Segmentation of OCT Images for detection of DME(Diabetic Macular Edema)

Dyuti Dasmhapatra, Nichenametla Karthik Raja

Computer Science Department, BML Munjal University, Gurgaon, Haryana, India.

Abstract

Background and Objectives: OCT imaging in ophthalmology is vital for diagnosing conditions like Diabetic Macular Edema (DME). Precise segmentation of OCT images is crucial for accurate disease detection and monitoring. This study aims to employ advanced segmentation techniques, including U-Net and Mask RCNN, to enhance the detection of DME in OCT images. The goal is to improve diagnostic accuracy and contribute to effective treatment strategies for diabetic patients.**Material & Methods**: The segmentation of OCT(Optical coherence tomography) images and the detection of DME disease is proposed in this work. The report reviews related work, highlighting diverse methodologies, including machine learning and deep learning approaches, for DME detection.The proposed study introduces three distinct architectures: U-Net, basic transformation techniques, and Mask RCNN. The dataset, sourced from the Duke University database, underwent preprocessing involving blurring methods to enhance ground truth determination. Gaussian blurring, Laplacian filtering, and adaptive thresholding were applied for noise reduction and edge enhancement.The U-Net architecture, renowned for its segmentation capabilities, was employed with promising results. **Results:** The performance of the proposed system was tested using IOU score, Dice coefficient, F1 score and pixel accuracy. It was observed that the IOU score, Dice coefficient, F1 score and pixel accuracy of UNETl was much higher than Mask RCNN. It has been seen that the UNET performs better than Mask RCNN .

Keywords: OCT, Diabetic Macular Edema (DME), Segmentation, U-Net, Mask RCNN, Ophthalmology, Image Processing, Diagnostic Accuracy

1 Introduction

Optical Coherence Tomography (OCT) has emerged as a pivotal diagnostic tool in ophthalmology over the past two decades, undergoing substantial advancements in performance within the last ten years [1]. Various iterations of OCT systems have been introduced and refined during this period. Notably, Spectral-Domain OCT (SD-OCT), introduced in 2006,[2] has played a crucial role in enabling high-resolution and swift acquisition OCT systems. SD-OCT employs a broadband optical source and an interferometer, probed by a high-speed spectrometer, with detected spectra exhibiting modulations corresponding to the depth of features within the studied object. Achieving impressive scanning speeds exceeding 300,000 A-scans per second [3].SD-OCT is extensively utilized in ophthalmic applications for imaging both anterior and posterior eye segments, including the cornea, macula, optical nerve, and choroidal layer. With its capability for long-range imaging and clear visualization, SD-OCT has simplified the detection of conditions such as diabetic macular edema (DME) and AMD[4].Diabetic patients develop the DME(Diabetic macular edema) with the progression of DR(Diabetic retinopathy) .[4,5] DME occurs when fluid builds up in the macula because the barrier in the eye is not working right. This is a big reason why people with diabetic retinopathy lose their sight. If we find DME early, we can cure it and make it better. Sadly, many people in rural areas don't know about it, so they end up suffering from DME, and it can make them go blind forever.[6]Detecting Diabetic Macular Edema (DME) in diabetic patients has been tough. Therefore, it's crucial to help eye doctors in finding DME, which can make their diagnosis work better.In this study we proposed some methods such as U-net, mask RCNN and some image processing techniques.

2 Related Work

In [7], the author analyzed and performed the classification and segmentation for the detection of DME.The algorithm used HOG descriptors as features for the SVM classification. Cross validation was performed on the OCT dataset which contained the images of 18 people. Out of them 8 were normal, another 9 images were the patients of diabetic macular edema (DME). The classifier performed with an accuracy of 78.65%

In[8], the author proposed a transfer learning based approach in which the CNN of AlexNet was used to extract the features from preprocessed SD-OCT Images.The SVM classifier was used to classify images and an eight fold cross validation strategy was used to evaluate the proposed method.This method has achieved an accuracy of 96%

In[9], The authors proposed an algorithm in which CNN which was trained on medical images GoogLeNet, can be fine-tuned to detect DME in OCT images. It was also said that fine-tuning produces different results in different experiments, and this model has achieved the highest accuracy of 86% for its best model. The model achieved a maximal cross validation.

