An Enhanced Segmentation and Deep Learning Architecture for Early Diabetic Retinopathy Detection

Renato R. Maaliw III
College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines
rmaaliw@slsu.edu.ph

Zoren P. Mabunga
College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines
zmabunga@slsu.edu.ph

Maria Rossana D. De Veluz
College of Engineering
Southern Luzon State University
Lucban, Quezon, Philippines
mrddeveluz@slsu.edu.ph

Alvin S. Alon
Digital Transformation Center
Batangas State University
Batangas City, Philippines
alvin.alon@g.batstate-u.edu.ph

Ace C. Lagman
Information Technology Dept.
FEU Institute of Technology
Manila, Philippines
aclagman@feutech.edu.ph

Manuel B. Garcia
Information Technology Dept.
FEU Institute of Technology
Manila, Philippines
mbgarcia@feutech.edu.ph

Luisito Lolong Lacatan
College of Computing & Engineering
Pamantasan ng Cabuyao
Cabuyao, Laguna, Philippines
louie1428@gmail.com

Rhowel M. Dellosa
Research and Development Office
University of Northern Philippines
Vigan City, Philippines
rhowel.dellosa@unp.edu.ph

2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC) | 979-8-3503-3286-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/CCWC57344.2023.10099069

***Abstract*—Diabetic retinopathy is a serious complication needing prompt diagnosis and medication to avert vision loss. Lesions caused by the condition are difficult to track because they are hidden behind the eye's structure in small and subtle forms. To extract relevant features, we created a robust pipeline using multiple preprocessing techniques, image segmentation architecture (DR-UNet) with atrous spatial pyramid pooling, and an attention-aware deep learning convolutional network with different modules based on ResidualNet. Empirical results show that our framework has segmentation accuracies of 87.10% (intersection over union) and 84.50% (dice similarity coefficient). Moreover, classification performance of 99.20% provided better results than existing schemes, as reinforced by the smooth convergence of training/validation loss and accuracy. This study has the potential to supplement traditional diagnosis to identify better the ailment in its early and advanced stages.**

Keywords—attention-aware DCNN, atrous spatial pyramid pooling, blindness, fundus images, lesion detection, ophthalmology

I. Introduction

Excessive sugar levels in the blood cause severe harm to body organs resulting in life-threatening illnesses if left unchecked. Diabetic retinopathy (DR) is a medical disorder affecting people with diabetes. It can be either mild (DR-NP, non-proliferative) or severe (DR-P, proliferative). At first, a person with DR-NP experiences blurred vision. However, as the condition progresses, new blood vessels (BV) emerge in the retina, further debilitating visual acuity. The complications from these leaks and bleeding trigger DR. At its worst, the disease causes complete obstruction of the BV, manifesting as various lesions. The most noticeable lesions are microaneurysms, the earliest indications of DR, and appear as little scattered dots with a round shape on fundus images, as depicted in Fig. 1(a) and Fig 1(b). At present, an ophthalmologist manually evaluates the scans to spot the abnormalities. An automated prognosis system can replace the technique to determine conditions rapidly and reliably. Both supervised and unsupervised approaches have been explored in the pursuit of highly accurate outcomes. Deep learning (DL) techniques are applied at the pixel level to recognize and segment retinal pictures. Convolutional neural networks (CNN) are used in deep learning. They are made up of interconnected 'neurons' of multiple layers and have widespread utilization in pattern recognition and image processing.

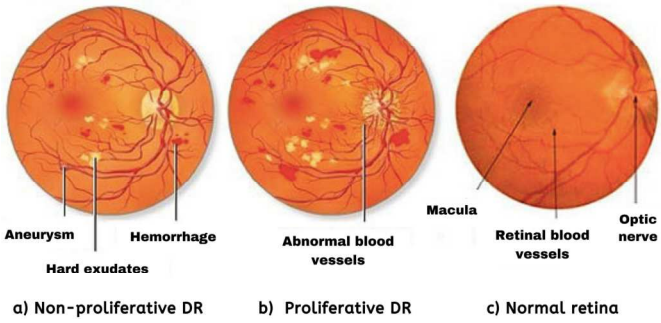

Fig. 1. Diabetic retinophaty cases (a & b) and normal retina (c) [1].

