SEGMENTATION OF DIABETIC RETINOPATHY IMAGES

Project supervisor-Dr. Anjali Gautam

Harsh Garg
iit2020082@iiita.ac.in
B.Tech IT
IIIT Alllahabad
UP, INDIA

Ramavath Vasanth Naik
iit2020096@iiita.ac.in
B.Tech IT
IIIT Alllahabad
UP, INDIA

Tarun Harishchandra Pal
iit2020098@iiita.ac.in
B.Tech IT
IIIT Alllahabad
UP, INDIA

Mohit Kumar Mina
iit2020100@iiita.ac.in
B.Tech IT
IIIT Alllahabad
UP, INDIA

***Abstract*—Diabetic retinopathy (DR) is an eye abnormality caused by chronic diabetes that affected patients worldwide. Hard exudate is an important and observable sign of DR and can be used for early diagnosis. This paper presents a method for segmenting the multiple lesions like Microaneurysms, Haemorrhages, Hard and Soft exudates that occurs on the surface of the retina due to diabetes..**

***Index Terms*—Diabetic Retinopathy , lesions , fundus ,exudates, UNet, IDriD.**

I. Introduction

Diabetic retinopathy (DR) stands as a prevalent microvascular complication arising from iabetes, with the potential to induce visual impairment if not diagnosed and treated in a timely manner. In recent years, the field of medical imaging has witnessed notable advancements, particularly in the domain of automated detection and segmentation of DR lesions within retinal fundus images, thanks to deep learning techniques. While these advancements offer promise, a critical problem remains—enhancing the accuracy and efficiency of DR lesion segmentation. The primary challenge lies in accurately distinguishing and delineating the affected regions within the retinal images. This is essential for effective diagnosis and treatment planning. Current methods, though promising, may still fall short in capturing the intricate spatial and contextual information required for precise segmentation

One noteworthy architectural choice deep learning for medical image segmentation is the Residual UNet model. This model combines the U-Net's capacity to capture local details with the principle of residual learning to promote information flow across network layers. However, even with these capabilities, there is room for improvement. The motivation for addressing this problem arises from the critical nature of Diabetic retinopathy. Timely and accurate detection is vital to prevent visual loss and facilitate early intervention. Fundus images often exhibit various forms of microvascular lesions associated with DR, such as microaneurysms, hemorrhages, hard exudates, and soft exudates. Each of these lesions presents unique challenges in segmentation, adding to the complexity of the problem. To overcome these challenges and enhance the field of Diabetic Retinopathy diagnosis and treatment, this thesis aims to contribute a solution by advancing automated DR lesion segmentation techniques. While the Residual UNet architecture is a powerful tool in this endeavor, the focus is on its potential for further improvement, rather than prescribing a specific solution. The primary goal is to develop a method that can better capture the intricate spatial and contextual details within retinal images, thus improving the accuracy of DR lesion segmentation. To evaluate the proposed method, a diverse dataset comprising both normal and DRaffected retinal images will be employed. Preprocessing techniques will be applied to optimize the quality of the input data. The evaluation will encompass training, validation, and testing phases to ensure the robustness and effectiveness of the developed approach. In conclusion, the emphasis here is on the critical problem of enhancing the accuracy and efficiency of DR lesion segmentation in the context of Diabetic Retinopathy diagnosis and treatment. While the Residual UNet architecture is a key tool, the thesis seeks to contribute to the field by addressing this problem rather than prescribing a specific solution. The ultimate goal is to improve patient outcomes by advancing the state of the art in automated DR lesion

II. Related Work

A. Paper 1

- **Author–** DUN Qomariah, H Tjandrasa, C Fatichah
- **Title–** "Qomariah, DUN, Tjandrasa, H Fatichah, C 2021, 'Segmentatio n of Microaneury sms for Early Detection of Diabetic Retinopathy using MResUNet', International Journal of Intelligent Engineering and Systems, vol. 14, no. 3, pp. 359-373. https://doi.or g/10.22266/ij ies2021.0630.30 "
- **Method/ approach used–** "The methodology of MResUNet involves combining UNet and Residual networks to create an efficient architecture for image segmentation,

specifically for detecting microaneurysms in medical images. This hybrid approach leverages UNet's data efficiency and Residual units' fast training capabilities. The network consists of contracting, bridge, and expansive paths, with a total of 68 layers and 21 convolutional layers. It uses identity mapping and batch normalization to enhance performance and standardizes input images to 256x256 for consistency. This methodology is designed to achieve accurate and efficient image segmentation results in medical applications."

