features without downgrading the feature map resolution. They got better performance without using any pre-trained model. They also provide U2-Net small version for memory issues. They did an experiment with their model on six different datasets. Sirojbek et al. [17] proposed an automatic polyp segmentation model. That model can segment polyps to prevent deadly colon cancer. They modified the U-Net architecture to segment polyp images. In their proposed model, they mentioned that UNET++ and DenseNets(densely connected convolutional network) are the base of their architecture. In this paper, they used a module that is a dense block. They tested their model on the Kvasir-SEG dataset and CVC-612 datasets. They have shown that their proposed model gained 90% dice coefficient score on Kvasir-SEG dataset and 91% on CVC-612 dataset. In the future, they will implement their model in real-time surgical robots. Kun Yang et al. [18]proposed deep learning model based on Mask R-CNN. can automatically detect and segment colorectal polyps. They introduced a shuffle-efficient channel attention network(sECA-Net) by obtaining cross-channel interaction. They used CVC-clinicDB, ETIS-Lbrib polyp DB and Kvasir-SEG dataset.They achieved 94.9% precision,96.9% recall,95.9%F1 and 96.5% F2 score on that uses datasets. They highlighted that their model has a good ability to detect polyps during the time of endoscopy and it can segment polyps early and timely before it turns into deadly cancer. Woo et.al. [9] proposed a convolution block attention module(CBAM) used to feed-forward convolutional neural networks. By leveraging the attention mechanism, it can boost representational power.The specialty of this paper is it has less overhead and easily copes with any kind of convolutional architecture. Inside this CBAM module, they make focus on both channel and spatial areas as convolutional neural network extract features by cross-channel and spatial information. They applied their module to several datasets and showed that their model has less number of parameters so the model is very lightweight. In this study, Jie Hu et.al[8] proposed a module that is used for increasing channel-wise interdependencies. It is a module used in Convolutional Neural Networks (CNN) to increase performance with a slight amount of computational cost. Squeeze and Excitation module only attention to channel-wise information. They suggest a feature recalibration approach that enables the network to learn how to selectively emphasize informative traits and suppress less helpful ones by using global information. Inside this module, there has a squeeze module, excitation module, and scale module. The squeeze and Excitation block is a very lightweight module so there has a slight amount of computation cost. Gendry et al. [19] mainly propose a model that segments the retina blood vessel and balances the classification of the blood vessel in a good manner where the rate of unbalanced classes is very high. The methodology of their model is created by two convolutional neural networks where chained between each other. In the architecture, they basically used two UNet where the first UNet is a general UNet and the other UNet is UNet with residual blocks. The first UNet is working for feature extraction whereas the second UNet with residual block helps to find the performance ambiguity or the new characteristic. They used DRIVE and CHASEDB retina datasets for developing their model. The final comparison their propose model works better than the older model in two matrices which is F1-score (0.8312) and Accuracy (0.9766). Jonthan et al. [20] on the segmentation of lung region which is overlapped by some abnormalities which may happen for some diseases in this paper. The method they are using consists of four sections which are image acquisition, initial segmentation, reconstruction and final segmentation. They use the Montgomery County dataset for their model.The final accuracy rate of this model is 96.97% and the Dice value and Jaccard index is 93.56% and 88.07%. The best medical picture segmentation techniques in recent years have been UNet and its most recent extensions, such TransUNet. These networks are parameter-heavy, computationally difficult, and slow to utilize, hence they cannot be efficiently used for quick image segmentation in point-of-care applications. They suggest "UNeXt: MLP-based Rapid Medical Image Segmentation Network"[21] as a means of assisting segmentation in order to decrease the number of parameters and computational complexity as well as to produce a better representation. They choose the International Skin Imaging Collaboration (ISIC 2018) and Breast Ultrasound Images (BUSI) datasets. They demonstrate that UNeXt reduces the number of parameters by testing it on several medical picture segmentation datasets and shows that it reduces the number of parameters by 72x, decreases the computational complexity by 68x, and improves the inference speed by 10x.

Zhenzhen et al. [22]proposed a new architecture called Densely Connected U-Net Architecture (DenseUNet). This proposed method uses a dense block to enhance feature extraction capacity and uses a multi-functional fuse block that combines feature maps from different levels to increase the correctness of feature extraction. XiwangXie et al. [23]proposed a hierarchical contextual integrated network called CHI-Net which is built with two main modules. One is Dense Dilated Convolution (DDC) and another is Stacked Residual Pooling (SRP). The Dense Dilated Convolution (DDC) module can hierarchically acquire substantial key features by joining four cascading branches of hybrid dilated convolutions, useful for extracting features at different scales. The SRP module integrates the detailed characteristics of the encoder through multiple effective visual fields, aimed at generating more distinctive characteristics. PengfeiSong et al. [24]proposed a new method called Multi-Scale Attention Decoding-AMNet. In this method, they first perform a multi-scale fusion of the high-level feature information extracted from the backbone network by oversampling and down-sampling while aggregating the high-level feature information to create an initial predictive segmentation map for the next generation of contextual instructions. Segmentation is an important step in biomedical image analysis tasks. Recently, convolutional neural networks (CNNs) are being used more and more in the field of medical imaging. However, the standard model still has some drawbacks. Due to the significant loss of spatial information during the encoding stage, it is often difficult to recover low-level details of visual features by simple deconvolution, resulting in feature maps that are sparse and perform poorly.So, they propose his new method called DCACNet[25].A Dual Context Aggregation and Attention-Driven Cross-Deconvolution Network for Medical Image Segmentation. Experimental results on two published medical data sets demonstrate that Dual Context Aggregation (DACNet) can be used to obtain more detailed semantic feature information from medical images. For their experiments, they use the CHAOS and Herlev datasets.

Zaiwang Gu et al. [26] propose a context encoder network (CE-Net) that can capture more high-level information and preserve spatial information. It contains a feature encoder module, a context extractor, and a feature decoder module. For the feature extractor, they used a pre-trained Res Net block and for the context extractor module, they propose a residual multi-kernel pooling block and a dense atrous convolution block. In this process, at first, the picture is fed are fed into the feature encoder module. Then it is replaced with the original U-Net encoder block by the ResNet-34 block pre-trained from ImageNet. Then The context extractor generates a dense atrous convolution (DAC) block and a residual multi-kernel pooling (RMP) block. And finally, in the feature decoder block, there will decode the data into a segmented image. For tooth root segmentation on the oral X-ray image, they propose[27] a novel end-to-end U-Net-like Group Transformer Network (GT U-Net). Here the encoders and decoders are replaced by a group. This grouping structure and the bottleneck structure reduce the computational cost. Actually, GT U-Net is a hybrid structure of convolution and Transformer. For shape prior knowledge they also used shape-sensitive Fourier Descriptor (FD) loss function. Zabir et al. [28] presented a deep learning model that can automatically segment skin lesions and detect melanoma from dermoscopy images. They used U-Net in order to segment out the lesion from the surrounding skin. And they evaluated the model on two different datasets. They resolve the overfitting problem by using the spatial dropout technique along with U-Net. And augmented the dataset to increase the total number of samples. Their motivation was to improve the diagnosis accuracy of melanoma (which is a malignant skin cancer). They find 0.87% and 0.80 Jaccard Index in ISIC 2018 dataset. Marayan et al. [29] proposed Squeeze and Excitation (SE) block, bi-directional ConvLSTM (BConvLSTM), and the mechanism of dense convolutions to improve the U-Net for medical image segmentation. Utilizing a self-gating mechanism, Squeeze and Excitation recalibrate the channel-wise feature responses. The mechanism of dense convolutions is used here to strengthen feature propagation and encourage feature reuse. The BConvLSTM is used at all network levels to combine the feature maps. Chiu-Han Hsiao et al [30] mainly focused on the preprocessing methods needed before processing medical images in a neural network model. The experimental results showed that the proposed preprocessing methods or models significantly improve accuracy compared to the case without data preprocessing. Specifically, the dice score was improved from 0.9436% to 0.9648% for renal segmentation and 0.7294% for all types of tumor detections. Mobeen et al. [31]proposed a BrainSeg-Net model which works with the encoder-decoder process. In this model, the Feature Enhancer (FE) block is used to extract the middle-level features from low-level features from the shallow layers and shares them with the dense layers. They used a custom-designed loss function to address the problem associated with the imbalance of class. For this architecture, they used BraTS2017, BraTS 2018, and BraTS 2019 datasets. This BrainSeg-Net segmentation process actually enhanced the MR brain tumor's location and spatial feature.

2.3 Gap Analysis

The authors of U2-Net only used this architecture in salient object detection. But, this model was not used first in medical image segmentation to handle medical image segmentation challenges. Secondly, in U2-Net architecture, the authors used a simple skip connection, just like U-Net, without reducing the semantic gaps between the feature maps of the encoder and decoder. Furthermore, U2-Net is a very deep network, so adding a new module will increase the number of parameters by a large amount, so keeping the number of parameters comparatively low, is another challenge. Lastly, in U2-Net, no attention module has been used to improve the performance of the overall architecture.

2.4 Summary

In this chapter, we talked about the related application to our project. We provide the uses of some features of these types of image segmentation applications. We also discussed the important part of the work which is a literature review. We found some extraordinary ideas from different types of medical image segmentation in some papers. We also found the drawback of these existing papers which helps us to do the Gap analysis for our project.

Chapter 3

Project Design

In this chapter, we will present the requirement analysis, project plan, and task allocation of our project and also show some portion of the Gantt chart.

3.1 Requirement Analysis

This section will discuss the two requirements which we have to consider when we want to build a system. One is Functional requirements and the other is Non-functional requirements.

3.1.1 Functional Requirements:

In our project, our main goal is to help doctors and clinicians for finding diseases properly from medical image analysis. For creating this model, we have to ensure some functional requirements which help our model to work properly. At first, we have to ensure that we have proper datasets. We use medical image datasets that are not properly organized. For this, we have to do data pre-processing and make the datasets suitable for our model. By using this dataset, we will build a model which will give us a segmented image from medical images so that the radiologist or doctor finds the problem from the images clearly.

3.1.2 Non-Functional Requirements:

Non-functional requirement defines the performance of the built model and the quality of this model or system in image segmentation nonfunctional requirements is very important because the quality of the segmented image should be as perfect as possible. For our model, some non-functional requirement defines by performance, accuracy, Stress, and Security. Performance means how fast and well our model segmented an image. It is very important for our model building because we need as fast as possible to find the segmented images for knowing the problem quickly for diagnosis. Also, we have to be concerned about the accuracy of our model. Because if we don't get the accurate segmented image it will hamper proper diagnosis. The other most important thing is Stress. When we deliver this model by an application the system user load has to be performed properly and should not being malfunctioned. Security is a challenging task for a machine-learning model. Because we should not accept any change in our model and we have to ensure data protection. These non-functional requirements are becoming our model more comfortable and ensure to follow the legal rules and originate security policy for our model.

3.2 Methodology and Design

In this section, we show the diagram design of our project.

3.2.1 Context Diagram

Context diagram is given in Figure 3.1

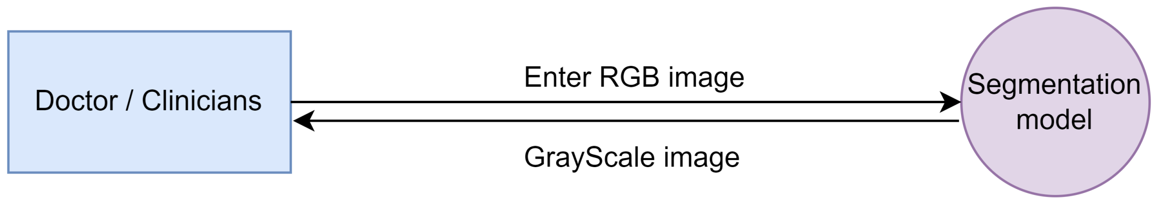


Figure 3.1: Context Diagram

3.2.2 Data Flow Diagram

Data flow diagram is given in Figure 3.2

3.3 Planning of the Project

In this section, we have described the timeline of our work process till now.

- Problem Characterization - 1 July – 21 July 2022;
- Literature Review and Skill development - 23 July – 31 Aug
- Data collection and Data Pre-processing - 2 Sep – 15-Sep
- Initial Model Implementation - 17 Sep – 15 Oct
- Metrics Selection and Definition - 16 Oct – 30 Oct

3.3.1 Ganchart Overview

3.3.2 Task Allocation

In this section, we will show the task allocations for each member for our project.