Fundus image-based pathological screening is an emerging study area in modern healthcare diagnostics. Experts in computer vision are starting to pay significant attention to the automated detection of DR. Since the segmentation's output is critical to classification success, it has been the subject of numerous papers. In their research, authors [2] outlined a method for accurately segmenting images of the retinal fundus in two phases through pre and post-processing of images via maximum principal curvature. Using mathematical morphology, proponents [3] segmented BV to assess and extract retinal image features using a smoothing technique. It effectively suppresses background information, while a K-means clustering algorithm improves the image quality. The method is compared to other alternative solutions achieving 95.10% precision evaluated on the DRIVE dataset. Scientists [4] introduced a segmentation framework to distinguish BV based on extended descriptions to address the difficulties of variations with thin and elongated morphologies. Its model training is done so that parameters are automatically learned by a support vector machine (SVM). The work by [5] aims to apply enhancement techniques to quantify the properties of optic BV. Partitioning was accomplished using three essential processes such as multi-scale analytics (MA), morphological geometry (MG), and Gaussian methods (GM). MG excels in

pinpointing the delicate intricacies of blood veins. If combined with image transformation and discriminant analysis, the MA was optimal because it can distinguish between large and small arteries without any background noise. At the same time, GM was potent in identifying thick vessels. Researchers [6] demonstrate automatic retinal vascular splitting using c-means (fuzzy) segmentation and intensity ranges. Restricted dynamic histogram normalization raised the contrast of retinal scans. At the same time, mathematical morphology approaches and matching filtering sequences through Frangi & Gabor filters boosted the clustering. The initial blood vein network is extracted based on a genetic algorithm utilizing an extended spatial fuzzy c-means scheme. An integrated level set enhances the segmentation, achieving an accuracy of 96.10%. A solution that shortens the amount of time needed for calculations while maintaining high accuracy and improving sensitivity compared to the Frangi method is presented by [7]. During its process, they tried to prevent potential problems (e.g., specular reflexes) using two open access (DRIVE & STARE) datasets with 93% and 95% precisions.

According to the literature, there are three popular segmentation schemes such as watershed transform (WST), super pixels (SP) & active contour (AC). The WST is a known and robust unsupervised algorithm capable of handling many image partitioning challenges. However, in some instances, it tends to excessively segment the original images as the technique is sensitive to noise and contrast leading to incorrect results [8]. SP enables image over-segmentation by arranging pixels into monolithic clusters based on derived characteristics (e.g., texture, intensity & other related features). Each pixel of the same region is represented on a map considering both local and nonlinear deformations. This process allows efficiency by minimizing mistakes of assigning wrong labels in making super pixel-based features with low computation and memory usage [9]. Conversely, AC is good at acquiring regions of interest (ROI) in medical images. It uses edge and region-based models to locate edge information, set the approximate shape boundaries, and compute sub-regions based on optimal fitness [10], respectively.