- **Achieved Performance –** "sensitivity values of 61.96% and 85.87% on the IDRID and DiaretDB1 datasets "
- **Advantages–** Improved Sensitivity: MResUNet outperforms other methods in detecting microaneurysms, enhancing early diabetic retinopathy diagnosis.
  Enhanced Feature Learning: The architecture effectively captures relevant information from images, overcoming feature degradation.
  Pixel Imbalance Handling: It addresses pixel imbalance issues, ensuring accurate segmentation even for minority classes.
  Multidata set Applicability: MResUNet is versatile and works well with different retinal fundus image datasets.
  Efficiency: Despite complexity, MResUNet offers efficient training and segmentation capabilities.
- **Dataset–** IDRID and DiaretDB1
- **Scope–** The scope of the MResUNet paper is to introduce and evaluate a modified deep learning network (MResUNet) for microaneurysm segmentation in diabetic retinopathy. The paper outlines the model's architecture, training techniques, and assesses its effectiveness by comparing it to other state-of-the-art methods.
- **Year–** 2021

B. Paper 2

- **Author–** Zong, Y., Chen, J., Yang, L., Tao, S., Aoma, C., Zhao, J. H., Wang, S.
- **Title–** "Zong, Y., Chen, J., Yang, L., Tao, S., Aoma, C., Zhao, J. H., Wang, S. (2020). U-Net based method for automatic hard exudates segmentation in FundUs images using inception module and residual connection. IEEE Access, 8, 167225–167235. https://doi.org/10.1109/access.2020.3023273 "
- **Method/ approach used–** "The methodology of the paper "U-Net Based Method for Automatic Hard Exudates Segmentation in Fundus Images" involves a multi-step approach. First, it preprocesses retinal fundus images by cropping and selecting the green channel for optimal contrast. Then, it utilizes super pixel segmentation to generate patches, ensuring each patch contains similar features. These patches are classified into those with and without hard exudates. To tackle data imbalance, a 2:1 ratio of hard exudate to background patches is maintained. The network architecture is based on U-Net with residual units featuring inception modules. Focal Loss is used as the loss function to address data imbalance and emphasize challenging samples. The network is trained using the Adam optimizer with a learning rate of 1e4 over 800 epochs, resulting in an effective method for hard exudates segmentation in fundus images. "
- **Achieved Performance –** "sensitivity, specificity, and accuracy achieve 96.38%, 97.14%, and 97.95% respectively, "
- **Advantages–** Early DR Diagnosis: The paper introduces an automatic method for early diabetic retinopathy (DR) diagnosis by segmenting hard exudates (HE), aiding timely treatment.
  Super pixels: Utilizing super pixels improves context capture, essential for segmentation, and mitigates dataset limitations.
  Advanced Network: The U-Net-based architecture with inception modules and residual connections enhances feature extraction and segmentation.
  Focal Loss: The use of Focal Loss addresses dataset imbalance, leading to better segmentation accuracy.
  High Performance: Extensive tests yield impressive results with high sensitivity, specificity, and accuracy.
- **Dataset–** IDRiD dataset
- **Scope–** The paper's scope includes DR diagnosis via hard exudate segmentation. Future research can explore its applicability to larger datasets and assess computational efficiency for practical use in clinical settings.
- **Year–** 2020

C. Paper 3

- **Author–** Guo, C., Szemenyei, M., Yi, Y., Zhou, W.
- **Title–** "Guo, C., Szemenyei, M., Yi, Y., Zhou, W. (2020). Channel Attention Residual U-NeT for retinal vessel segmentation. ResearchGate. https://www.researchgate.net/publication/340523534 Channel Attention Residual U-Net for Retinal Vessel Segmentation "
- **Method/ approach used–** "The study uses a modified U-Net architecture called CAR-UNet for retinal vessel segmentation. CAR-UNet has two paths: a contracting path with Channel Attention Double Residual Blocks and max-pooling, and an expansive path with transposed convolution. They introduce Multi-Element Channel Attention (MECA) for better feature extraction, which involves a lightweight 1D convolutional layer to generate a channel attention map. Additionally, they consider using Channel Attention Double Residual Blocks for complex feature extraction, inspired by DRNet. "
- **Achieved Performance –** "CAR-UNet has the highest AUC (0.45% / 0.73% / 0.53% "
- **Advantages–** Achieves state-of-the-art performance in retinal vessel segmentation. Introduction of Modified Efficient Channel Attention (MECA) and Channel Attention Double Residual Block (CADRB) for improved feature extraction. Consideration of feature channel relationships enhances segmentation accuracy. Lightweight MECA offers performance gains without significant complexity.