In[10] of their work, the researchers employed Convolutional Neural Network (CNN) segmentation techniques on a dataset comprising 289 OCT macular images. This segmentation methodology underwent rigorous cross-validation, iterating through a substantial 200,000 cycles. Notably, the model showcased remarkable performance, achieving a peak cross-validation Dice coefficient of 0.911..

The researchers proposed a network model to address the problem of small datasets, as well as to overcome the previous method's poor capacity to adapt to variations between datasets, with a new concept of the reusable feature. The ability to reuse features boosts the neural network's performance, suggesting that data amplification, data augmentation, and generative adversarial networks (GAN) can be examined as well[11]

In the [12] of their research, the authors introduced a novel approach. They proposed a method involving the extraction of features both before and after the fully connected layers of a pre-trained VGG16 architecture. These extracted features were then subjected to classification using both K-Nearest Neighbors (KNN) and a random forest classifier. Impressively, the model achieved a high accuracy rate of 87.5%. This innovative methodology showcases the effectiveness

of leveraging pre-trained neural network architectures for feature extraction and subsequent classification tasks.

In their 2019 study, the authors directed their efforts toward the segmentation of Diabetic Macular Edema (DME) disease.[13] The process involved a comprehensive approach, including denoising, segmentation of intraretinal layers, and the removal of unwanted elements. Notably, the authors employed AdaBoost for seed and region extraction, complemented by a graph cut algorithm for precise fine segmentation. The evaluation, conducted with 3D OCT images, yielded promising results for DME volume segmentation, with metrics indicating 84.6% True Positive Volumetric Fraction (TPVF), 1.7% False Positive Volumetric Fraction (FPVF), and an overall accuracy of 91.7%. This method showcases a robust strategy for effectively delineating DME in OCT images.

In this study[14], a deep learning algorithm aimed to automate segmentation and quantification of macular fluids in exudative macular diseases, including AMD, DME, and RVO, was introduced. The algorithm, seamlessly integrated with two OCT devices, demonstrated remarkable performance with mean AUROC values of 0.94 and 0.92 for detecting IRC and SRF. These results highlight its effectiveness in accurately identifying and quantifying features associated with exudative macular diseases. Additionally, the study explored the U-net deep learning algorithm, revealing its significant proficiency with an impressive 86.8% accuracy in segmenting arteries and veins in OCT images for DME. This dual-pronged approach showcases the versatility of deep learning in ophthalmology and suggests enhanced diagnostic capabilities for macular diseases.

In this study [15], a novel method was introduced for the identification of drusen shapes within the Retinal Pigment Epithelium (RPE) layer in Optical Coherence Tomography (OCT) images. Leveraging this information, the researchers precisely located the Retinal Fluid Layer (RFL) and successfully detected blood bubbles. The study then utilized binary classification to effectively distinguish between Age-Related Macular Degeneration (AMD) and Diabetic Macular Edema (DME) diseases. Despite working with a small dataset comprising 16 OCT images (10 AMD, 6 DME), the method exhibited noteworthy performance, achieving an accuracy of 87.5% in accurately classifying the diseases. This research represents a valuable contribution to the field, showcasing the potential for precise identification and classification of pathological features in OCT imagery.

In this study of [16] Retinal image analysis, centered on Optical Coherence Tomography (OCT), accelerates the diagnosis of retinal disorders, notably targeting AMD and DME. Despite OCT's pivotal role, existing methods lack an end-to-end

system. This study introduces a GAN-based solution, seamlessly integrating Gabor features for precise segmentation and non-local denoising for effective noise reduction. The outcomes showcase an impressive 92.42% accuracy and a robust F1-score of 0.79, highlighting GANs' effectiveness in reshaping OCT-based retinal image analysis. This research underscores advanced machine learning's transformative potential in medical imaging, promising a more efficient and reliable era for diagnosing ophthalmic conditions.

2 Methodology

2.1 Data-Set

OCT(Optical coherence tomography) images and dataset[17] were downloaded from kaggle which has images in .mat format(matlab).The dataset has a collection of 10 OCT images of the patients taken from the duke university database.Our main objective is to detect the DME in the following images.

2.2 Pre-Processing

Segmentation of OCT images and their images were collected from different sources. Thus, in the preprocessing we had used some methods of blurring.In this process we were about to find the ground truth of the image that we wanted to find out.