There are multiple papers on identifying and categorizing DR in fundus images. For instance, [11] demonstrates an alternative morphological function that emphasizes optical disc (Opt-D) recognition. They attained specificity and sensitivity of 88% and 94%. Authors [12] offer unsupervised retinal Opt-D histogram-based template fitting detection techniques comprising three phases: fine-tuning, centroid, and boundary identification. Other hemorrhage research has been conducted without segmentation or ROIs, utilizing retinal characteristics as features and classified by partial least square [13]. Data scientist [14] presented an iterative technique for enhanced BV detection utilizing Opt-D algorithms in five steps, with fovea localization as one of its strengths. A machine learning approach aiding DR identification was developed by [15]. They validated the results from thousands of eye scans (MESSIDOR dataset) with the macula as the primary priority. A kappa (*kp*) value compared two out of three doctor's in grading DR against the machine learning performance with a mean average accuracy of 90%. With the same data, researchers [16] proposed a four-layered CNN starting with preprocessing, transformation, and normalization of input images. The effect was a quality improvement in DR grading accomplished via clustering with accuracy of 82%. Finally, the probe of [17] devised a hybrid method to develop a hybrid pipeline using Hough transformation and CNN to detect fluids (exudates) leaking from BV. It calculates the categorization using pixel probabilities. The reliability of their assessments based on three public datasets (DiaretDB0, DiaretDB1 & DrimDB) averages 98% accuracy. Several models used deep learning (neural networks) for DR diagnostics based on recent publications. Attention-aware mechanism [18], multi-task [19] & self-attention CNNs [20], and auto encoders [21] provides optimized processes representation sharing during modeling, thus improving the overall rate of classification. Moreover, it includes histogram equalization and sparse filters that automatically extracts different features from fundus images. Other benchmarks [22] utilize long short-term memory networks (LSTM) to enhance classification performance further. Parameter reduction was explored by [23] by modifying the residual pathways of networks via sampling modules for challenging datasets.

Each method stated in the literature has its advantages but we are confident that there are still gaps in the exactitude of early lesion detection. In the advanced stage of DR, noticeable vascular growth markings are conveyed easily by the current state-of-the-art image processing techniques. However, this is not the case for mild conditions, which may not have enough signs and proliferations of the eye's vascular abnormalities using only the naked eye. As a contribution to science, our proposed novel pipeline comprises extensive preprocessing techniques; enhanced segmentation algorithms with atrous spatial pyramid pooling (ASPP) for extracting the maximum lesion features efficiently, and a modularized attention-aware deep CNN architecture (DCNN) for DR classification. This study can assist ophthalmologists in reliably predicting its progression at the earliest possible time to give patients the necessary and proper treatment to prevent permanent blindness.

II. Methodology

A. Datasets and Preprocessing

We collected 388 high-resolution retinal images divided into two equally independent groups. The first set contains 194 data exhibiting DR (DiaretDB0 & DiaretDB1), while the other includes 25 healthy eyes. Due to an imbalance ratio of 7:1, we acquired 169 additional non-DR images from various sources [24] to solve these vast discrepancies. In order to improve training and validation of deep learning models, it is crucial to have a well-balanced dataset to prevent under and over-fitting during training. Moreover, all images are converted to 1440 x 900 pixels to compensate for size difference then normalized into 512 x 512 pixel images after background cropping. Fig. 2 shows excerpts of sample images annotated by medical experts.