- **Dataset–** DRIVE, CHASE DB1 and STARE.
- **Scope–** Future research can optimize model complexity. Explore transferability to other medical image tasks. Improve feature interpretability. Assess generalization across diverse datasets and challenges in medical image analysis.
- **Year–** 2020

D. Paper 4

- **Author–** Author Bowen Cheng, Ishan Misra, Alexander G. Schwing, Alexander Kirillov, Rohit Girdhar
- **Title–** "Maskedattention Mask Transformer for Universal Image Segmentation"
- **Method/ approach used–** "Masked-attention Mask Transformer (Mask2Former) is capable of addressing any image segmentation task (panoptic, instance or semantic). Its key components include masked attention, which extracts localized features by constraining crossattention within predicted mask regions "
- **Achieved Performance –** "Mask2Former sets for panoptic segmentation (57.8 PQ on COCO), instance segmentation (50.1 AP on COCO) and semantic segmentation (57.7 mIoU on ADE20K). "
- **Advantages–** Mask2Former obtains top results in all three major image segmentation tasks (panoptic, instance, and semantic) on four popular datasets, outperforming even the best-specialized models designed for each benchmark while remaining easy to train. Mask2Former saves 3× research effort compared to designing specialized models for each task, and it is accessible to users with limited computational resources In addition to reducing the research effort by at least three times, it outperforms the bestspecialized architectures by a significant margin on four popular datasets
- **Dataset–** COCO [35] (80 "things" and 53 "stuff" categories), ADE20K [65] (100 "things" and 50 "stuff" categories), Cityscapes (8 "things" and 11 "stuff" categories) and Mapillary Vistas (37 "things" and 28 "stuff" categories)
- **Scope–** Mask2Former trained on panoptic segmentation only performs slightly worse than the exact same model trained with the corresponding annotations for instance and semantic segmentation tasks across three datasets. This suggests that even though Mask2Former can generalize to different tasks, it still needs to be trained for those specific tasks. In the future, goal is to develop a model that can be trained only once for multiple tasks and even for multiple datasets.
- **Year–** 2022

E. Paper 5

- **Author–** Pavani, P. G., Biswal, B., Gandhi, T. K.
- **Title–** "Pavani, P. G., Biswal, B., Gandhi, T. K. (2023). Simultaneous multiclass retinal lesion segmentation using fully automated RILBP-YNet in diabetic retinopathy. Biomedical Signal Processing and Control, 86, 105205. https://doi.org/10.1016/j.bspc.2023.105205 "
- **Method/ approach used–** "introduced a novel methodology for simultaneous multiclass retinal lesion segmentation in diabetic retinopathy using the RILBP-YNet model. Our approach begins with meticulous pre-processing to enhance fundus images, followed by the computation of Local Binary Patterns (LBP) to capture essential texture information. RILBP-YNet, with its dual-encoder architecture, then takes these representations and the original image as inputs. The model employs Residual-Inception blocks to effectively extract features and utilizes a decoder to generate segmented masks for various retinal lesion classes. This methodology empowers RILBP-YNet to provide accurate and efficient retinal lesion segmentation, contributing to the diagnosis and treatment of diabetic retinopathy. "
- **Achieved Performance –** "an average dice coefficient of 95 % for OD, 70 % for SE, 75 % for EX, 50 % for HE, and 35 % for ME respectively.
"
- **Advantages–** Accurate Segmentation: RILBP-YNet accurately segments diabetic retinopathy lesions, aiding in early detection and treatment.
Dual Encoder: Its dual encoder architecture captures both traditional features and textural patterns for improved lesion detection.
Residual-Inception: Residual-Inception modules enhance multi-scale feature extraction, improving lesion pattern recognition.
Class Imbalance Mitigation: The model addresses class imbalance issues, enhancing lesion pixel detection.
Competitive Performance: Achieves competitive results, outperforming existing methods in specific lesion segmentation tasks.
- **Dataset–** IDRiD Dataset. The proposed method is also tested on datasets like DRIVE, STARE, CHASE-DB1
- **Scope–** Generalization: Enhancing model generalization across diverse datasets and imaging techniques.
Real-time Applications: Adapting for real-time clinical use.
Interpretability: Developing methods for result interpretation.
Clinical Integration: Collaborating with healthcare institutions for practical implementation.
- **Year–** 2023