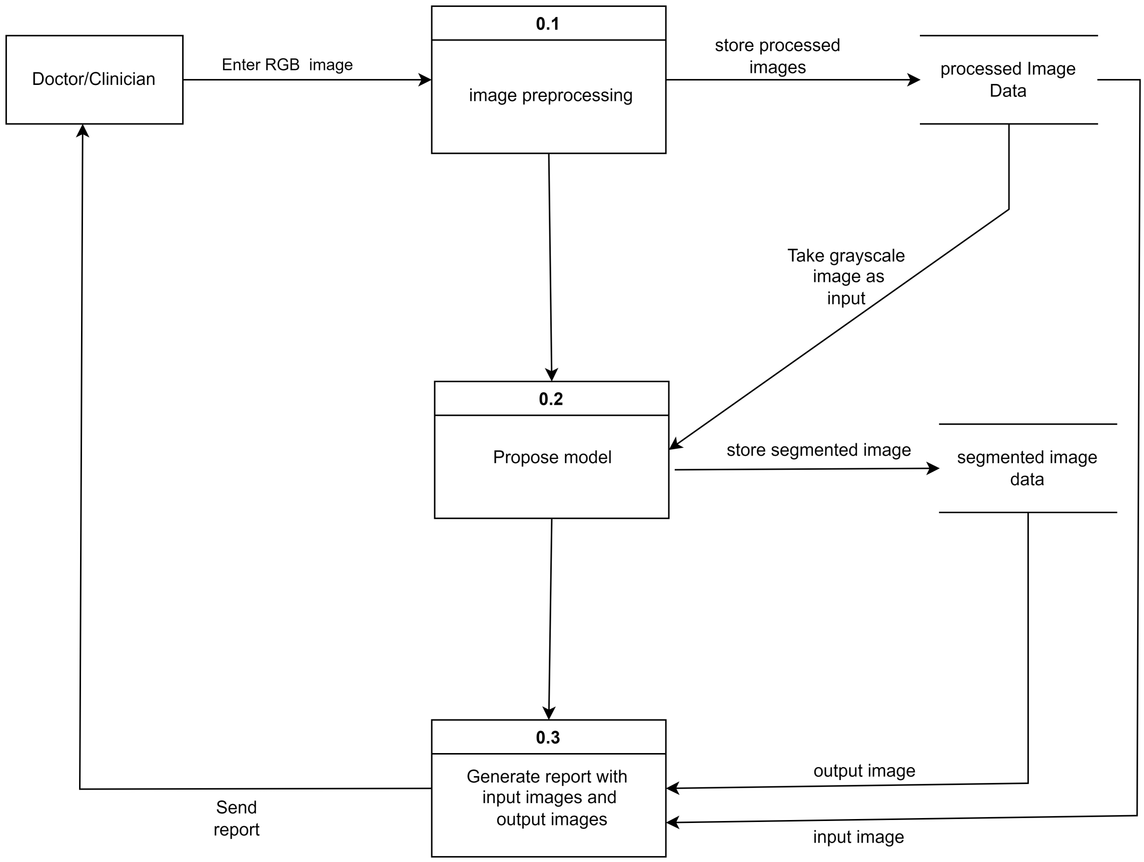
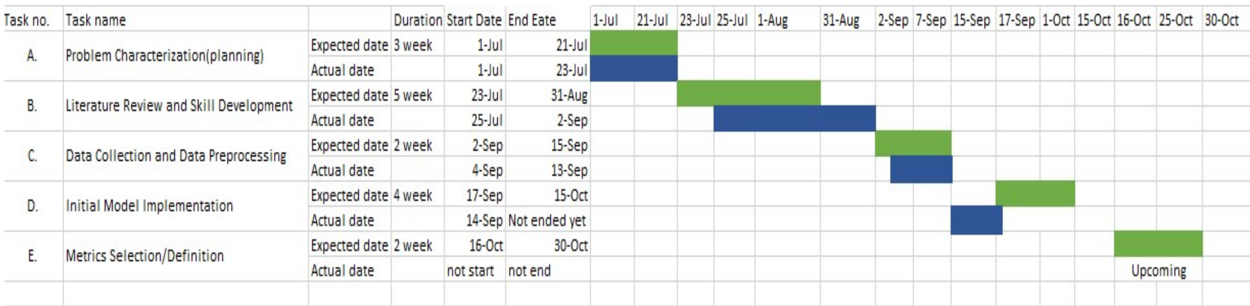
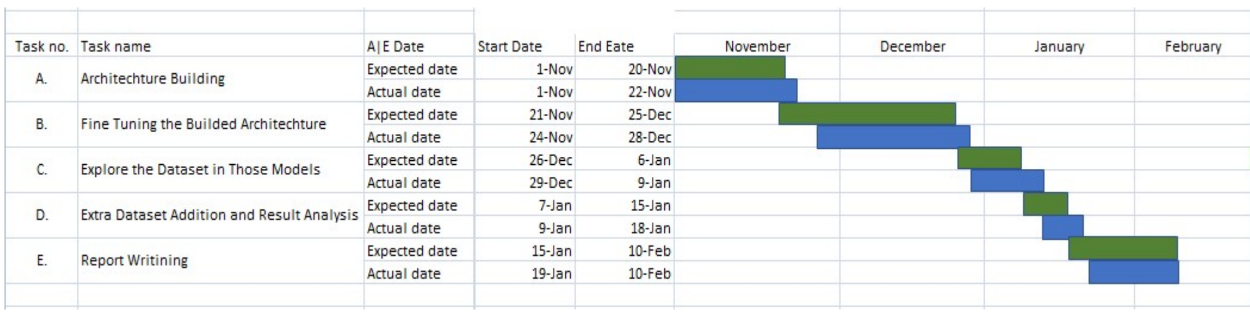


Figure 3.2: Data Flow Diagram



Task no.	Task name		Duration	Start Date	End Eate	1-Jul	21-Jul	23-Jul	25-Jul	1-Aug	31-Aug	2-Sep	7-Sep	15-Sep	17-Sep	1-Oct	15-Oct	16-Oct	25-Oct	30-Oct
A.	Problem Characterization(planning)	Expected date	3 week	1-Jul	21-Jul															
		Actual date		1-Jul	23-Jul															
B.	Literature Review and Skill Development	Expected date	5 week	23-Jul	31-Aug															
		Actual date		25-Jul	2-Sep															
C.	Data Collection and Data Preprocessing	Expected date	2 week	2-Sep	15-Sep															
		Actual date		4-Sep	13-Sep															
D.	Initial Model Implementation	Expected date	4 week	17-Sep	15-Oct															
		Actual date		14-Sep	Not ended yet															
E.	Metrics Selection/Definition	Expected date	2 week	16-Oct	30-Oct															
		Actual date		not start	not end													Upcoming		

Figure 3.3: Gantt Chart for FYDP-1



Task no.	Task name	A|E Date	Start Date	End Eate	November	December	January	February
A.	Architechture Building	Expected date	1-Nov	20-Nov				
		Actual date	1-Nov	22-Nov				
B.	Fine Tuning the Builded Architechture	Expected date	21-Nov	25-Dec				
		Actual date	24-Nov	28-Dec				
C.	Explore the Dataset in Those Models	Expected date	26-Dec	6-Jan				
		Actual date	29-Dec	9-Jan				
D.	Extra Dataset Addition and Result Analysis	Expected date	7-Jan	15-Jan				
		Actual date	9-Jan	18-Jan				
E.	Report Writining	Expected date	15-Jan	10-Feb				
		Actual date	19-Jan	10-Feb				

Figure 3.4: Gantt Chart for FYDP-2

3.3.3 Fydp-1 Task

- Problem Characterization - everyone
- Literature Review and Skill development - everyone
- Data collection and Data Preprocessing - Jayed, Sazzad
- Initial Model Implementation - Shaown, Jayed, Sazzad
- Metrics Selection and Definition - Shaown, Jayed, Sazzad
- Writing the first version of FYDP report - Eram, Shaown, Sazzad, Jayed, Sumon, Rahul

3.3.4 Fydp-2 Task

- Architecture Building - Shaown, Bakku, Jayed
- Fine Tuning the Builded Architecture - Shaown, Bakku, Jayed
- Explore the Dataset in Those Models - Shaown, Bakku, Jayed, Eram
- Extra Dataset Addition and Result Analysis - everyone
- Metrics Selection and Definition - Shaown, Jayed, Sazzad
- Writing the first version of FYDP report - Eram, Shaown, Sazzad, Jayed, Sumon, Rahul

3.4 Summary

In this chapter, we have explained the requirements and function of our project. We discussed the details of the functional and non-functional requirements with respect to our project. In the second section, we add the different types of diagrams like context diagrams and data-flow diagrams of our project. Also, we include and described the project plan where we add the Gantt chart and task allocation for the whole project plan.

Chapter 4

Implementation and Results

In this chapter, we discussed our project's environmental setup, discussions, results, testing, and evaluation.

4.1 Environment Setup

4.1.1 Data Sets

Our model is based on a segmentation algorithm where we have used multi-domain biomedical image datasets. For training and testing purposes, we collected some datasets and we have run those datasets into our proposed model. So in our whole work, we have used four datasets of the different domains which are DRIVE, CHASEDB1, ISBI-2012, and HRF. Here DRIVE, CHASEDB1, and HRF are retinal image datasets and ISBI-2012 is a microscopic image dataset.

The **DRIVE** dataset contains some retina images that can be used to segment blood vessels in those images. From the segmented images, it's possible to diagnose various diseases such as diabetes, arteriosclerosis, and hypertension and helps in the treatment of those diseases. This DRIVE database was collected from a diabetic retinopathy screening program in The Netherlands. They obtained this database by screening 400 diabetic patients of age between 25-90 years. In this dataset, there are 20 images and their corresponding masks in each of the training and testing datasets.

In **ISBI 2012** dataset 30 serial section Transmission Electron Microscopy (ssTEM) images of the ventral nerve cord of a Drosophila larva are included in the dataset (VNC). The pictures show a sequence of 3D slices taken one after another. This dataset also includes corresponding segmentation ground truths.

CHASEDB1 is a public retinal vascular reference dataset that has been made available by St. George's University of London and Kingston University of London. This is a subset of the Child Heart and Health Study in England (CHASE) dataset that contains retinal scans of children of different ethnic backgrounds. This group includes 28 retinal photos from 14 of the study's recruited kids, each one taken from both eyes. Each retinal image in this collection is accompanied by two ground truth images. This is presented as

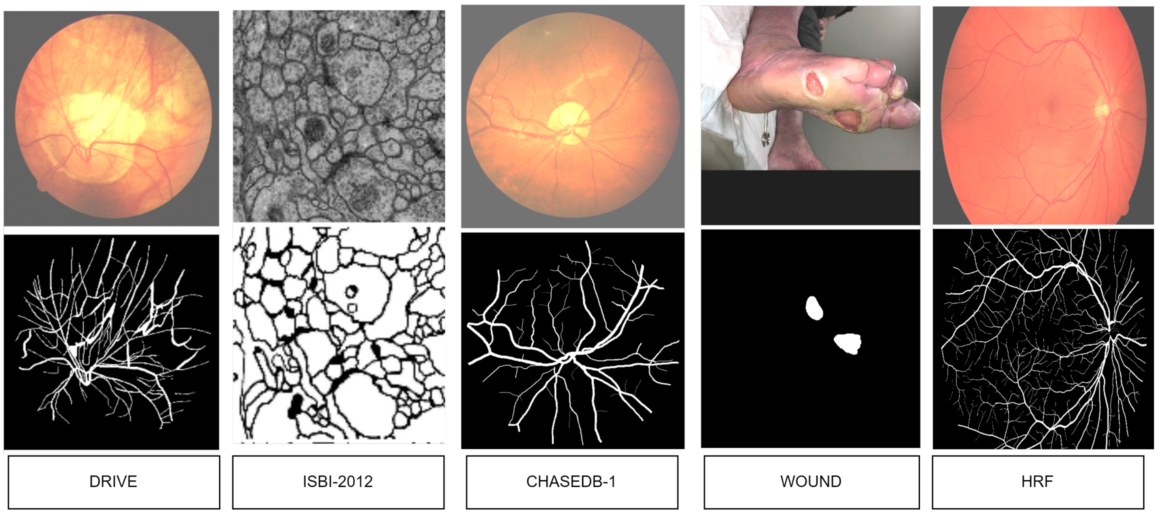


Figure 4.1: Input images and their corresponding segmentation masks in the dataset

two hand vessel segmentations produced by two separate human observers for each image. Each pixel is given a "1" label if it is a blood vessel and a "0" label otherwise. By making this subgroup accessible, the scientific community can practice and evaluate computer vision algorithms (specifically vessel segmentation methodologies). Most crucially, this subset enables performance comparisons since many algorithms can be tested on the same database and their results may be directly compared.