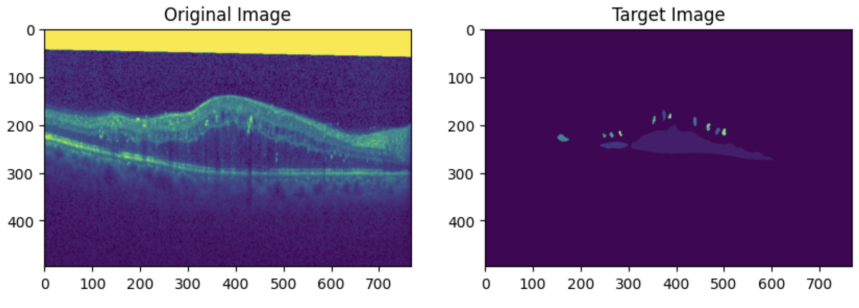


Fig 1: a) Original Image b) Image which has ground truth

Gaussian blurring was implemented inorder to reduce the noise or some details in an image and sharpening enhances the edges and fine details in an image this will be achieved by using the high frequency components mostly focussed on the edges which creates some changes in the intensity.

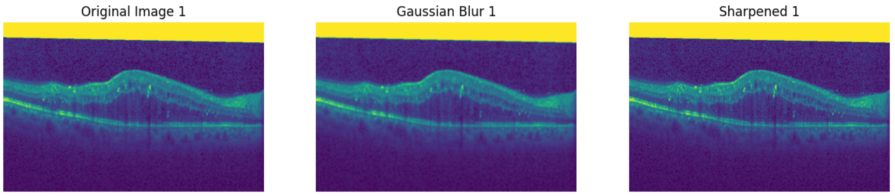


Fig 2 : Representation of the image using gaussian blurring

2.3 Model Architecture

We propose 3 different architectures: UNet, basic transformation techniques and MaskRCNN(Mask Recurrent neural networks) to detect the DME(diabetic

U-NET Architecture:

U-Net is an image segmentation deep learning algorithm which got its name from its U-like architecture which helps in image downscaling and upscaling to get the desired output. the classic U-net compromises of 2 convolutional blocks in each layer along with max pooling layer and Relu Activation layer.the first half of the U-net architecture extracts features and second half of the U-net borrows features from first half and output of flatten layer to form the desired image.

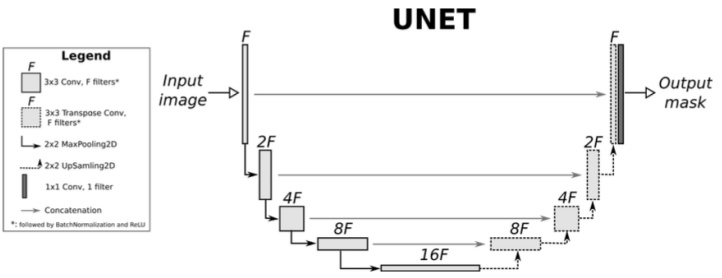
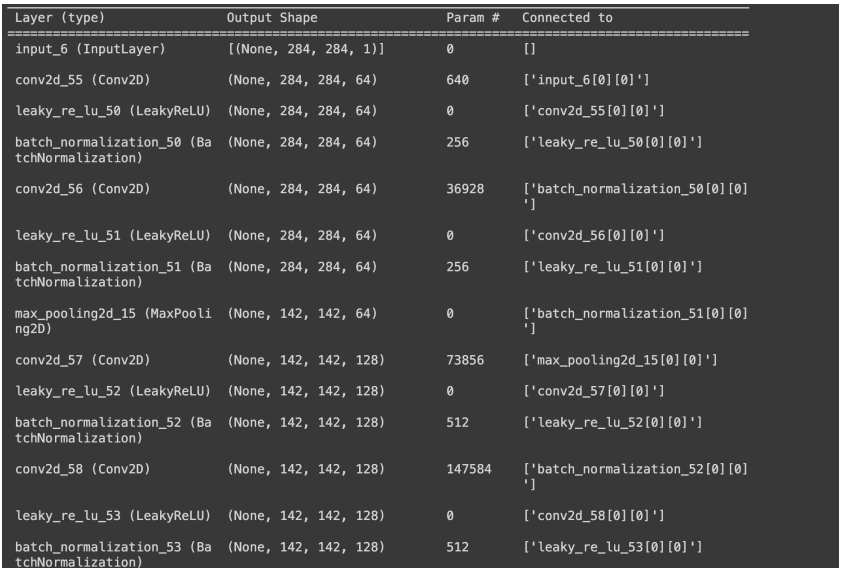
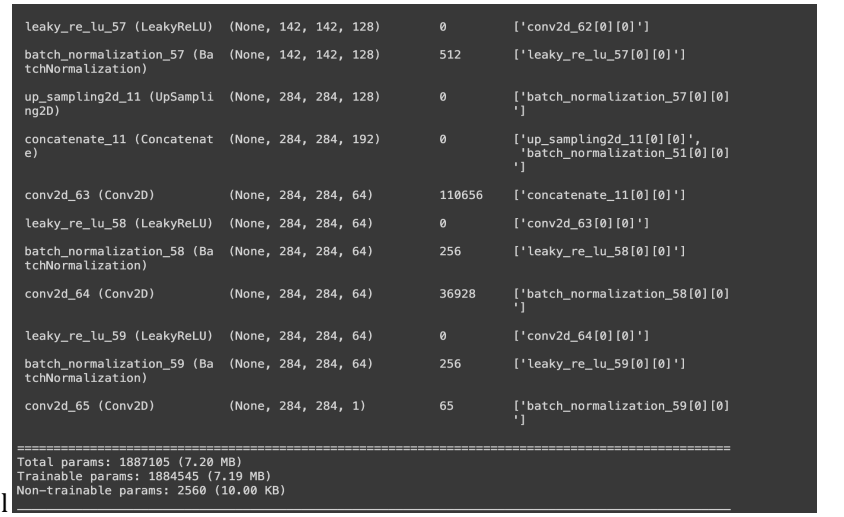