F. Paper 6

- **Author–** Huang, Ko-Wei, Yao-Ren Yang, Zih-Hao Huang, Yi-Yang Liu, Shih-Hsiung Lee
- **Title–** "Huang, Ko-Wei, Yao-Ren Yang, Zih-Hao Huang, Yi-Yang Liu, and Shih-Hsiung Lee. 2023. "Retinal Vascular Image Segmentation Using Improved UNet Based on Residual Module" Bioengineering 10, no. 6: 722. https://doi.org/10.3390/bioengineering10060722 "
- **Method/ approach used–** "The proposed methodology involves enhancing retinal vessel segmentation using a modified U-Net architecture with added residual blocks,

full-scale skip connections, and inception blocks. Data preprocessing techniques, including grayscale conversion, filtering, contrast enhancement, and gamma correction, are applied to improve image quality. The residual blocks address the vanishing gradient problem, while skip connections integrate information from various feature scales. Inception blocks increase feature diversity. This comprehensive approach enhances accuracy and robustness in retinal vessel segmentation. "

- **Achieved Performance –** "F1-measure reaching 77.8
- **Advantages–** Advanced Segmentation Model: The paper introduces an advanced retinal blood vessel segmentation model, promising improved diagnostic efficiency.
  Full-Scale Skip Connection: The inclusion of full-scale skip connections enhances segmentation accuracy by combining fine details with high-level features.
  Experimental Validation: The proposed method is rigorously tested and outperforms existing models, indicating its practical utility.
- **Dataset–** DRIVE and ROSE
- **Scope–** The paper contributes to the evolving field of medical image analysis by presenting an innovative approach to retinal vessel segmentation. Its success in outperforming existing models suggests potential applications in clinical diagnosis and beyond. Further research could explore its adaptability to other medical imaging tasks and address specific challenges in retinal vessel segmentation in more detail.
- **Year–** 2023

G. Paper 7

- **Author–** Fu Yinghua, Ge Zhang, Xin Lu, Honghan Wu, and Dawei Zhang.
- **Title–** "Fu, Yinghua, Ge Zhang, Xin Lu, Honghan Wu, and Dawei Zhang. "RMCA U-net: Hard exudates segmentation for retinal fundus images." Expert Systems with Applications 234 (2023): 120987. "
- **Method/ approach used–** "This paper proposes RMCA U-net which features a U-shape framework combined with a residual structure to obtain the subtle features of hard exudate (the percentage of hard exudates in the whole fundus image is relatively small, and their shapes are often irregular and the contrasts are usually not high enough). A multi-scale feature fusion (MSFF) module and an improved channel attention (CA) module are designed to segment sparse small lesions. "
- **Achieved Performance –** "F1 score was 54.43%. The method in this paper is increased by 6% higher in PR-MAP than U-net on the IDRID dataset, increased by 10% in Recall than U-net on the Kaggle dataset, and increased by 20% in F1-score than U-net on the local dataset. "
- **Advantages–** Enhanced Accuracy: The RMCA U-net achieves superior segmentation accuracy for hard exudates compared to other CNNs.
  Multi-scale Feature Fusion: The MSFF module enhances robustness by capturing features from various receptive fields.
  Channel Attention: Improved channel attention aids in effective feature extraction.
  Residual Structure: Inclusion of a residual structure helps capture subtle exudate characteristics.
  Generalizability: The method is validated on ultra-widefield fundus images, demonstrating its broader applicability.
- **Dataset–** IDRID, Kaggle and one local data set
- **Scope–** Clinical Applications: The method has potential applications in computer-aided diagnosis systems for diabetic retinopathy and macular edema.
  Further Optimization: Future research can explore optimizations to reduce computational complexity while maintaining accuracy.
  Transfer Learning: Applying transfer learning techniques could enable the adaptation of the model to different retinal imaging modalities and conditions.
  Integration: Integration into existing medical imaging systems and collaboration with healthcare providers could facilitate real-world clinical deployment.
- **Year–** 2023