A collaborative research team created **HRF** database to facilitate a comparative study on automatic segmentation methods using retinal fundus pictures. Currently, the public database includes 15 photos of healthy patients, 15 photographs of diabetic retinopathy patients, and 15 images of glaucoma patients. For every image, there are binary, gold-standard vessel segmentation images accessible. For some datasets, field of view (FOV) masks are also offered. A team of retinal image analysis specialists and doctors from collaborative ophthalmology clinics create the standard gold data.

Table 4.1: **Dataset Analysis**

Datasets	**Training Samples**	**Modality**	**Sample After Augmentation**
DRIVE	20	Retinal	336
ISBI-2012	30	Electron Microscopy	504
CHASEDB-1	28	Retinal	472
HRF	45	Retinal	757
Wound	1210	Human Body Wound	1210(no augmentation)

4.2 Evaluation

After collecting those datasets for our model, we fitted them into the three different models we proposed. We skipped the feature selection part because there is no need for feature selection. But, we need to do the preprocessing steps such as image augmentation, because we didn't have enough images for training. And by augmenting the images, we increased the total amount of images.

4.2.1 Proposed Model-1

In our proposed model we used the concept of an architecture called U square net for semantic segmentation[7]. This architecture gives appreciable performance for medical image segmentation. We also do fine-tune this architecture and we use the concept of Squeeze and excitation[8] block and add this block in the encoder part of the U-square net architecture. By using the SE block the performance of the metrics is increased tremendously. We have used 512 X 512 image resolution and divided the whole datasets into training (80%), testing (10%), and validation(10%) datasets.

Residual U-blocks

We used Residual U-Block[Figure 4.1] to capture multi-scale features, and we got this idea from U-Net architecture.RSU block is itself a U-Net. But, to keep the number of trainable parameters less than enough, we used U-Net of a different number of layers. To extract local features we simply used convolutional layers. And then we used the number of layers, L, so that, as the outer Unet gets deeper, the number of layers, L gets smaller. Greater L results in a deeper residual U-block (RSU), more pooling activities, a wider variety of receptive fields, and richer local and global features. By setting this option, it is possible to extract multi-scale features from input feature maps with any desired level of spatial resolution. Progressive upsampling, concatenation, and convolution encode the multi-scale features into high-resolution feature maps from successively downsampled feature maps. The loss of fine features brought on by direct upsampling at large scales is reduced by this method. Then we added local feature maps and multi-scale feature maps using a residual connection.

Squeeze and Excitation Block (SE Block)

In this study, Jie Hu[8] et.al proposed a module that is used for increasing channel-wise interdependencies. It is a module used in Convolutional Neural Networks (CNN) to increase performance with a slight amount of computational cost. The Squeeze and Excitation module only applies attention to channel-wise information. They suggest a feature recalibration approach that enables the network to learn how to selectively emphasize informative traits and suppress less helpful ones by using global information. Inside this module, there has a squeeze module, excitation module, and scale module. The squeeze and Excitation block is a very lightweight module so there is a slight amount of computation cost.

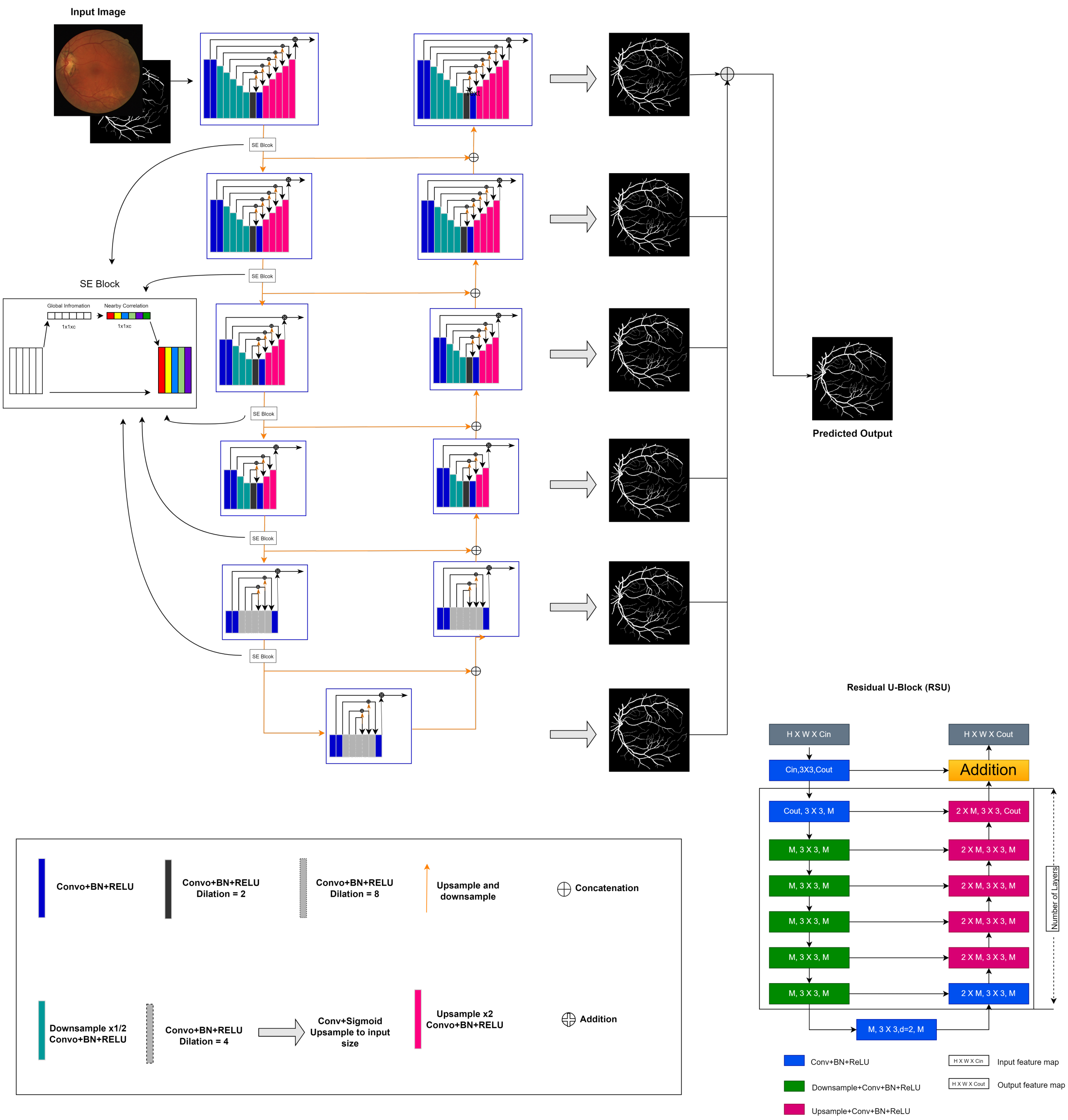


Figure 4.2: Architecture of our proposed model-1

4.2.2 Proposed Model-2

We used U2̂-Net [7] as a backbone model of our proposed model, where we used Residual U Block (RSU), Parallel Convolutional Block Attention Module (RPCBAM), and Attention Fusion Block (AFB). In this model, we tried to add more attention modules, for example: after every encoder, we added an attention block and designed a novel block to capture spatial information along with reducing the semantic gap between the feature maps from the encoder and the feature maps from the decoder.

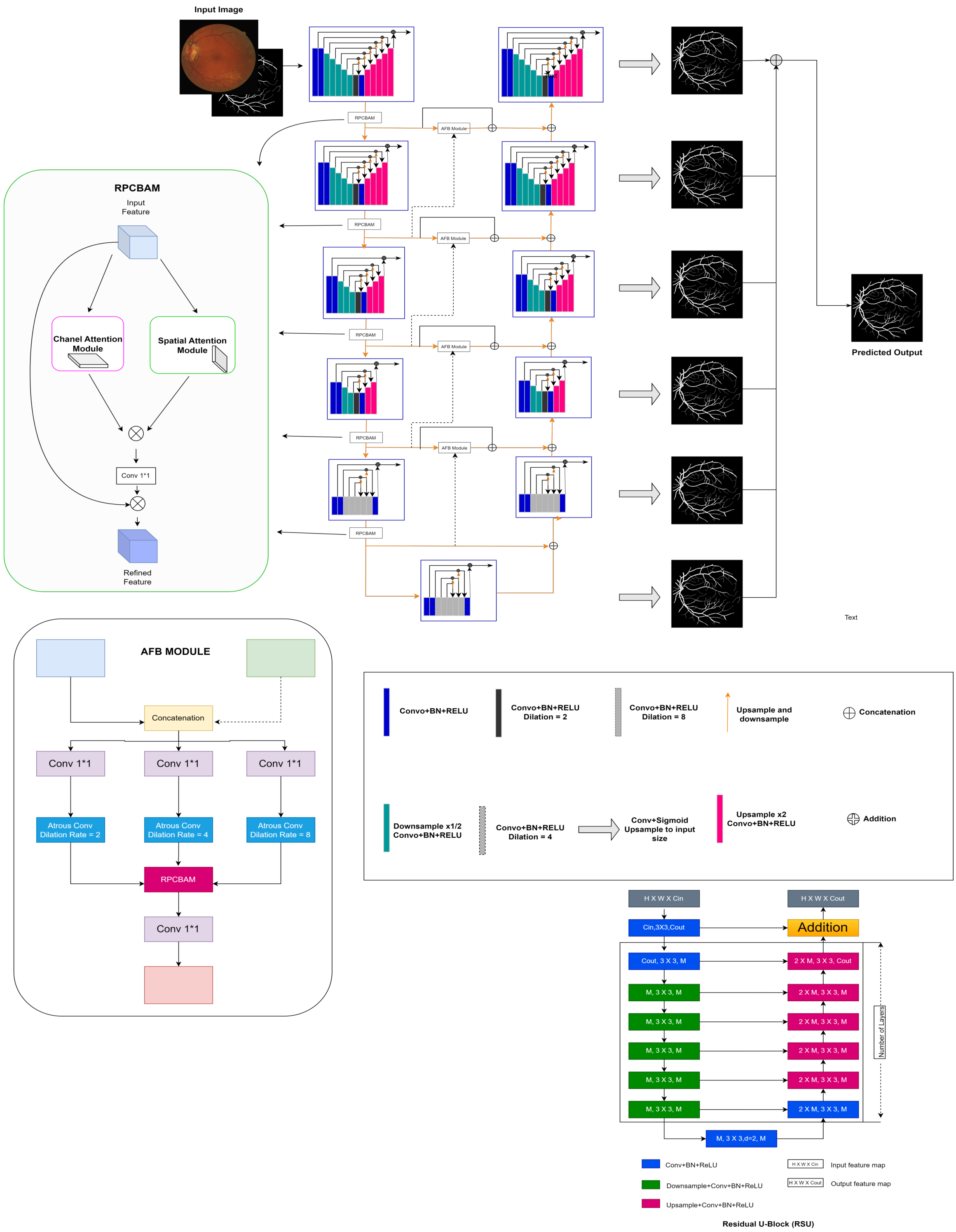


Figure 4.3: Architecture of our proposed model-2

Residual U-blocks

We used Residual U-Block to capture multi-scale features, and we got this idea from U-Net architecture. RSU block is itself a U-Net. But, to keep the number of trainable parameters less than enough, we used U-Net of a different number of layers. To extract local features we simply used convolutional layers. And then we used the number of layers, L, so that, as the outer Unet gets deeper, the number of layers, L gets smaller. Greater L results in a deeper residual U-block (RSU), more pooling activities, a wider variety of receptive fields, and richer local and global features. By setting this option, it is possible to extract multi-scale features from input feature maps with any desired level of spatial resolution. Progressive upsampling, concatenation, and convolution encode the multi-scale features into high-resolution feature maps from successively downsampled feature maps. The loss of fine features brought on by direct upsampling at large scales is reduced by this method. Then we added local feature maps and multi-scale feature maps using a residual connection.

Residual Parallel Convolutional Block Attention Module

Sanghyun Woo introduced the Convolutional Block Attention Module [9] for feature refinement. In feature maps, we introduce spatial attention blocks, so that the model can learn to focus more on the crucial area of the whole image, and channel attention blocks to address attention on different channels or different features so that the model can learn to focus more on the essential features as well. To retain the effects of both attention modules, we used the Residual Parallel Convolutional Block Attention Module (RPCBAM), which multiplies the image with the output features maps of both attention modules. We used the RPCBAM after every encoder block of the outer U-Net.