Fig 3: Figure representing the proposed U-Net architecture model



Layer (type)	Output Shape	Param #	Connected to
input_6 (InputLayer)	[(None, 284, 284, 1)]	0	[]
conv2d_55 (Conv2D)	(None, 284, 284, 64)	640	['input_6[0][0]']
leaky_re_lu_50 (LeakyReLU)	(None, 284, 284, 64)	0	['conv2d_55[0][0]']
batch_normalization_50 (BatchNormalization)	(None, 284, 284, 64)	256	['leaky_re_lu_50[0][0]']
conv2d_56 (Conv2D)	(None, 284, 284, 64)	36928	['batch_normalization_50[0][0]']
leaky_re_lu_51 (LeakyReLU)	(None, 284, 284, 64)	0	['conv2d_56[0][0]']
batch_normalization_51 (BatchNormalization)	(None, 284, 284, 64)	256	['leaky_re_lu_51[0][0]']
max_pooling2d_15 (MaxPooling2D)	(None, 142, 142, 64)	0	['batch_normalization_51[0][0]']
conv2d_57 (Conv2D)	(None, 142, 142, 128)	73856	['max_pooling2d_15[0][0]']
leaky_re_lu_52 (LeakyReLU)	(None, 142, 142, 128)	0	['conv2d_57[0][0]']
batch_normalization_52 (BatchNormalization)	(None, 142, 142, 128)	512	['leaky_re_lu_52[0][0]']
conv2d_58 (Conv2D)	(None, 142, 142, 128)	147584	['batch_normalization_52[0][0]']
leaky_re_lu_53 (LeakyReLU)	(None, 142, 142, 128)	0	['conv2d_58[0][0]']
batch_normalization_53 (BatchNormalization)	(None, 142, 142, 128)	512	['leaky_re_lu_53[0][0]']