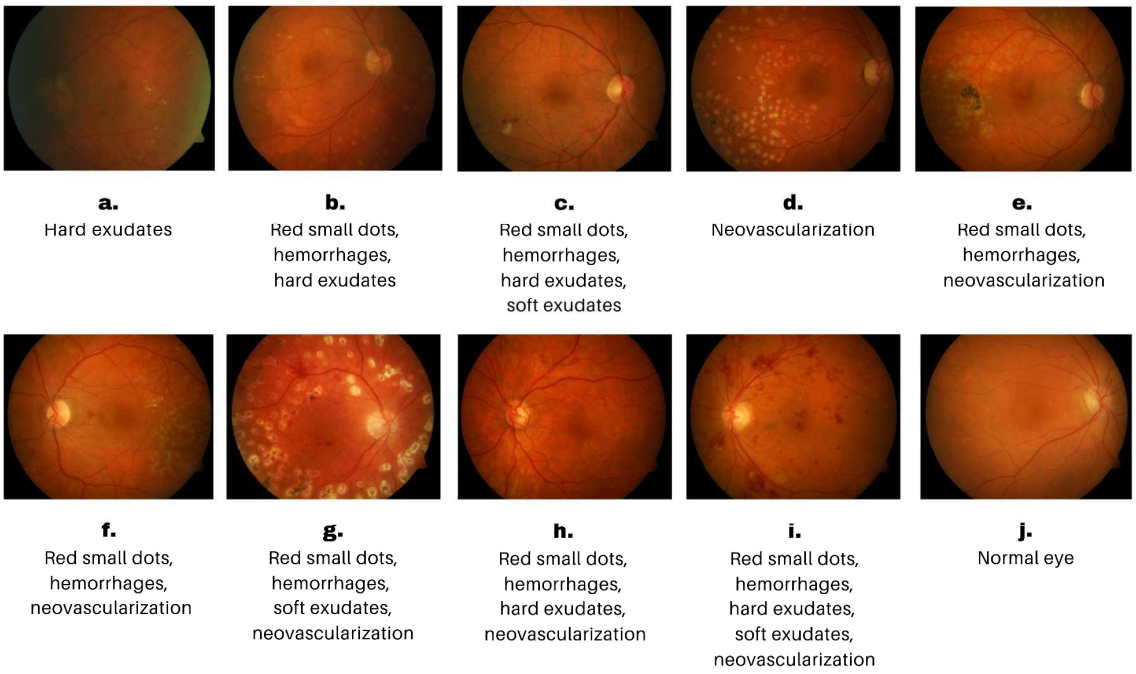

Fig. 2. Excerpts of fundus images showing examples of diabetic retinopathy (a – i) and normal eye (j) [24].

B. Data Augmentation

Inadequate medical images for training can make it hard to improve pattern recognition as it affects how well a machine learning model can perform. Therefore, data expansion is necessary to increase the size of the dataset by artificial means to achieve robustness and generalization. We applied geometric transformations such as flipping (vertical & horizontal) and rotations (45 & 135 degrees) to produce additional 1552 synthetic data with a total of 1940 images, 970 for each class. Through this process, it can help improve the robustness and generalization ability of the model to make accurate predictions. Lastly, we divided the dataset into training (80%), testing (20%), and excluded ten early case DR samples using a 10-fold cross-validation.

C. Color Adjustments for Maximum Feature Extraction

When processing images, color calibration is vital since it enriches the visual quality and suits a specific context or application. Our strategy is to convert each image to grayscale to reduce data size to cut computation time by a factor of 0.2815 using the values red (R) = 0.27, green (G) = 0.56, and blue (B) = 0.17 (Fig. 3). Moreover, this also increases training time efficiency during test runs.

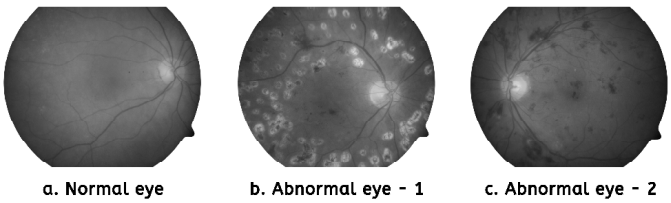

Fig. 3. Sample grayscaled fundus images, normal (a) and abnormal (b & c).

It is obvious to give emphasis to the red channels to highlight blood vessels and clots but this is not the case according to our observations. Instead, we prioritized the green colors over the red and blue channels with a second adjustment of R = 0.28, G = 0.57, and B = 0.15 as it reveals more details. Fig. 4 illustrates the comparative color shift prioritization.

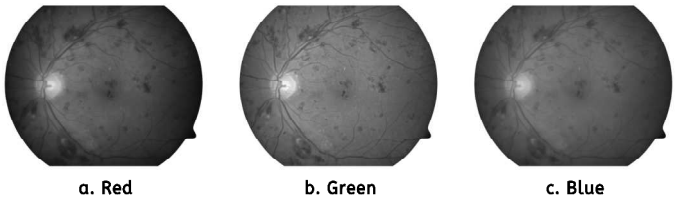

Fig. 4. Prioritizing the green channel produced finer and greater information (b) than the other channels (a & c).