AFB Module

The U-Net architecture used skip connections to capture spatial information. But, this does not reduce the semantic gap between low-feature maps and high-feature maps. Yanhong Liu[32] introduced the AFB module in ResDO-UNet to mitigate the effect of the semantic gap. To achieve huge receptive fields while dealing with multi-scale feature expression, convolution layers with large kernel sizes are always introduced. Different receptive fields could be acquired in conjunction with the convolution layers with various kernel sizes to represent multi-scale features. However, convolution layers with big kernel sizes will result in high computation costs, which will reduce the effectiveness of the computation. With the help of a convolution layer with various expansion rates, the ASPP block offers a powerful method for expressing multiscale features, which may well acquire various receptive fields. They suggested an attention fusion block along with the benefits of the ASPP block. The attention fusion block receives input from the lower-level and higher-level feature maps that have been concatenated together to add more context information. In order to maintain the same size as the lower-level feature maps, bilinear interpolation is specifically performed to get higher-level feature maps. We added a novel block that contains atrous convolution of different dilation rates and then concatenated those feature maps and passed them through a parallel CBAM module. In atrous convolutional operations, different receptive fields are chosen. But selecting a big kernel increases the number of the trainable parameters by a large amount, So, in atrous convolution, the kernel is created in such a way that it can capture more spatial information without significantly increasing the total number of trainable parameters.We have used 256 X 256 image resolution and divided the whole dataset into training (80%), testing (10%), and validation(10%) datasets.

4.2.3 Proposed Model-3

Residual U-blocks

We used Residual U-Block to capture multi-scale features, and we got this idea from U-Net architecture. RSU block is itself a U-Net. But, to keep the number of trainable parameters less than enough, we used U-Net of a different number of layers. To extract local features we simply used convolutional layers. And then we used the number of layers, L, so that, as the outer Unet gets deeper, the number of layers, L gets smaller. Greater L results in a deeper residual U-block (RSU), more pooling activities, a wider variety of receptive fields, and richer local and global features. By setting this option, it is possible to extract multi-scale features from input feature maps with any desired level of spatial resolution. Progressive upsampling, concatenation, and convolution encode the multi-scale features into high-resolution feature maps from successively downsampled feature maps. The loss of fine features brought on by direct upsampling at large scales is reduced by this method. Then we added local and multi-scale feature maps using a residual connection.

Squeeze and Excitation Block (SE Block)

In this study, Jie Hu [8] et.al proposed a module that is used for increasing channel-wise interdependencies. It is a module used in Convolutional Neural Networks (CNN) to increase performance with a slight amount of computational cost. The Squeeze and Excitation module only applies attention to channel-wise information. They suggest a feature recalibration approach that enables the network to learn how to selectively emphasize informative traits and suppress less helpful ones by using global information. Inside this module, there has a squeeze module, excitation module, and scale module. The squeeze and Excitation block is a very lightweight module so there is a slight amount of computation cost.

Dense Convolutional Block Attention Module

Sanghyun Woo first introduced the Convolutional Block Attention Module [9] for feature refinement. In feature maps, we introduce spatial attention blocks, so that the model can learn to focus more on the vital area of the whole image, and channel attention blocks to address attention on different channels or different features so that the model can learn to focus more on the essential features as well. Gao Huang [33] et. al proposed a densely connected network to reduce the vanishing gradient problem and allow the reuse of

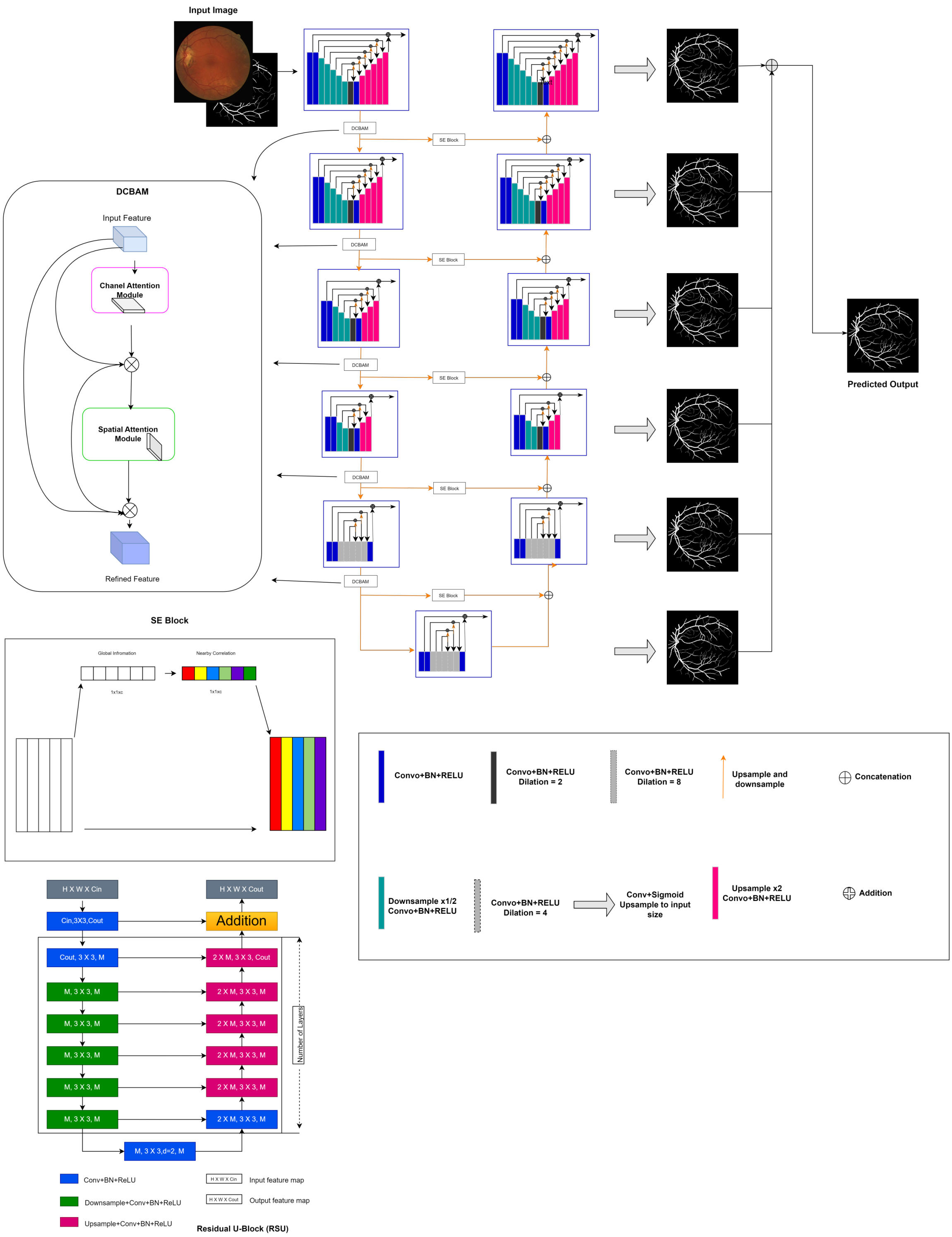


Figure 4.4: Architecture of our proposed model-3

features and strengthen feature propagation. We used this idea of the Densely Connected Network in the CBAM Module. Instead of using Resblock + CBAM just like the proposed module in the CBAM paper, we added more dense connections in the intermediate CBAM module to address the alleviation of the vanishing gradient problem and reuse of features, and improve the flow of feature propagation. In the dense CBAM module, we added a dense connection, where we take features from the previous convolution block of CBAM and add them to the feature maps that we get after the channel attention module. Then we take the output feature maps, resulting from dense connection, and add them to the feature maps that we get after the spatial attention module. Thus, applying two dense connections in the sequential CBAM increases the performance of the whole architecture.

4.3 Results and Discussion

In this section, we represent the result of the Datasets for our three proposed models. We will show how better the different datasets are performed for different.

4.3.1 Result Analysis for Proposed Model-1

In our proposed model-1 we have run five datasets which are DRIVE, ISBI-2012, CHASEDB-1, HRF, and Wound. Here DRIVE, CHASEDB-1, and HRF are the retinal datasets and ISBI-2012 is a microscopic image dataset and the Wound dataset is the human body wound dataset. The datasets performed really well in testing scores. But DRIVE, CHASEDB-1, and ISBI-2012 performed very well and gave very good results than so many existing segmentation models.

Table 4.2: **Result Analysis for Proposed Model-1**

Datasets	**Dice Coff**	**IoU**	**Recall**	**Precision**
DRIVE	87	77.06	58.72	98.55
ISBI-2012	98	95.91	85.89	99.17
CHASEDB-1	90.19	82.25	78.30	97.56
HRF	88.65	79.72	78.18	96.09
Wound	88.22	80.17	86.06	91.92

4.3.2 Result Analysis for Proposed Model-2

In our proposed model-2 we make some changes and try different module and block like Parallel CBAM and Attention Fusion Block in this architecture. We have run three datasets which are DRIVE,CHASEDB-1, and HRF. Here DRIVE, CHASEDB-1, and HRF all are the retinal datasets. As we use 256 x 256 image resolution here because of the lack of resources that's why this model gives a little bit lower result than the proposed model-1.But this result also gives well segmentation output for retinal images.

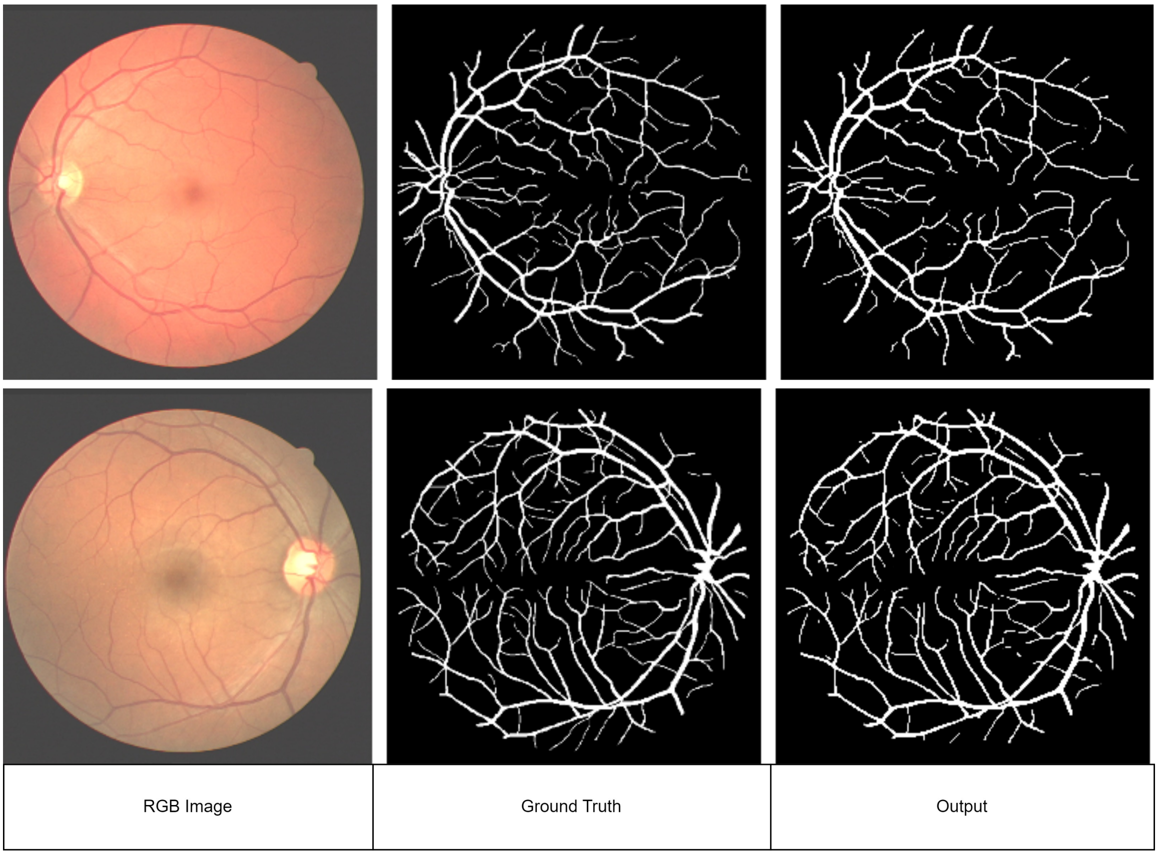


Figure 4.5: Output images of DRIVE dataset for Proposed model-1

Table 4.3: **Performance of Proposed Model-2**

Datasets	**Dice Coff**	**IoU**	**Recall**	**Precision**
DRIVE	84.45	73.21	55.56	97.47
CHASEDB-1	88	78.30	74.72	96.01
HRF	86.47	72.54	70.77	94.34

4.3.3 Result Analysis for Proposed Model-3

In our proposed model-3 we make some changes and introduced a different new attention module which is the Dense Convolutional Block Attention Module(Dense CBAM) in the encoder part and also the Squeeze and Excitation Block in the skip path of the main architecture. We have run four datasets which are DRIVE, ISBI-2012, CHASEDB-1, and HRF. Here DRIVE, CHASEDB-1, and HRF all are the retinal datasets and one is the microscopic image dataset which is ISBI-2012. As we use 256 x 256 image resolution here because of the lack of resources that why this model gives a little bit lower result than the proposed model-1 but give as good as performance like the proposed model-2.But this model also gives better segmentation output for multi domain images.