Fig 4: Representing the model summary of the UNET model



leaky_re_lu_57 (LeakyReLU)	(None, 142, 142, 128)	0	['conv2d_62[0][0]']
batch_normalization_57 (BatchNormalization)	(None, 142, 142, 128)	512	['leaky_re_lu_57[0][0]']
up_sampling2d_11 (UpSampling2D)	(None, 284, 284, 128)	0	['batch_normalization_57[0][0]']
concatenate_11 (Concatenate)	(None, 284, 284, 192)	0	['up_sampling2d_11[0][0]', 'batch_normalization_51[0][0]']
conv2d_63 (Conv2D)	(None, 284, 284, 64)	110656	['concatenate_11[0][0]']
leaky_re_lu_58 (LeakyReLU)	(None, 284, 284, 64)	0	['conv2d_63[0][0]']
batch_normalization_58 (BatchNormalization)	(None, 284, 284, 64)	256	['leaky_re_lu_58[0][0]']
conv2d_64 (Conv2D)	(None, 284, 284, 64)	36928	['batch_normalization_58[0][0]']
leaky_re_lu_59 (LeakyReLU)	(None, 284, 284, 64)	0	['conv2d_64[0][0]']
batch_normalization_59 (BatchNormalization)	(None, 284, 284, 64)	256	['leaky_re_lu_59[0][0]']
conv2d_65 (Conv2D)	(None, 284, 284, 1)	65	['batch_normalization_59[0][0]']

Total params: 1887105 (7.20 MB)
Trainable params: 1884545 (7.19 MB)
Non-trainable params: 2560 (10.00 KB)

l

The table represents a neural network model with a U-Net architecture for image segmentation. It takes a grayscale image of size (284, 284, 1) as input and outputs a segmentation mask of the same size. The model consists of convolutional layers with leaky ReLU activation and batch normalization. Max pooling and up-sampling layers are used for spatial down-sampling and

up-sampling, respectively. Skip connections via concatenation help preserve spatial information. The final convolutional layer outputs a single-channel mask. The model has a total of 1,887,105 parameters, with 1,884,545 trainable parameters and 2,560 non-trainable parameters, indicating moderate model complexity.

Mask - RCNN architecture :

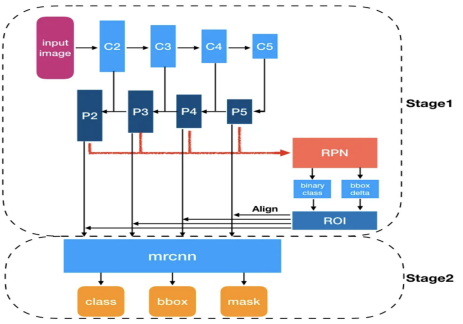


Illustration of Mask RCNN structure

Mask RCNN is a two-stage instance segmentation neural network. In the first stage, a Region Proposal Network (RPN) scans a Feature Pyramid Network (FPN) to propose regions using predefined anchors. The anchors help associate features with raw image locations. In the second stage, another network refines proposals using a technique called ROIAlign, predicts object classes, bounding boxes, and generates pixel-level masks. The FPN backbone extracts features with different scales, enhancing the network's ability to capture object details at various resolutions. Anchors guide proposal generation, while ROIAlign assists in accurate feature localization without anchors, making Mask RCNN effective for object instance separation.

Model summary of the R CNN architecture:

Transform Parameters	
Normalize	Mean=[0.485, 0.456, 0.406], Std=[0.229, 0.224, 0.225]
Resize	min_size=(800,), max_size=1333, mode='bilinear'
Backbone Parameters	
Backbone type	ResNEt with bottleneck architecture

No of layers	4
No of blocks per layer	[3,4,6,3]
FPN Parameters	
Inner Blocks	1x1 convolutions for feature map scaling
Layer Blocks	3x3 convolutions for feature map processing
Extra Blocks	LastLevelMaxPool for additional feature map scaling
ROI Heads Parameters	
Box ROI Pool	MultisclaeROIAlign with output size(7,7) and sampling ratio 2
Box Head	Two layer MLP with output features 1024
Box predictor	Linear layer with 64 output features for bounding box prediction
Mask Head	Four 3x3 convolutions with RELU activation
Mask predictor	2x2 transposed convolution layer with 128 output channels,Relu activation and 1x1 convolutions layer with 16 output channels for mask prediction

3 Results and evaluation:

We have used k fold cross validation for the evaluation of the accuracy.The performance of the segmented oct images are tested through multiple evaluation metrics such as IOU score, Dice coefficient, f1 score and the pixel accuracy of the model.

$$Dice = \frac{2 * area\,(S_1 \cap S_2)}{area\,(S_1) + area\,(S_2)}$$

$$Iou = \frac{area\,(S_1 \cap S_2)}{area\,(S_1 \cup S_2)}$$

$$R = \frac{TP}{TP + FN}$$

$$P = \frac{TP}{TP + FP}$$

S1 = represents the region of the edema

S2 = represents the region of the non - edema(which is not edema)

TP - true number of pixels for the prediction of edema

FN - The number of pixels with true edemas predicted as non-edemas.