III. Dataset

IDRiD The IDRiD (Indian Diabetic Retinopathy Image Dataset) is a publicly available dataset specifically designed for Diabetic Retinopathy research, including segmentation tasks. Here is some information about the IDRID dataset that can be useful for your thesis work:– The IDRID dataset consists of retinal fundus images captured using a Topcon TRC-NW6 nonmydriatic camera. – It contains a total of 81 fundus images. – Each image has a resolution of 4288 x 2848 pixels. – The IDRID dataset includes manual annotations for two main Diabetic Retinopathy lesions: microaneurysms (MA) and exudates (EX). – The annotations are provided as pixel-level masks, where each lesion is labeled with a unique identifier. – The annotations can be used to train and evaluate segmentation models to automatically detect and delineate the lesions

IV. Material & Methods

In this paper, we have used IDRID dataset for training a Spatial Attention Modified U-Net model with attention mechanisms for Segmentation of Diabetic Retinopathy lesions from fundus pictures using deep learning techniques The general workflow of our methodology is shown in Fig. 1 and sample images of Diabetic Retinopathy lesions are shown in Fig. 2.

U-Net-based models are widely used for the image segmentation task and it has an encoder and decoder for performing this operation. The Encoded part takes the volume as input of size $512 \times 512 \times 3$. There are a total of 4 layers in the model including the input layer. Channel-wise attention mechanisms are applied at different levels of the model during encoding and decoding which help in avoiding feature loss. In encoder, the number of feature maps increases as we go from top to bottom and reduces the size of feature maps. In the decoder, the number of feature map doubled after each upsampling. The top layer of the decoder contains output made up of a convolution layer and a sigmoid activation function. The architectural model of our approach is shown in Fig. 3, 4, and 5.

In the methodology, before doing any dataset training, we have done pre-processing on the dataset and details of it are given below. Thereafter, proposed architectural model as mentioned in Algorithm 1 and Fig. 3, 4, and 5 is used for the final model training.

A. Pre-processing

A custom generator is generated and preprocessing is done for the IDRID dataset. The IDRID dataset has high quality scans of retinal lesions with their corresponding segmentation masks.

Steps:

- Define a MinMaxScaler object that scales the intensity values of the MRI scans.
- The scaler is used to normalize the data to a range of 0 to 1 to improve the convergence of the deep learning model.
- The scans are stored in separate lists for each Diabetic Retinopathy lesions: microaneurysms (MA) and exudates (EX), while the masks are stored in a separate list.
- Load the Retinal scans and their corresponding masks using the nibabel library.
- Scale the intensity values of the Retinal scans using the MinMaxScaler object.
- Crop the image volume and its corresponding mask to remove any empty borders.
- Convert the mask to a one-hot encoded representation using the to_categorical function from tensorflowk̇erasu̇tils.
- Save the preprocessed image and mask as NumPy arrays.
- Then the preprocessed data is divided into training and validation sets of 75:25.

B. Implementation Details and Results

In this section, we have mentioned details of various loss functions used during the training and validation phase. The details of hyperparameters and results obtained on testing sets are also mentioned.

1) Loss Function: Dice Loss Function: It finds the similarity between two samples and is represented as:

$$DiceLoss = 1 - (2 * TP)/(2 * TP + FP + FN) \tag{1}$$

where TP is True Positive, FP is False Positive, and FN is False Negative.

2) Categorical Focal Loss: A Focal Loss function addresses class imbalance during training in tasks like object detection.

$$FocalLoss = -\alpha(1-p)^{\gamma}\log(p) \tag{2}$$

where α is the balancing parameter, p is the predicted probability, and γ is the focusing parameter. The balancing parameter α is used to balance the contribution of each class to the loss function, while the focusing parameter γ is used to control the rate at which the loss function focuses on hard-to-classify samples.