D. Enhanced Segmentation Technique

Retinal image segmentation is an indispensable process that helps diagnose DR. It locates specific areas that are hard to see otherwise. This study aims to divide the visuals into pixels or regions concerning the optic disc, BV, lesions (hemorrhages), exudates, and microaneurysms using semantic segmentation. We modified the U-Net architecture to address this dilemma because the original design was ineffective in our segmentation contexts after several tests.

Our framework comprises three major parts: encoder, bottleneck, and decoder (Fig. 5) similar to the article by [25] but with additional mechanics. The encoder with four blocks (BLK) is on the left, with the first receiving the input images and the subsequent BLK receiving the previous output with a lower subsampled resolution. Its initial BLK comprises three convolutions, the first two layers having 32 feature maps (FM) and 64 in the third. The same structure is observed in the second unit but with 64 and 128 FM on its layers while the remaining BLKs (3rd and 4th) offer only two convolutions

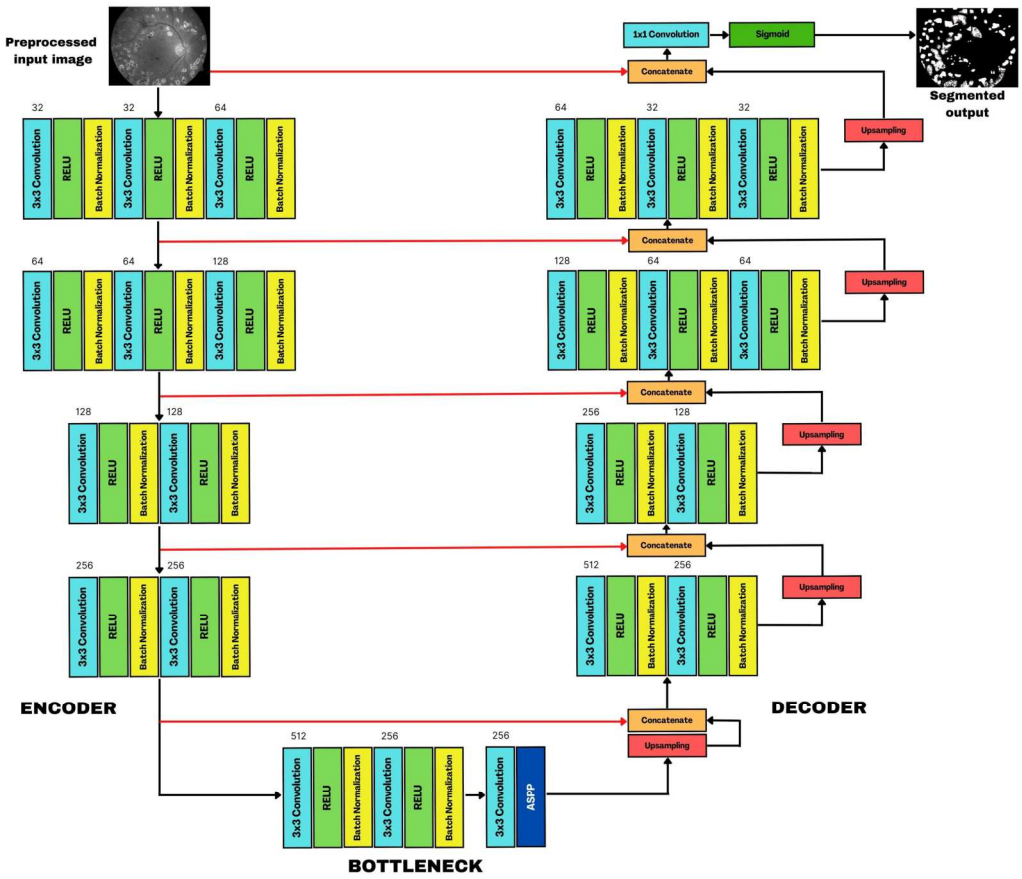

Fig. 5. Proposed DR-UNet for lesion segmentation for detection of diabetic retinopathy in fundus images with three parts (encoder, bottleneck & decoder).

while doubling subsampled FM at 128 and 256. A *Maxpooling* (MXP) operation is executed at the end of each BLK to decrease the input's sample size while keeping most details. Before each MXP layer, the preservation of a duplicate copy connects the contraction to the expansion path leading to precise segments.