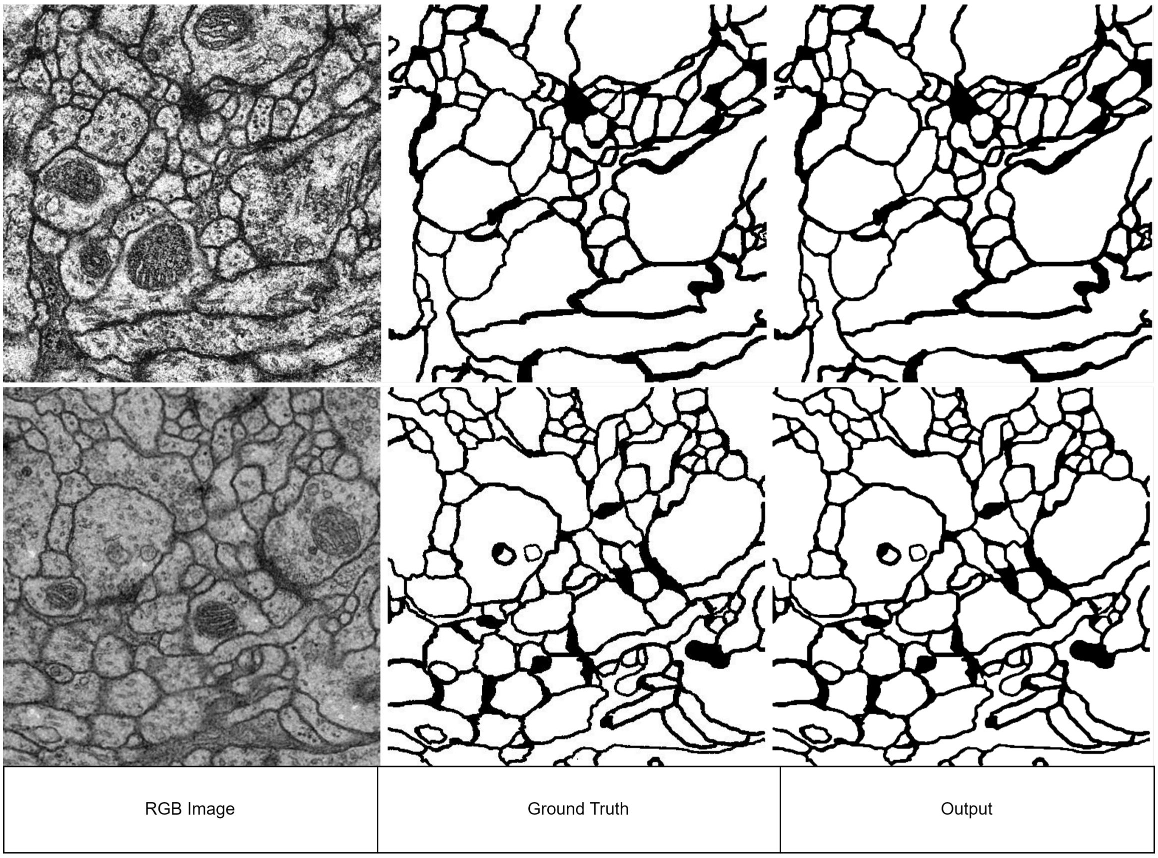


Figure 4.6: Output images of ISBI-2012 dataset for Proposed model-1

Table 4.4: **Performance of Proposed Model-3**

Datasets	**Dice Coff**	**IoU**	**Recall**	**Precision**
DRIVE	84.70	73.60	56.29	97.76
ISBI-2012	96.73	93.67	81.90	99.67
CHASEDB-1	88.34	78.54	75.25	96.56
HRF	86.73	72.86	71.34	94.91

4.4 Result Analysis and Discussion

This section shows our result from the analysis of our proposed three proposed models over other segmentation models. We report the result on five medical image datasets and compare them with other segmentation models. For all three models, the dataset preprocessing and train-test split ratio are reaming the same. We used 512x512 size images for the proposed model-1 and 256x256 size images for the proposed model-2 and model-3.

4.4.1 Result Analysis for DRIVE

Proposed Model-1:

On the drive dataset, our model outperforms the other segmentation architectures.

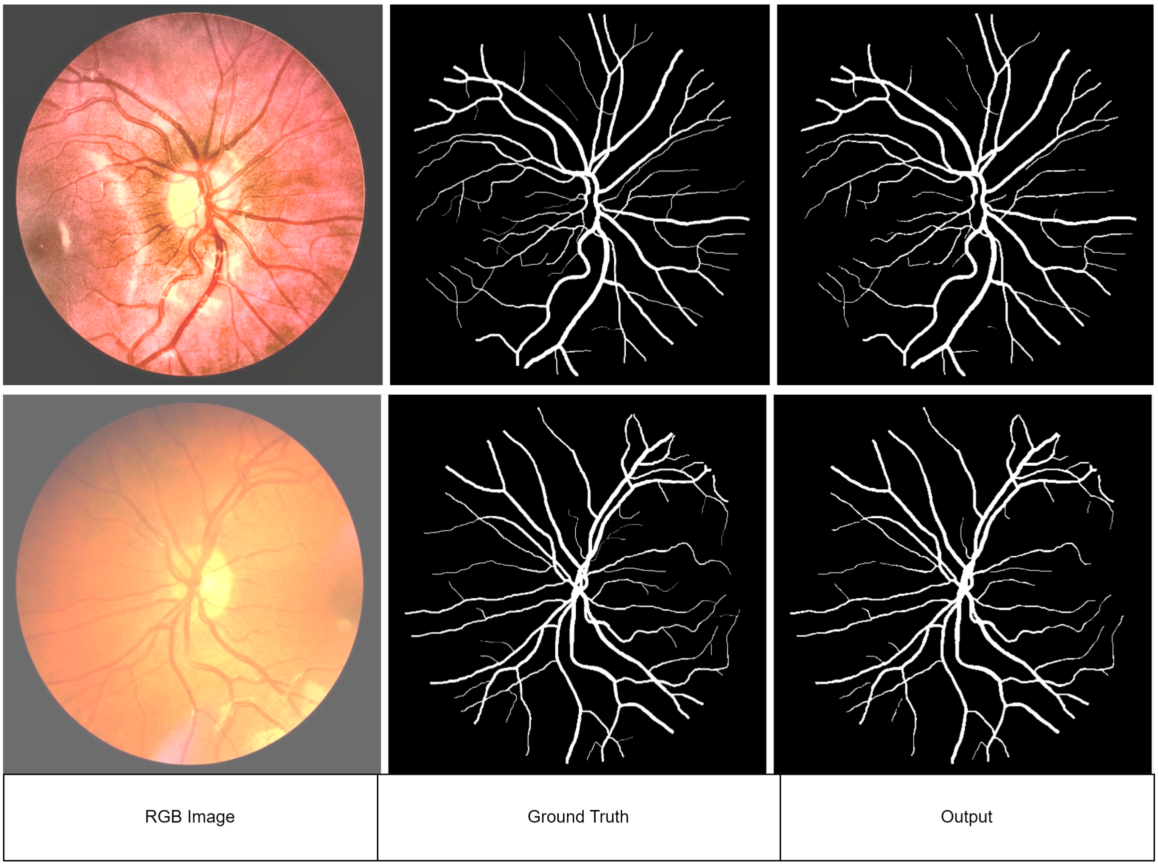


Figure 4.7: Output images of CHASEDB-1 dataset for Proposed model-1

With a dice score of 87%, IOU of 77.06%, Recall of 58.72%, and Precision of 98.55%. It is observed that our proposed model exceeds the performance of the recently published U2-net[7].

Proposed Model-2:

On the DRIVE dataset, our model outperforms the several segmentation architectures. With a dice score of 84.45%, IOU of 73.21%, Recall of 55.56%, and Precision of 97.47%. It is observed that our proposed model exceeds the performance of the recently published U2-net[7].

Proposed Model-3:

On the DRIVE dataset, our model outperforms the several segmentation architectures. With a dice score of 84.70%, IOU of 73.60%, Recall of 56.29%, and Precision of 97.76%. It is observed that our proposed model exceeds the performance of the recently published U2-net[7].

4.4.2 Result Analysis for ISBI-2012

Proposed Model-1:

In the ISBI 2012 datasets our first model achieved 98% of dice, 95.91% of IOU, 85.89% of Recall, and 99.17% of Precision which improve results compared to U-Net, U-Net++,

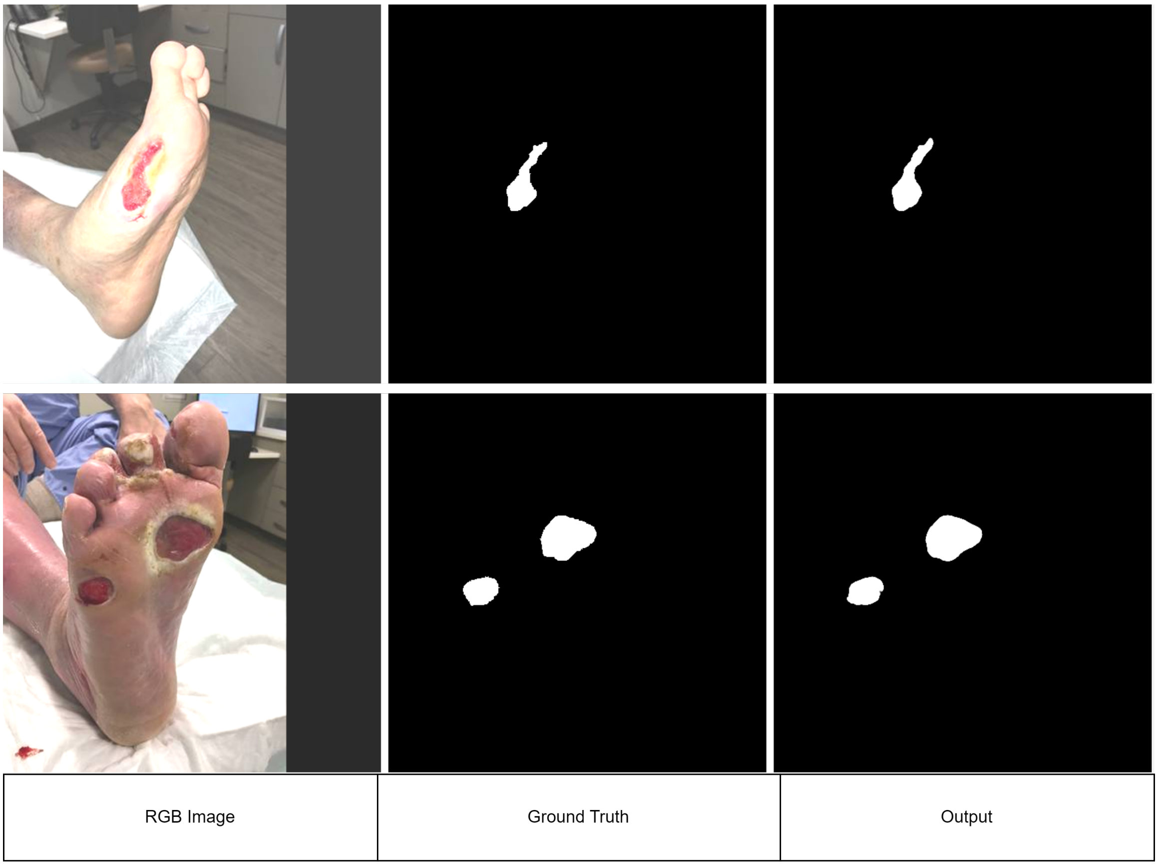


Figure 4.8: Output images of Wound dataset for Proposed model-1

Attention U-Net, MultiResU-Net.

Proposed Model-3:

In the ISBI 2012 datasets our first model achieved 96.73% of dice, 93.67% of IOU, 81.90% of Recall, and 99.67% of Precision which improve results compared to U-Net, U-Net++, Attention U-Net, MultiResU-Net.