Algorithm 1 Model Training

```
procedure TRAINMODEL
    Define loss as loss
    Define metrics as metrics
    Define optimizer as optimizer
    Set batch_size
    Set steps_per_epoch
    Build U-NET model with attention
    Train the model for specified epochs
    for each epoch in range(num_epochs) do
        for each step in range(steps_per_epoch) do
            Load batch of images and masks from train_gen
            Perform forward and backward passes
            Update model parameters
            Record loss and metrics
        end for
        Load saved model
        Initialize mIOU as 0
        for each batch of validation images and masks in val_gen do
            Calculate mIOU for the batch
            Predict masks using loaded model
            Calculate mIOU between predicted masks and ground truth masks
        end for
        Select an image for prediction
        Load and preprocess test image
        Predict mask for the test image
    end for
end procedure
```

3) Metrics: The total loss function is a combination of the Dice Loss and the Categorical Focal Loss. The weights of the two loss functions are adjusted using a hyperparameter to control the trade-off between them during the training process. In the stated paper, the weight of the Categorical Focal Loss is 1.

The formula for Intersection over Union (IoU) Score is as follows:

$$IoUScore = TP/(TP + FP + FN) \tag{3}$$

4) Optimiser - Adam optimiser: The Adam optimizer is a popular optimization algorithm for training neural networks. It is a stochastic gradient descent method that uses adaptive learning rates to update the weights of the neural network. The algorithm combines the advantages of two other popular optimization algorithms, Adagrad and RMSprop. The Adam optimizer uses two momentum parameters, beta1 and beta2, to control the decay rate of the moving averages of the gradient and the squared gradient, respectively. The update rule for the

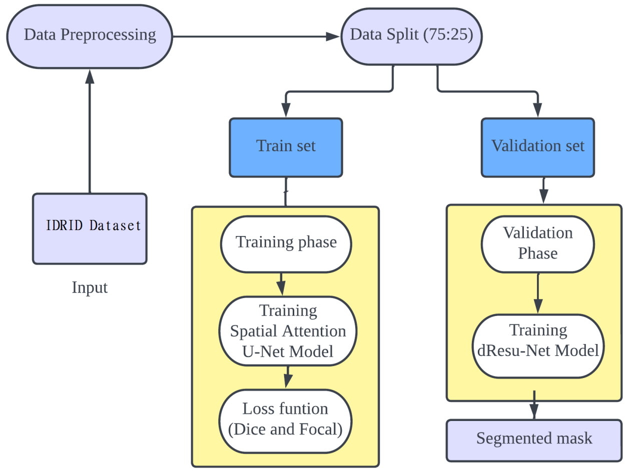

Fig. 1. The Proposed Methodology's General Flow.

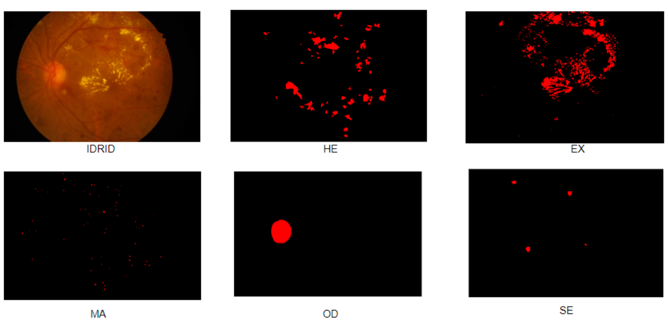

Fig. 2. Sample image of Retinal scans in IDRID dataset.

weights is given by [?]:

$$w(t+1) = w(t) - \eta \frac{m(t)}{\sqrt{v(t)} + \epsilon} \quad (4)$$

where w(t) is the weight at time step t, η is the learning rate, m(t) and v(t) are the moving average of the gradient and squared gradient respectively, and epsilon is a small value in denominator to just prevent division by zero.

5) Evaluation Measure:

- Dice similarity coefficient (DSC), specificity, and sensitivity are used to assess the proposed approach

$$Dice\_coefficient = 2*TP/(2*TP+FP+FN) \quad (5)$$

- A binary cross-entropy loss function has also been used.

V. Implementation Details and Results

The proposed model (dResU-Net with spatial attention) was implemented using Python programming language, Keras library, and TensorFlow as the backend. An ADAM optimizer was used with a learning rate of 0.0001, ReLU as the activation function with batch normalization was also used.