The decoder on the right side mirrors the encoder but includes an upsampling (2D) at the end. Its feature map decreases by half from the previous block during each sampling stage. Furthermore, the red arrows (Fig. 5) recreate the FM previously stored before each maximum clustering layer in each encoder block and compare these with the decoder's oversampled outputs. We applied subsampling at the end of the decoder block to improve the ability to detect retinal hemorrhages, further enhanced by concatenating the input image with the network results. To complete the process, a final convolution (1 x 1 kernel) and a sigmoid activation function compute the maps' outcomes.

In the middle of the architecture is a bottleneck connecting the two pathways with two convolutions (3 x 3 kernel, MXP = 2 x 2 kernel & stride = 2) for enriching and easing FMs. Rectified learning unit (RELU) followed by batch normalization (BN) is a standard configuration used to stabilize network weights and avoids gradient degradation. We also introduced ASPP to improve the training speed further for a large dataset. It increases the receptive field for capturing finer details of the input without raising the number of parameters leading to segmentation efficiency.

E. A Modified Attention-Aware Deep Convolutional Neural Network Architecture

The ability of a model to zero in on specific parts instead of equally treating the entire image makes attention-aware deep neural networks invaluable. Certain features are undoubtedly more crucial than others in identifying DR, which is particularly useful. Our network was built of three sections: segmentation adjustment (SA), lesion-conscious (LC), and DR detection sub-networks based on ResidualNet *(RNet)* and *MaskRCNN* [26].

The SA identified the overall segmentation fitting with pre-trained weights on *ImageNet* initialized on *RNet* to create the base network. Each selection was subtly adjusted for contraction to prevent over segmentation that may cause irregularities. Then, the LC generates mask images to identify different lesions existing in retinal images via quadrant search. As part of this section, we used non-maximal suppression to select the most objective bounding box with the highest accurate scores from among several predicted ones. The process continues until the completion of all boxes. Ultimately, the DR classifications are generated thru combined characteristics from SA and LC. The pre-trained DR-based network fused with the attribute extractor of the LC collects a more detailed and granular representations, conserving needed image information. Fig. 6 exhibits the classification pipeline.

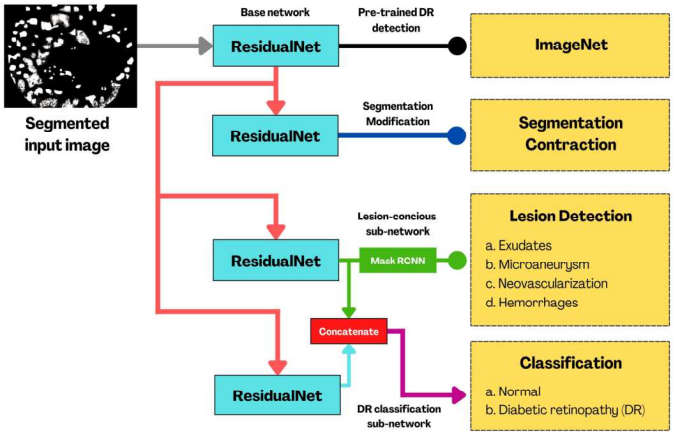

Fig. 6. An attention-aware deep convolutional neural network with three sub-networks for detection of diabetic retinopathy.