4.4.3 Result Analysis for CHASEDB-1

Proposed Model-1:

In the CHASEDB-1 datasets, our first model achieved 90.19% of dice, 82.25% of IOU, 78.30% of Recall, 97.56% of Precision

Proposed Model-2:

In the CHASEDB-1 datasets, our second model achieved 88% of dice, 78.30% of IOU, 74.72% of Recall, and 96.01% of Precision.

Proposed Model-3:

In the CHASEDB-1 datasets, our second model achieved 88.34% of dice, 78.54% of IOU, 75.25% of Recall, and 96.56% of Precision.

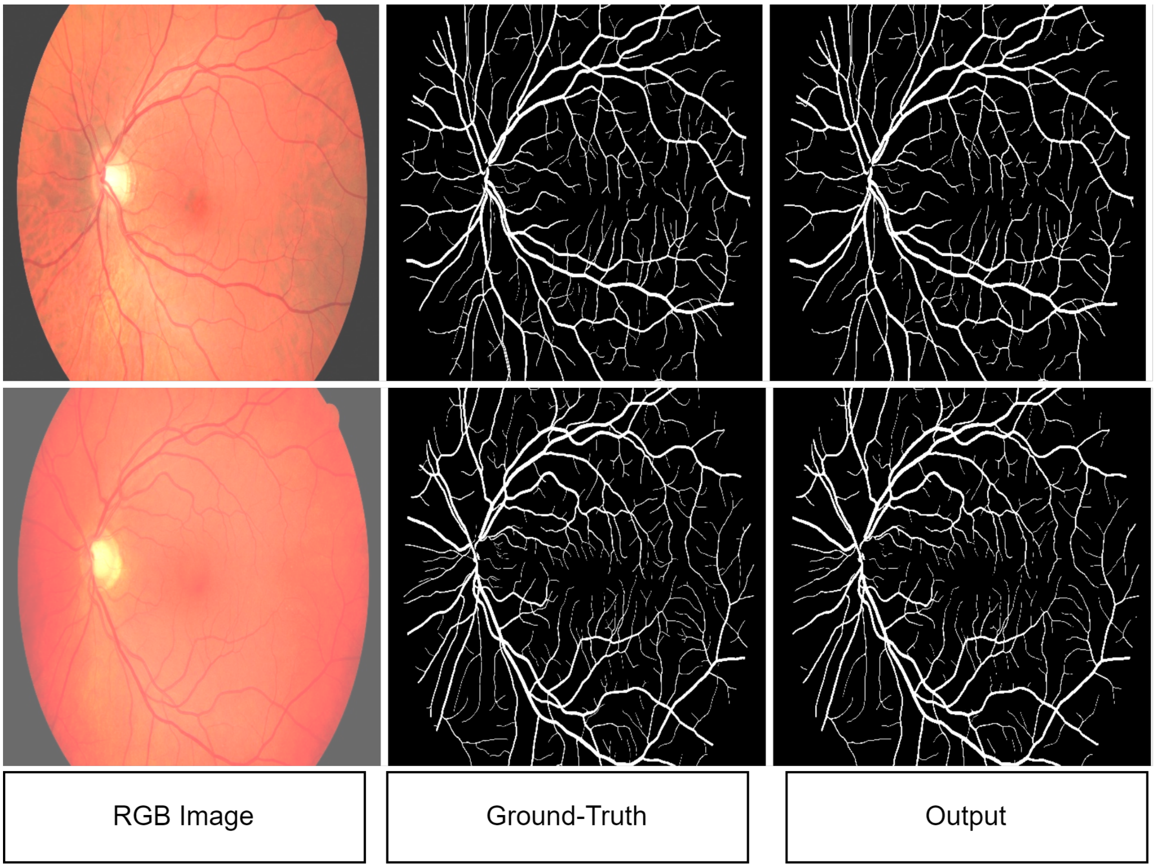


Figure 4.9: Output images of HRF dataset for Proposed model-1

4.4.4 Result Analysis for HRF

Proposed Model-1:

In the HRF dataset, our first model achieved 87.71% of dice, 77.10% of IOU, 74.18% of Recall, and 95.38% of Precision.

Proposed Model-2:

In the HRF dataset, our second model achieved 86.47% of dice, 72.54% of IOU, 70.77% of Recall, and 94.34% of Precision

Proposed Model-3:

In the HRF dataset, our second model achieved 86.73% of dice, 72.86% of IOU, 71.34% of Recall, and 94.91% of Precision

4.5 Summary

In this chapter, we discussed our data collection, data preprocessing, and testing of our model on multiple datasets. and at last, we discussed the performance of our model.