FP - The number of pixels where non-edemas are predicted to be edemas.

3.1 U-Net Architecture:

The model is trained with U-Net with 380 epochs of training and testing the dataset.In this method we have used the input images.Table 1 shows the various performance of U-net architecture.

IOU score	Dice coefficient	F1 Score	Accuracy of the pixel
0.6121	0.7593	0.75939	0.9814

The metrics for evaluating a segmentation model are as follows: IOU score (0.1503), Dice coefficient (0.2077), F1 Score (0.1631), and pixel accuracy (0.9958). These metrics provide insights into the model's performance in terms of overlapping predictions, overall segmentation accuracy, and pixel-level agreement with ground truth.

3.2 Mask RCNN

The model is trained with mask rcnn with 500 epochs of training and testing the dataset.In this method we have used the input images.Table 2 shows the various performance mask rcnn architecture.

IOU score	Dice coefficient	F1 Score	Accuracy of the pixel
0.1503	0.2077	0.1631	0.9958

The metrics for evaluating a segmentation model are as follows: IOU score (0.1503), Dice coefficient (0.2077), F1 Score (0.1631), and pixel accuracy (0.9958). These metrics provide insights into the model's performance in terms of overlapping predictions, overall segmentation accuracy, and pixel-level agreement with ground truth.

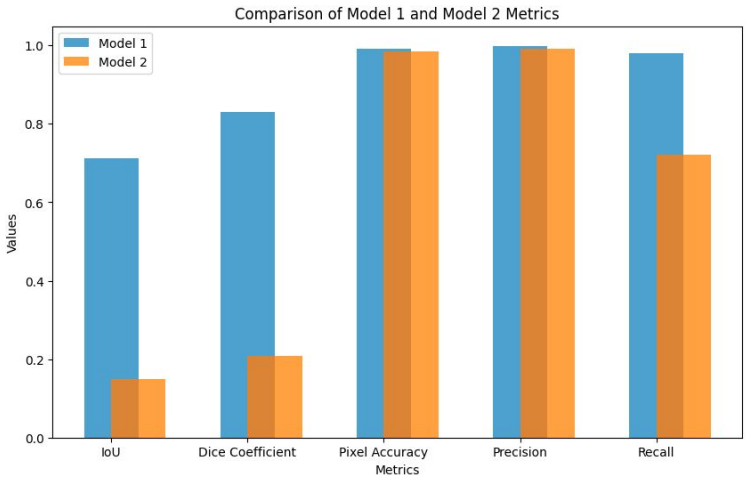


Fig 5 : Representing the evaluation metrics used for the project

The above figure represents the respective scores achieved by the model according to the evaluation metrics.