The model was trained for 100 epochs on a batch size of 2 due to the limited computational resources on RTX4090. The experiments were conducted on the IDRID benchmark dataset, from which 70% of the data was used for training, 20% data for testing, 10% for validation.

Table I depicts the values of the hyperparameters, which have been used in the experimental setup.

TABLE I
Hyperparameters of Proposed Spatial Attention Model

Hyperparameters	**values**
Input size	$4288 \times 2848 \times 3$
Learning rate	0.0001
Batch Size	2
The hidden layer activation function	ReLU
Optimizer	ADAM
Loss function	binary cross-entropy loss function
No. of epochs	100
Dropout	0.1 - 0.2
Output size	$512 \times 512 \times 3$
Output layer activation function	Softmax

A. Analysis of the results

1) Training and Validation Details: The training loss per epoch and validation loss per epoch are shown with The Accuracy Per epoch of training and validation.

- The Training and validation were done on the IDRID dataset which consists of 81 images
- After preprocessing and Data Augumentation we got 126 useful images.
- These are further split into 70:20:10 ratio for training, testing and validation.

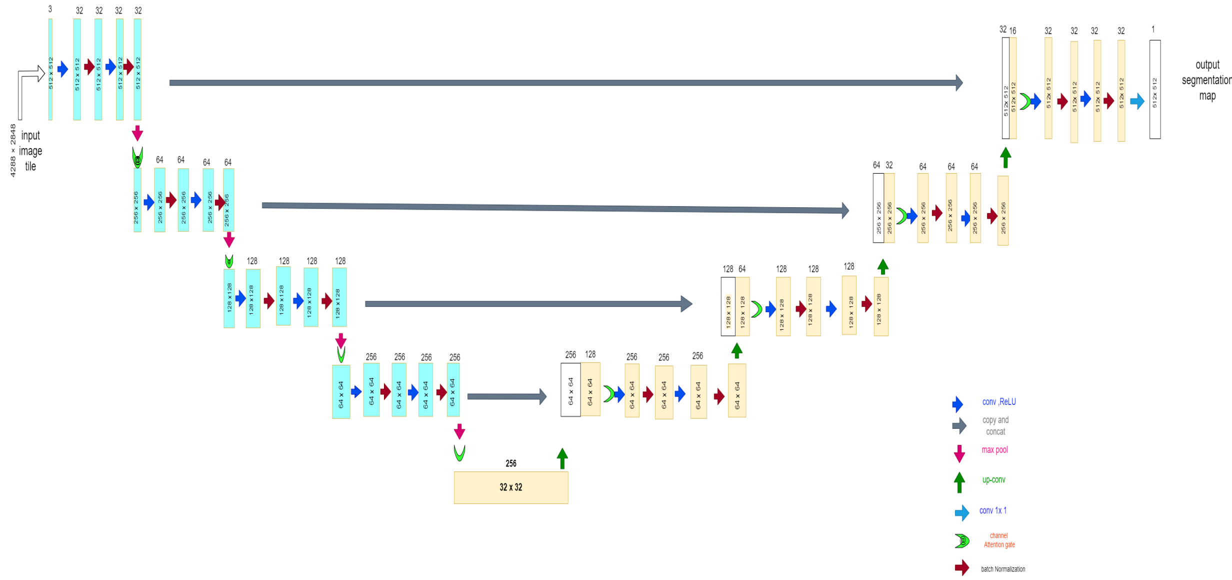

Fig. 3. Spatial Attention Based Architecture

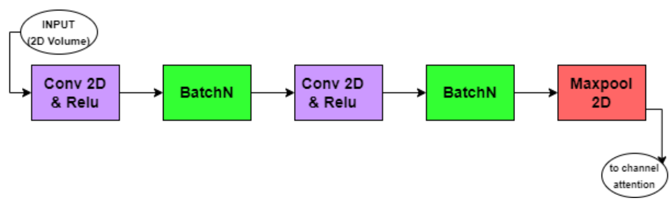

Fig. 4. Operations in each layer of contracting path.