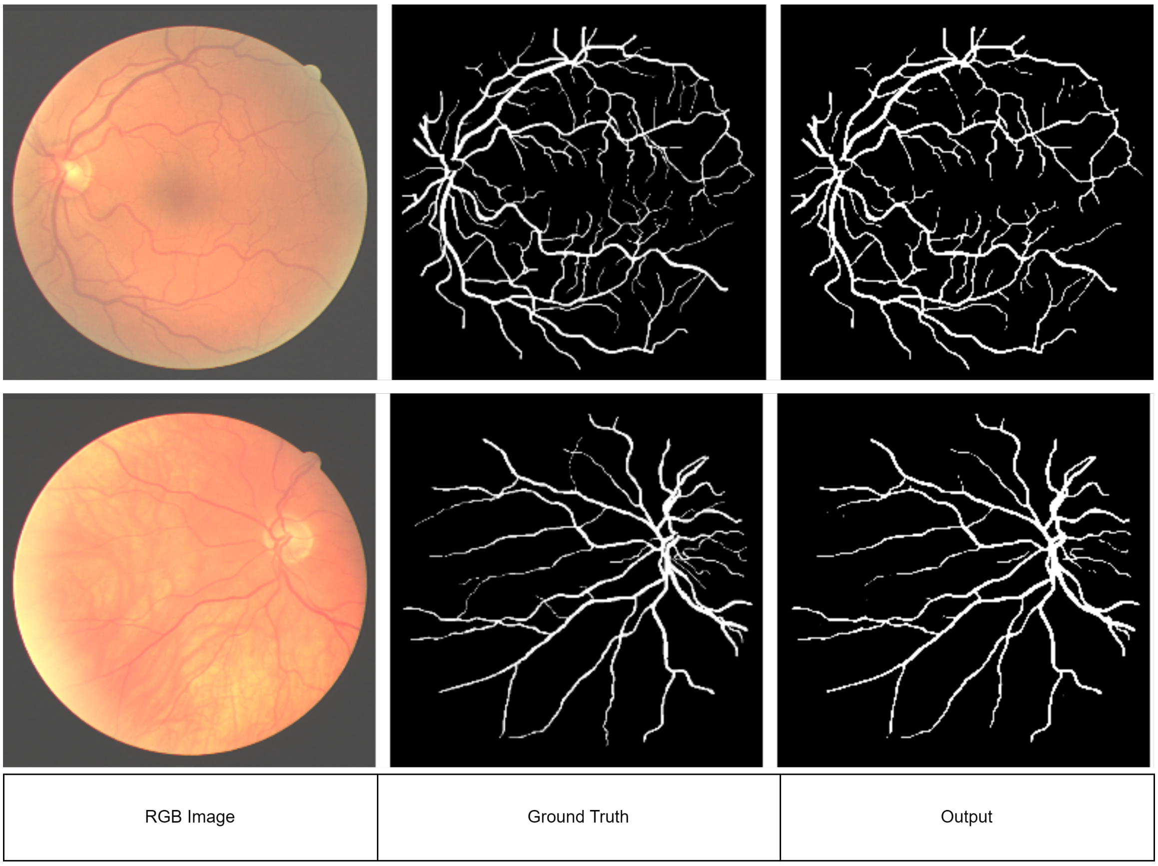


Figure 4.10: Output images of DRIVE dataset for Proposed model-2

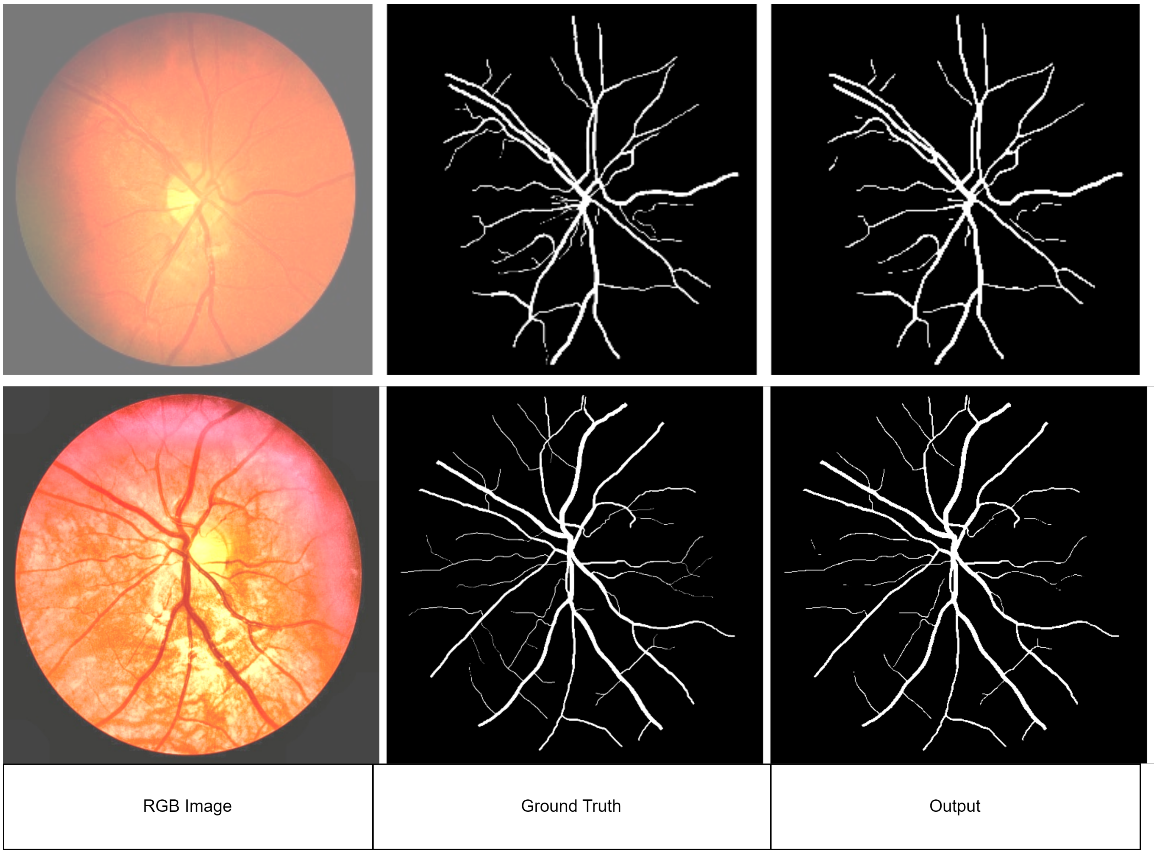


Figure 4.11: Output images of CHASEDB-1 dataset for Proposed model-2

Figure 4.12: Output images of HRF dataset for Proposed model-2

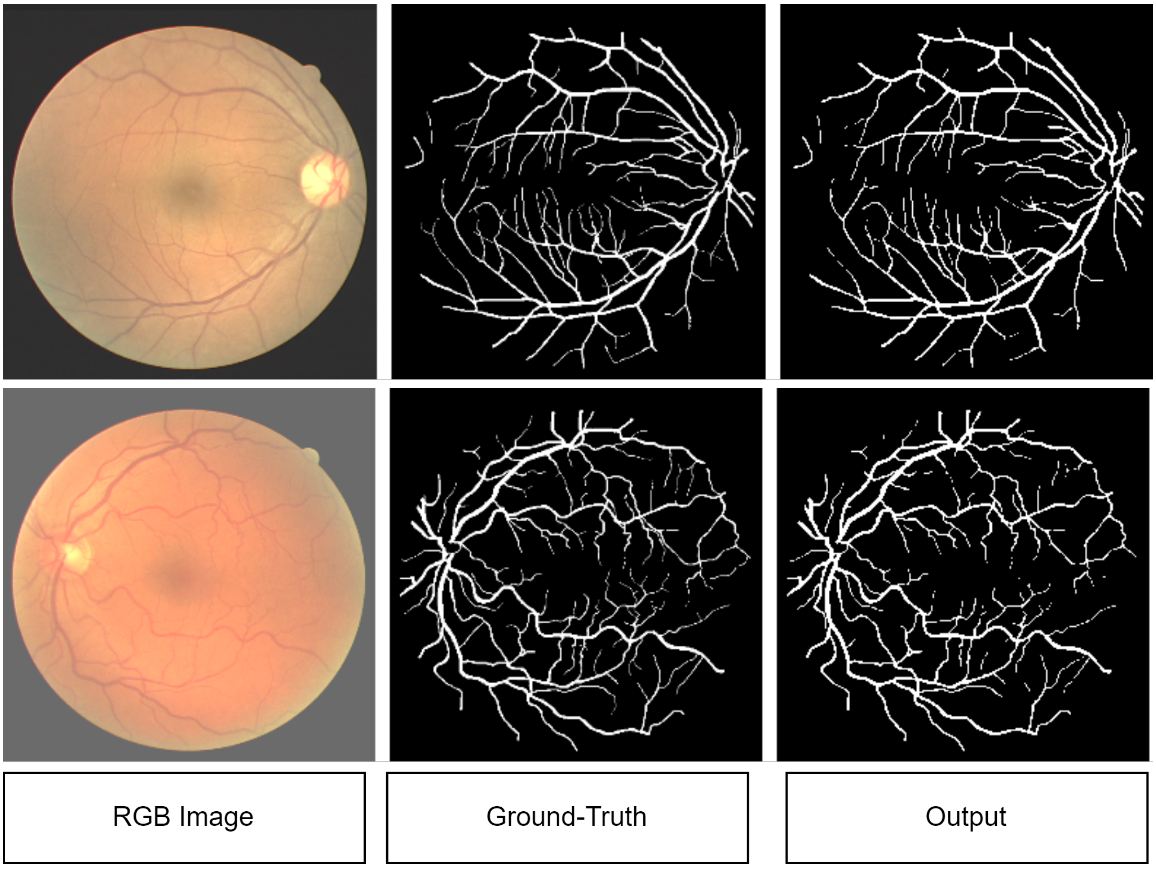


Figure 4.13: Output images of DRIVE dataset for Proposed model-3

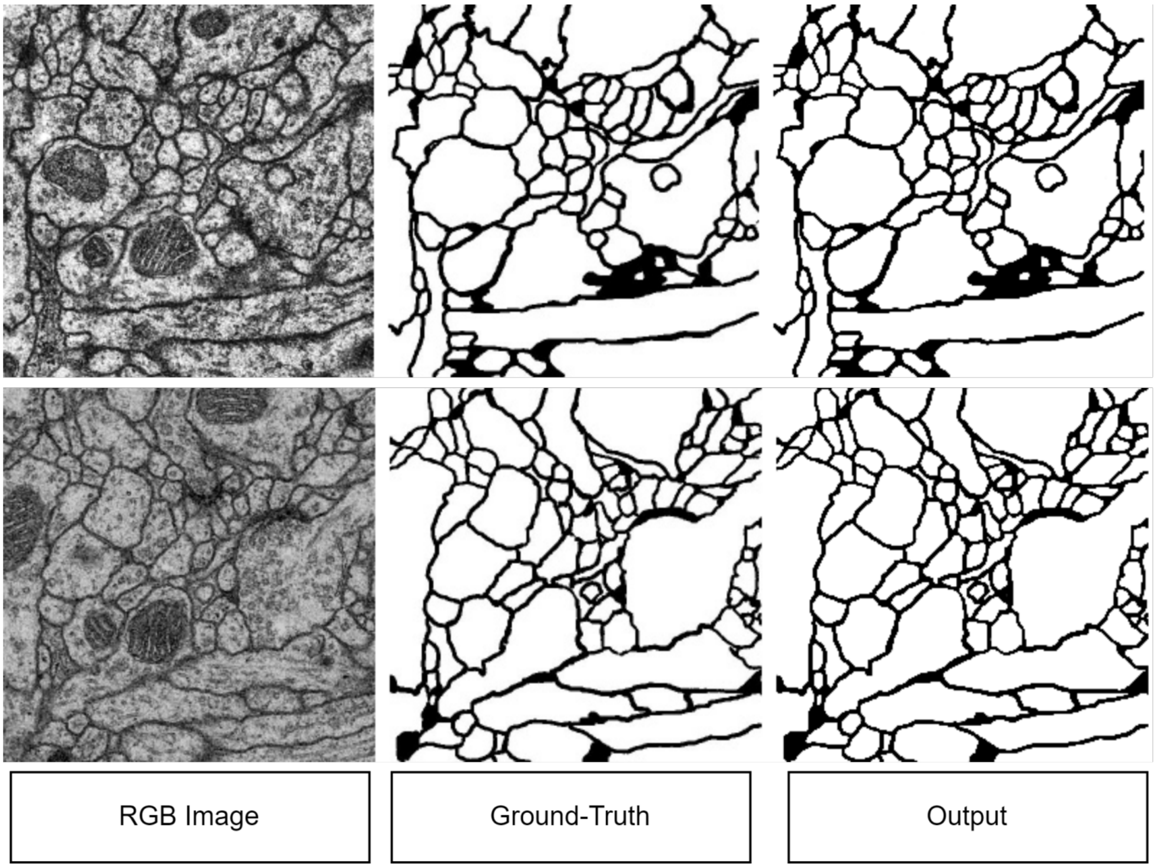


Figure 4.14: Output images of ISBI-2012 dataset for Proposed model-3

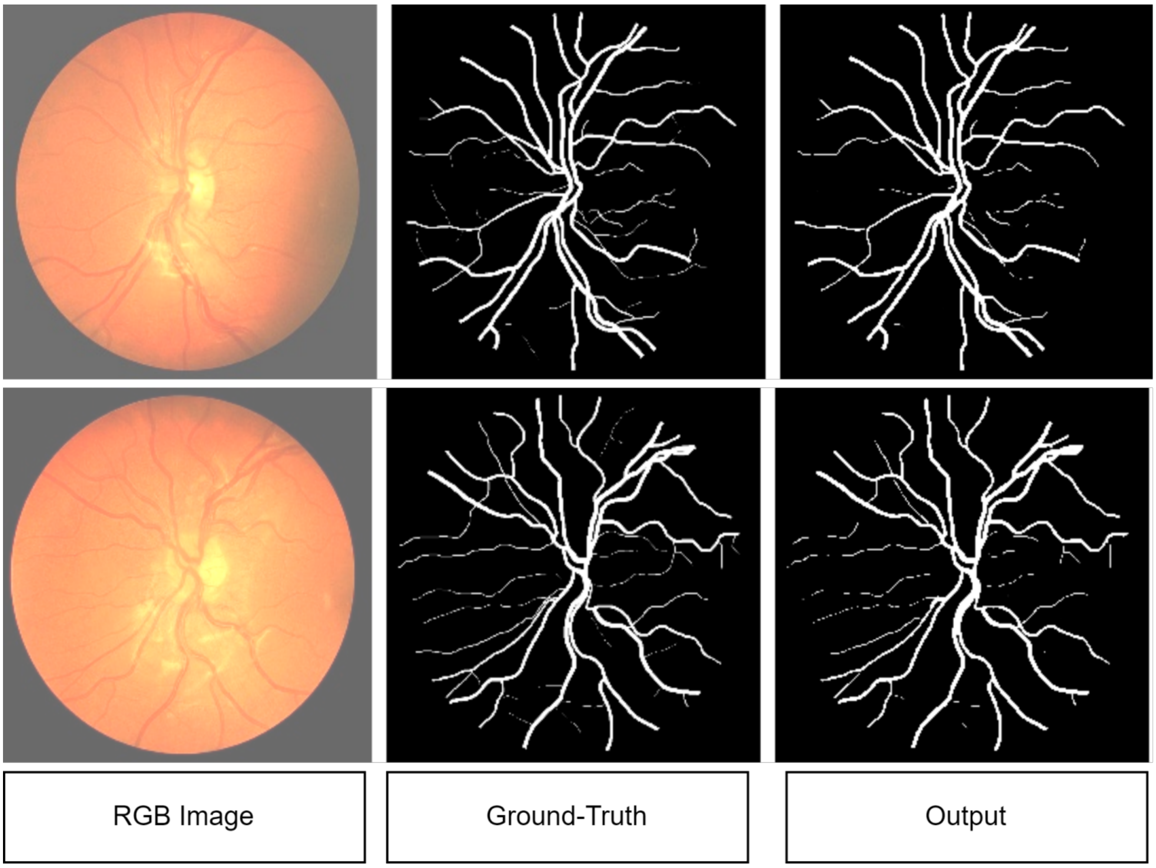
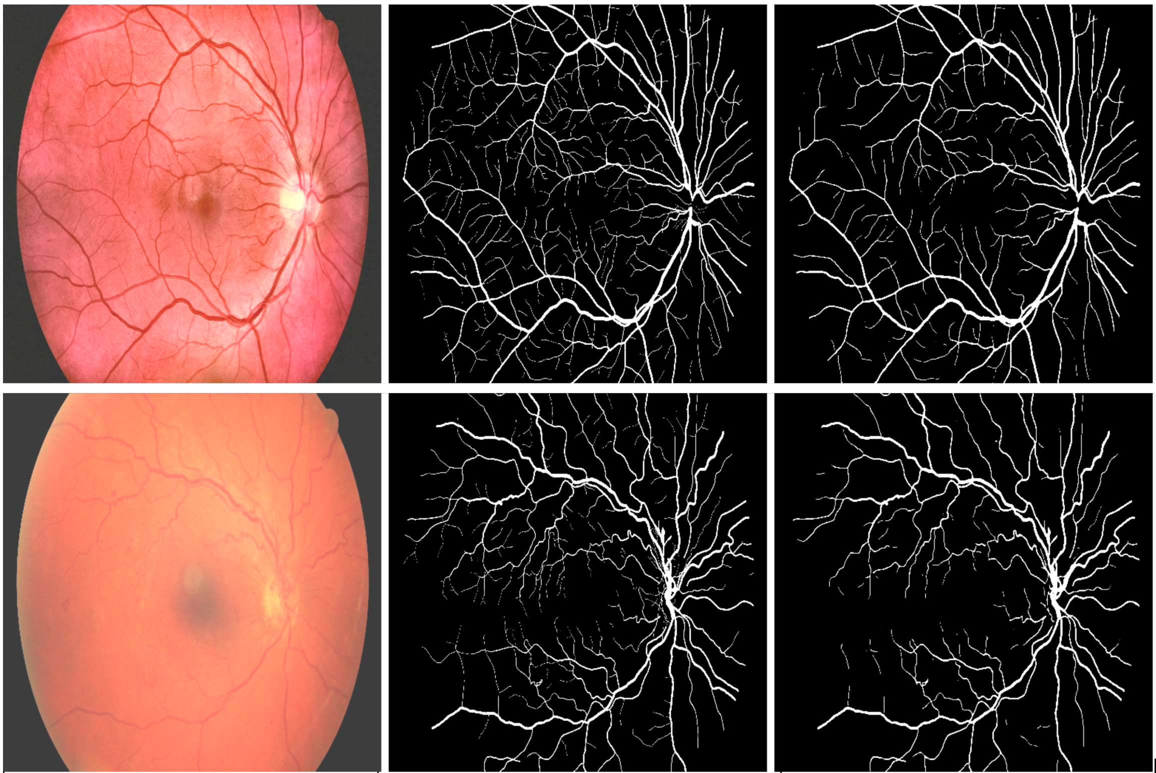


Figure 4.15: Output images of CHASEDB-1 dataset for Proposed model-3



RGB Image	Ground Truth	Output

Figure 4.16: Output images of HRF dataset for Proposed model-3

Chapter 5

Standards and Design Constraints

In this chapter, we, define the standards, design constraints, and complex engineering problems addressed in this project.

5.1 Compliance with the Standards

When we build a system or model we have to follow some software standards and hardware standards. Now we give some software and hardware standards for building our model.

5.1.1 Software Standard

There are some software standards we must have to follow. The software standards for our project:

- **Python (PEP8):** When writing python code it follows the best standards and rules;
- **ISO/IEC JTC 1/SC 42 - Artificial Intelligence:** For developing an AI model that follows these standards.
- **ISO/IEC/IEEE 26515:2018 - System and software engineering:** Application developing using Agile Methodolog's fundamental concepts.
- **ISO 9241-161:2016 (REVIEWED AND CONFIRMED IN 2021)-Guidance on visual user**-interface elements: This standard is used for developing user Interface.

5.1.2 Hardware Standard

- Intel Core i5-8400 CPU
- Colab Pro premium with 4GB GPU Support
- For System, Software, Hardware Verification and Validation we have to follow IEEE 1012-2016 Standards.

5.2 Design Constraints

Now we will describe the ethical issues and sustainability that are related to our project in this section.