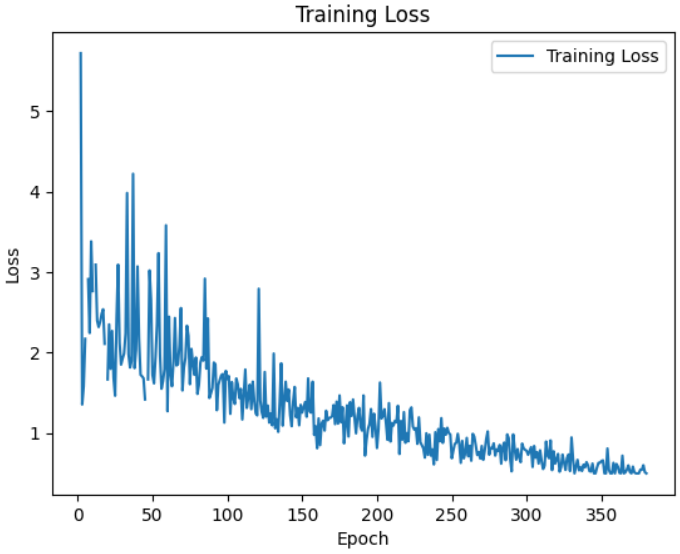


Fig 6 : The figure representing the training loss vs epochs using U-Net

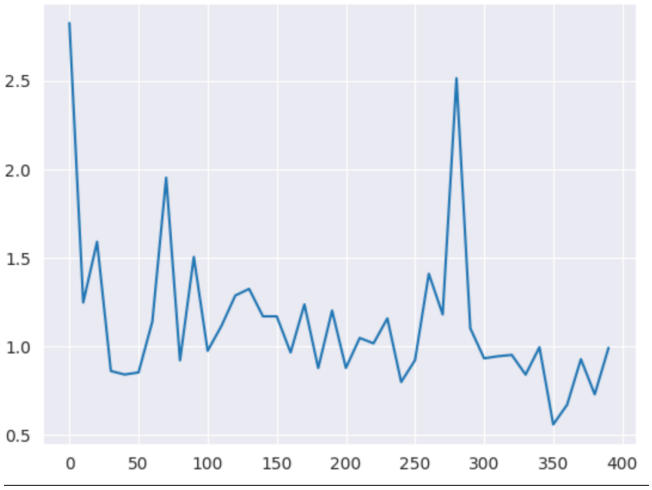


Fig : The figure representing the loss vs epochs using mask RCNN

Segmented and target image using U net architecture:

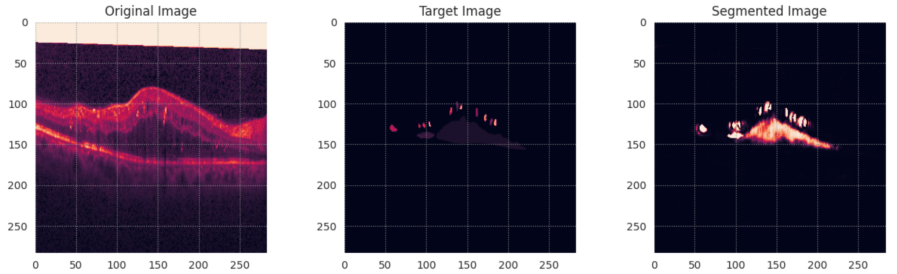


Fig 7: Representing the segmented and target image with Unet

In the segmented image figure it shows the detection of DME(diabetic macular edema) using U-Net architecture.

Segmented and target image using Mask RCNN:

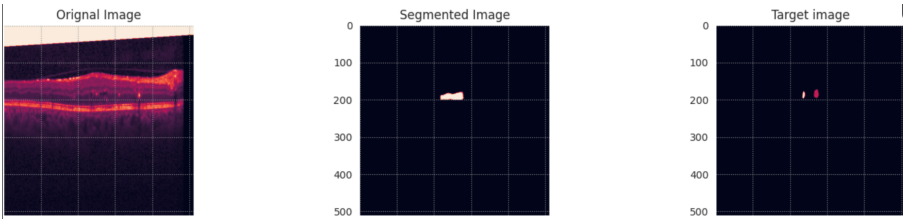


Fig 8 : Representing the segmented and target image using MASK RCNN

In the segmented image figure it shows the detection of DME(diabetic macular edema) using Mask Rcnn

METRIC	VALUE (MODEL 1)	VALUE (MODEL 2)
TRAINING ACCURACY	0.996290418	0.9858
TESTING ACCURACY	0.99620817	0.9808
PIXEL ACCURACY	0.9905	0.9828
IOU	0.712	0.1503
DICE COEFFICIENT	0.83145	0.2077
PRECISION	0.999	0.992
RECALL	0.980	0.722

Fig 9: The table representing the metrics of the models used.

4 Conclusion

In the realm of ophthalmology, where precise diagnosis is paramount, Optical Coherence Tomography (OCT) imaging plays a pivotal role, particularly in the identification of conditions like Diabetic Macular Edema (DME). The objective of this study was to elevate the accuracy of DME detection through advanced segmentation techniques, namely U-Net and Mask RCNN. Our investigation delved into the intricate world of OCT images, emphasizing the need for meticulous segmentation to enhance disease identification and monitoring. The significance of DME diagnosis cannot be overstated, especially in the context of diabetic patients, for whom early detection is instrumental in preventing vision impairment. To do this, a reliable system that can detect dme accurately is necessary. In this study, we propose Mask RCNN and U-Net model for DME detection. The comparative analysis of the performance metrics, including IOU score, Dice coefficient, F1 score, and pixel accuracy, revealed a

notable trend. The U-Net architecture consistently outperformed Mask RCNN across these metrics. The higher values for IOU score, Dice coefficient, F1 score, and pixel accuracy observed with U-Net underscore its efficacy in precise segmentation for DME detection. In the future, Expanding the dataset to include a more diverse set of OCT images from various sources and populations can contribute to the generalizability of the models. Assessing the performance on datasets from different demographics and imaging conditions will strengthen the reliability of the proposed methods.