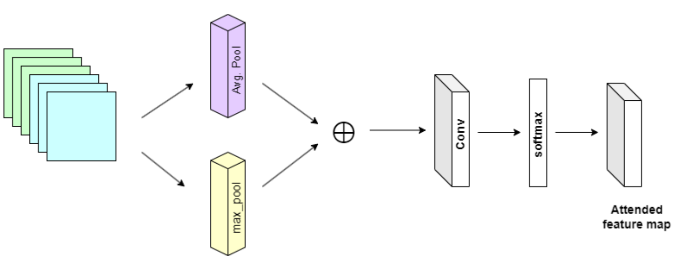

Fig. 5. Spatial Attention Module Overview.

The visual results of the experiments are shown in Fig. 6, 7 and 8, and their training and validation losses are shown in Fig. 9, 10 and 11 for simple UNET, modified UNET and Spatial attention UNET respectively.

VI. Conclusion

Paper introduced an approach for Segmentation of Diabetic Retinopathy using a modified UNET with spatial attention. Accurate segmentation is crucial for diagnosis and treatment planning. Our model's channel attention mechanism effectively preserves local details during encoding and decoding, leading to impressive segmentation results

This work holds promise for advancing brain tumor diagnosis and treatment by combining UNET's power with strategic attention mechanisms, paving the way for further research in the field.

References

1) Qomariah, DUN, Tjandrasa, H Fatichah, C (2021). Segmentation of Microaneurysms for Early Detection of Diabetic Retinopathy using MResUNet. International Journal of Intelligent Engineering and Systems, 14(3), 359-373. https://doi.org/10.22266/ijies2021.0630.30 doi:10.22266/ijies2021.0630.30.
2) Zong, Y., Chen, J., Yang, L., Tao, S., Aoma, C., Zhao, J. H., Wang, S. (2020). U-Net based method for automatic hard exudates segmentation in FundUs images using inception module and residual connection. IEEE Access, 8, 167225–167235. https://doi.org/10.1109/access.2020.3023273 doi:10.1109/access.2020.3023273.
3) Guo, C., Szemenyei, M., Yi, Y., Zhou, W. (2020). Channel Attention Residual U-NeT for retinal vessel segmentation. ResearchGate. https://www.researchgate.net/publication/340523534_Channel_Attention_Residual_UNet_for$_{Retinal_Vessel}$ $\_SegmentationLink.$
4) Pavani, P. G., Biswal, B., Gandhi, T. K. (2023). Simultaneous multiclass retinal lesion segmentation using fully automated RILBP-YNet in diabetic retinopathy. Biomedical Signal Processing and Control, 86, 105205. https://doi.org/10.1016/j.bspc.2023.105205 doi:10.1016/j.bspc.2023.105205.
5) Huang, Ko-Wei, Yao-Ren Yang, Zih-Hao Huang, Yi-Yang Liu, and Shih-Hsiung Lee. (2023). Retinal Vascular Image Segmentation Using Improved UNet Based on Residual Module. Bioengineering, 10(6), 722.

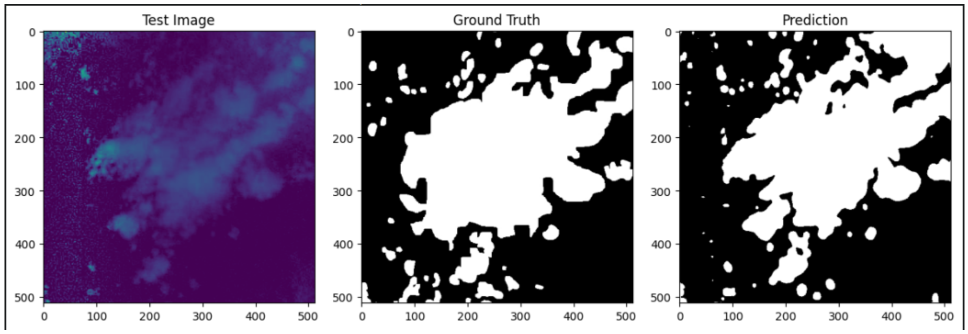

Fig. 6. Simple UNet architecture prediction image.

https://doi.org/10.3390/bioengineering10060722 doi:10.3390/bioengineering10060722.

6) Fu, Yinghua, Ge Zhang, Xin Lu, Honghan Wu, and Dawei Zhang. (2023). RMCA U-net: Hard exudates segmentation for retinal fundus images. Expert Systems with Applications, 234, 120987.