5.2.1 Ethical Constraint

The ethical issue arrives when an activity conflict with society's principle. In our project, we are working with human data. So it Is a very sensitive issue. Basically, we are working with patients' data so we have to ensure the privacy of these patients. The collection and distribution of patient data for any purpose is the responsibility of the people and organizations who are involved with that. But sometimes it is not as simple as we think. Sometimes some complex scenario could happen. Thus, there is some accredited organization that can review the ethical issues related to a project. To distribution, collection, and use of human data for research purposes are controlled by two accredited sources that are European Union(EU) and National Level. They give some rules for each data provider so that ethical issues can handle easily.

5.2.2 Manufacturability and Cost Analysis

Now we will discuss the cost analysis for building our project. To implement this project, we need some basic elements which are:

- GPU (NVidia GTX 1650) - 20000 TK
- Colab pro(per month 1000 TK.total used 4 month .total cost = 4000)
- Google storage for dataset (300$)

Total Cost : Almost 54000 TK

5.2.3 Sustainability

In today's healthcare services, Sustainability is a very important area. We also have to ensure that we achieved some sustainability with our project. Our project directly helps in medical science. Like when doctors or clinicians are getting a clear view of the segmented biomedical images they can identify the problem and diseases of patients. It is directly related to the life of people worldwide. Because when doctors and clinicians are able to find a disease in the preliminary stage they take proper steps to prevent the disease. It saves the lives of so many patients.

5.3 Complex Engineering Problem

In this section, we will discuss why our project is a complex engineering problem and which engineering constraints are related to our project.

Table 5.1: Mapping with complex problem solving.

P1 Dept of Knowledge	P2 Range of Conflicting Requirements	P3 Depth of Analysis	P4 Familiarity of Issues	P5 Extent of Applicable Codes	P6 Extent of Stakeholder Involvement	P7 Interdependence
√	√	√			√	√

- **P1 Depth Knowledge :** In our engineering journey, we do so many courses. Actually in CSE by doing these courses we gradually increase our knowledge about mathematics, Data structure algorithm, machine learning algorithm and Artificial Algorithm and so many things which enrich our knowledge and increase our thinking capability which helps us to work with new problems and find the solution optimally our research we are actually working with Biomedical Image segmentation where we will find the gap of the existing algorithm and try to create an optimized algorithm which probably does better performance from already exist algorithm. So, for Image segmentation, we are actually working with Convolutional Neural Network (CNN) algorithm. So, we need deep knowledge about CNN. Also, we need to play with Matrix which was included in our Linear Algebra courses. Also, we have to work with some machine learning algorithms that we learned from Machine learning and AI Courses. For optimizing coding purposes we need Data structure and algorithm knowledge. These types of knowledge will help us to create a new model for Biomedical Image Segmentation.

- **P2 Range of Conflicting Requirements:** We have to check the conflict between pre-defined benchmark data vs real data.

- **P3 Depth of Knowledge:** We are studying some research papers and we also study some architecture of Biomedical Image Segmentation. So, we found some issues in some papers. In some papers, we find some time-consuming architecture that takes more time for the training dataset. So, if we will find a new architecture model which are reducing the time for training it will be a great contribution for us. Also, if we will reduce the cost (like reducing parameters) of architecture that is lightweight it will also help us. Another approach is if we will reduce the error rate and increase the accuracy of a model then it will be great work for us. If we are not able to increase the accuracy rate but if we increase the Dice value, and Jaccard value then the image segment the picture perfectly. Because in image segmentation Dice value and Jaccard value are more important than accuracy. Because more Dice value and Jaccard value segment images more clearly. If our project's Dice value and Jaccard value are more but accuracy becomes low it is also a contribution for us. So, we say that though there is a solution for Biomedical Image Segmentation we have a lot of sides to working in Biomedical Image Segmentation.

- **P6 Extent of stakeholder's involvement:** Doctors, Clinicians will be our stakeholders.

- **P7 Inter Dependence:** - The project has numerous independent components, such as data collection, data analysis, and insights, model training, and prediction module.

5.4 Summary

In this chapter, we have discussed the Standards and Design Constraints of our project. We talked about some software, and hardware standards that we have to follow for developing our project. Then we discussed some constraints like Ethical and sustainability for our project. We do our cost analysis which very important part of our project development. At the end of the chapter, we explain why our project is a complex engineering problem.

Chapter 6

Conclusion

To analyze biomedical images, segmentation of biomedical images are very important. So, it is necessary to segment images properly. Recently, a lot of models have been developed to segment biomedical images such as U-Net[34], U-Net++[35], and MultiRes-Unet[36]. We proposed a novel architecture based on U square 2-Net[7]and added more modules to improve the performance of the overall architecture. U square 2-Net was used to segment salient object detection. We used this architecture as the backbone of our architecture in biomedical image segmentation.

This architecture has a large number of trainable parameters that present one difficulty. We intend to scale back both the number of parameters and the processing complexity in the future. Additionally, we intend to modify the network's design to make it flexible in the 3D picture domain. In the future, we will use different transformers for better performance.

References

[1] Ozan Oktay, Enzo Ferrante, Konstantinos Kamnitsas, Mattias Heinrich, Wenjia Bai, Jose Caballero, Stuart A. Cook, Antonio de Marvao, Timothy Dawes, Declan P. O'Regan, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Anatomically constrained neural networks (acnns): Application to cardiac image enhancement and segmentation. *IEEE Transactions on Medical Imaging*, 37(2):384–395, 2018.

[2] Tim McInerney and Demetri Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1(2):91–108, 1996.

[3] Anne L. Martel, Steven J. Allder, Gota S. Delay, Paul S. Morgan, and Alan R. Moody. Measurement of infarct volume in stroke patients using adaptive segmentation of diffusion weighted mr images. In Chris Taylor and Alain Colchester, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI'99*, pages 22–31, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg.

[4] Itamar Arel, Derek C. Rose, and Thomas P. Karnowski. Deep machine learning - a new frontier in artificial intelligence research [research frontier]. *IEEE Computational Intelligence Magazine*, 5(4):13–18, 2010.

[5] S.-C.B. Lo, S.-L.A. Lou, Jyh-Shyan Lin, M.T. Freedman, M.V. Chien, and S.K. Mun. Artificial convolution neural network techniques and applications for lung nodule detection. *IEEE Transactions on Medical Imaging*, 14(4):711–718, 1995.

[6] Fabian Isensee, Paul F. Jaeger, Simon A. A. Kohl, Jens Petersen, and Klaus H. Maier-Hein. nnU-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2):203–211, dec 2020.

[7] Xuebin Qin, Zichen Zhang, Chenyang Huang, Masood Dehghan, Osmar R. Zaiane, and Martin Jagersand. U2-net: Going deeper with nested u-structure for salient object detection. *Pattern Recognition*, 106:107404, oct 2020.

[8] Jie Hu, Li Shen, Samuel Albanie, Gang Sun, and Enhua Wu. Squeeze-and-excitation networks, 2017.

[9] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module, 2018.

[10] Li Deng and Dong Yu. Deep learning: Methods and applications. *Found. Trends Signal Process.*, 7(3–4):197–387, jun 2014.

[11] Swarnendu Ghosh, Nibaran Das, Ishita Das, and Ujjwal Maulik. Understanding deep learning techniques for image segmentation. *ACM Comput. Surv.*, 52(4), aug 2019.

[12] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)*, pages 1–6, 2017.

[13] Volodymyr Mnih, Nicolas Heess, Alex Graves, and koray kavukcuoglu. Recurrent models of visual attention. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014.

[14] Liangliang Liu, Jianhong Cheng, Quan Quan, Fang-Xiang Wu, Yu-Ping Wang, and Jianxin Wang. A survey on u-shaped networks in medical image segmentations. *Neurocomputing*, 409:244–258, 2020.

[15] Qiangguo Jin, Zhaopeng Meng, Tuan D. Pham, Qi Chen, Leyi Wei, and Ran Su. DUNet: A deformable network for retinal vessel segmentation. *Knowledge-Based Systems*, 178:149–162, aug 2019.

[16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.

[17] Sirojbek Safarov and Taeg Keun Whangbo. A-denseunet: Adaptive densely connected unet for polyp segmentation in colonoscopy images with atrous convolution. *Sensors*, 21(4), 2021.

[18] Kun Yang, Shilong Chang, Zhaoxing Tian, Cong Gao, Yu Du, Xiongfeng Zhang, Kun Liu, Jie Meng, and Linyan Xue. Automatic polyp detection and segmentation using shuffle efficient channel attention network. *Alexandria Engineering Journal*, 61(1):917–926, 2022.

[19] Gendry Alfonso Francia, Carlos Pedraza, Marco Aceves, and Saúl Tovar-Arriaga. Chaining a u-net with a residual u-net for retinal blood vessels segmentation. *IEEE Access*, 8:38493–38500, 2020.

[20] Johnatan Carvalho Souza, João Otávio Bandeira Diniz, Jonnison Lima Ferreira, Giovanni Lucca França da Silva, Aristófanes Corrêa Silva, and Anselmo Cardoso de Paiva. An automatic method for lung segmentation and reconstruction in chest x-ray using deep neural networks. *Computer Methods and Programs in Biomedicine*, 177:285–296, 2019.

[21] Jeya Maria Jose Valanarasu and Vishal M. Patel. Unext: Mlp-based rapid medical image segmentation network, 2022.

[22] Zhenzhen Yang, Pengfei Xu, Yongpeng Yang, and Bing-Kun Bao. A densely connected network based on u-net for medical image segmentation. *ACM Trans. Multimedia Comput. Commun. Appl.*, 17(3), jul 2021.

[23] Xiwang Xie, Xipeng Pan, Weidong Zhang, and Jubai An. A context hierarchical integrated network for medical image segmentation. *Computers and Electrical Engineering*, 101:108029, 2022.

[24] Pengfei Song, Jinjiang Li, and Hui Fan. Attention based multi-scale parallel network for polyp segmentation. *Computers in Biology and Medicine*, 146:105476, 2022.

[25] Hongchun Lu, Shengwei Tian, Long Yu, Lu Liu, Junlong Cheng, Weidong Wu, Xiaojing Kang, and Dezhi Zhang. Dcacnet: Dual context aggregation and attention-guided cross deconvolution network for medical image segmentation. *Computer Methods and Programs in Biomedicine*, 214:106566, 2022.

[26] Zaiwang Gu, Jun Cheng, Huazhu Fu, Kang Zhou, Huaying Hao, Yitian Zhao, Tianyang Zhang, Shenghua Gao, and Jiang Liu. CE-net: Context encoder network for 2d medical image segmentation. *IEEE Transactions on Medical Imaging*, 38(10):2281–2292, oct 2019.

[27] Yunxiang Li, Shuai Wang, Jun Wang, Guodong Zeng, Wenjun Liu, Qianni Zhang, Qun Jin, and Yaqi Wang. GT u-net: A u-net like group transformer network for tooth root segmentation. In *Machine Learning in Medical Imaging*, pages 386–395. Springer International Publishing, 2021.

[28] Zabir Al Nazi and Tasnim Azad Abir. Automatic skin lesion segmentation and melanoma detection: Transfer learning approach with u-net and dcnn-svm. In Mohammad Shorif Uddin and Jagdish Chand Bansal, editors, *Proceedings of International Joint Conference on Computational Intelligence*, pages 371–381, Singapore, 2020. Springer Singapore.

[29] Maryam Asadi-Aghbolaghi, Reza Azad, Mahmood Fathy, and Sergio Escalera. Multilevel context gating of embedded collective knowledge for medical image segmentation, 2020.

[30] Chiu-Han Hsiao, Tzu-Lung Sun, Ping-Cherng Lin, Tsung-Yu Peng, Yu-Hsin Chen, Chieh-Yun Cheng, Feng-Jung Yang, Shao-Yu Yang, Chih-Horng Wu, Frank Yeong-Sung Lin, and Yennun Huang. A deep learning-based precision volume calculation approach for kidney and tumor segmentation on computed tomography images. *Computer Methods and Programs in Biomedicine*, 221:106861, 2022.

[31] Mobeen Ur Rehman, SeungBin Cho, Jeehong Kim, and Kil To Chong. Brainsegnet: Brain tumor mr image segmentation via enhanced encoder–decoder network. *Diagnostics*, 11(2), 2021.

[32] Yanhong Liu, Ji Shen, Lei Yang, Guibin Bian, and Hongnian Yu. Resdo-unet: A deep residual network for accurate retinal vessel segmentation from fundus images. *Biomedical Signal Processing and Control*, 79:104087, 2023.

[33] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. Densely connected convolutional networks, 2016.

[34] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

[35] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation, 2018.

[36] Nabil Ibtehaz and M. Sohel Rahman. Multiresunet : Rethinking the u-net architecture for multimodal biomedical image segmentation. *Neural Networks*, 121:74–87, 2020.