References

[1]. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson,W. Chang, M. R. Hee, T.Flotte, K. Gregory, C. A. Puliafito, et al., Optical coherence tomography, Science (New York, NY) 254 (1991) 1178.

[2]. M. J. Marquesa, S. Rivetb, A. Bradua, A. Podoleanua, Polarization-sensitive plug-in optical module for a fourier-domain optical coherence tomography system, in Proc. of SPIE Vol, volume 10053, pp. 100531B–

[3]. J. F. De Boer, B. Cense, B. H. Park, M. C. Pierce, G. J. Tearney, B. E. Bouma, Improve signalto-noise ratio in spectral-domain compared with time-domain optical coherence tomography, Optics letters 28 (2003) 2067–2069

[4]. T. A. Ciulla, A. G. Amador, and B. Zinman, "Diabetic retinop-athy and diabetic macular Edema: Pathophysiology, screening,and novel therapies,"Diabetes Care, vol. 26, no. 9, pp. 2653–2664, 2003.

[5]. A. Das, P. G. McGuire, and S. Rangasamy, "Diabetic macularedema: pathophysiology and novel therapeutic targets,"Oph-thalmology, vol. 122, no. 7, pp. 1375–1394, 2015

[6]. R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," Ophthalmology, vol. 124,no. 7, pp. 962–969, 2017.

[7]. Detection of DME by Classification and Segmentation Using OCT Images January 2022 Webology 19(1):601-612

[8]. Karri SP, Chakraborty D, Chatterjee J. Transfer learning-based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration. Biomed Opt Express. 2017;8:579–

[9]. P. Zhou, M. Wang, and H. Cao, "Research on features of retinal images associated with hypertension and diabetes," in 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, pp. 6415–6417, IEEE, 2006.

[10]. Deep-Learning Based, Automated Segmentation of Macular Edema in Optical Coherence Tomography CECILIA S. LEE MD MS, 1 ARIEL J. TYRING MD,1

NICOLAAS P. DERUYTER BA, 2YUE WU PHD,1 ARIEL ROKEM PHD,3 AND AARON Y. LEE MD MSCI1,3* .

[11]. Wang D, Wang L. On OCT image classification via deep learning. In IEEE Photonics Journal. 2019; 11:1–14. doi:10.1109/JPHOT.2019.2934484.M. Arsalan, M. Owais, T. Mahmood, S. W. Cho, and K. R. Park, “Aiding the diagnosis of diabetic and hypertensive retinopathy using artificial intelligence-based semantic segmentation,” Journal of clinical medicine, vol. 8, no. 9, p. 1446, 2019.

[12]. Awais M, Müller H, Tang TB, Meriaudeau F. Classification of SD-OCT images using a deep learning approach. IEEE International Conference on Signal and Image Processing Applications (ICSIPA) 2017:489–92. doi:10.1109/ICSIPA.2017.8120661N..

[13]. Wang X, Tang F, Chen H, Luo L, Tang Z, Ran A, et al. UD-MIL: Uncertainty-driven deep multiple instance learning for OCT image classification. IEEE J Biomed Health Inform. 2020; 24:3431–42.

[14]. Smitha, A., Jidesh, P. Detection of retinal disorders from OCT images using generative adversarial networks. Multimed Tools Appl 81, 29609–29631 (2022). https://doi.org/10.1007/s11042-022-12475-1Q.

[15]. M. Arsalan, A. Haider, J. Choi, and K. R. Park, “Diabetic and hypertensive retinopathy screening in fundus images using artificially intelligent shallow architectures,” Journal of personalized medicine, vol. 12, no. 1, p. 7, 2021.

[16]. S. Naz, A. Ahmed, M. U. Akram and S. A. Khan, "Automated segmentation of RPE layer for the detection of age macular degeneration using OCT images," *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Oulu, Finland, 2016, pp. 1-4, doi: 10.1109/IPTA.2016.7821033.

[17] https://www.kaggle.com/datasets/paultimothymooney/chiu-2015/data