F. Evaluation Metrics

To assess the quantitative reliability of segmentation, we used intersection over union (*IoU*) that calculates the ground truth (G) and the algorithm's results (R) by dividing the intersection of two datasets. Another measurement is the Dice similarity coefficient (*DSC*) equivalent which measures the overlaps of two datasets by dividing the sum of the first and second sets with values ranging between 0 and 1. A value of 1 means it is perfectly identical; otherwise, it is not. Equations 1 and 2 display the formulas:

$$IoU = \frac{|G \cap R|}{|G| + |R| - |G \cap R|} \quad (1)$$

$$DSC = \frac{2|G \cap R|}{|G| + |R|} \quad (2)$$

For classification measurements, we evaluated the performance of the deep learning architecture in terms of accuracy, precision, sensitivity, and specificity. These are standard benchmarks in machine learning, as summarized by true positives (TP) and true negatives (TN), false positives (FP), and false negatives (FN) denoted in Equations 3 to 6:

$$Accuracy\ (AC) = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Precision\ (PR) = \frac{TP}{TP + FP} \quad (4)$$

$$Sensivitiy\ (SE) = \frac{TP}{TP + FN} \quad (5)$$

$$Specificity\ (SP) = \frac{TN}{TN + FP} \quad (6)$$

III. Results

We conducted the experiments using the fastest computer available with a Core-i9 processor (8 cores, 3.60 GHz, 16 MB cache, 16 threads), 64 GB RAM, and a graphics processing unit (RTX390Ti, 1.86 GHz & 24 GB). *Keras* and *TensorFlow* libraries allowed the creation of the models in a structured and efficient manner. The succeeding section outlines the results.

A. Attention-Aware DCNN's Hyperparameters Optimization

For a model to work effectively, adjusting the hyperparameters is essential. These settings unlike parameters are set before modeling. We utilized a Bayes search optimization technique to lessen the time-consuming process of manual configuration through trial and errors. Table I shows the fine-tuned options.

TABLE I. Attention-aware DCNN's Optimized Configurations

Hyperparameters	Settings
Batch size	16
Decay	0.001 per epoch
Environment	GPU
Epoch	100
Loss	Cross-entropy (Binary)
Learning rate	0.001
Optimizer	ADAM
Shuffling	Per epoch

B. Segmentation Performance

Table II compared the results of our modified segmentation framework (DR-UNet) with the original U-Net and Res-UNet using 50 random images with different severities of DR. The results show that DR-UNET outperforms the other models with IoU and DSC mean scores of 0.871 and 0.845 in multiple folds, respectively. This demonstrates that our technique can represent identified lesions accurately with reference to the ground truth. Fig. 7 provides examples of the performance of our networks versus the ground truth.

TABLE II. Comparative Segmentation Evaluations

K-fold	DR-UNET		U-NET		RES-UNET	
	IoU	DSC	IoU	DSC	IoU	DSC
1	0.863	0.850	0.721	0.778	0.803	0.818
2	0.882	0.849	0.713	0.743	0.799	0.815
3	0.861	0.843	0.749	0.734	0.813	0.809
4	0.863	0.842	0.811	0.789	0.821	0.813
5	0.871	0.848	0.789	0.812	0.834	0.814
6	0.872	0.853	0.748	0.803	0.817	0.821
7	0.874	0.847	0.758	0.787	0.818	0.820
8	0.861	0.842	0.748	0.791	0.809	0.818
9	0.889	0.839	0.801	0.813	0.818	0.826
10	0.878	0.841	0.812	0.812	0.805	0.821
Mean	**0.871**	**0.845**	0.765	0.786	0.813	0.817

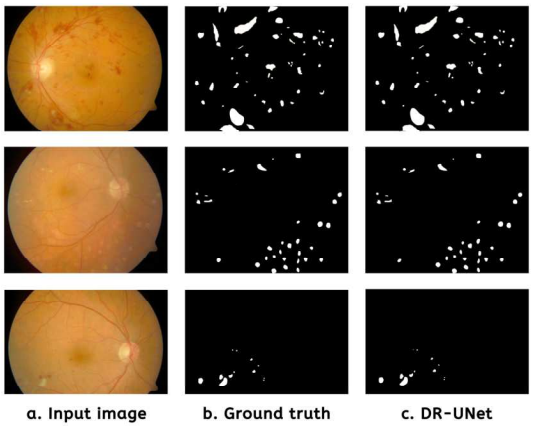

Fig. 7. Segmentation comparison of ground truth vs. DR-UNet.

C. Classification Performance

Table III gauged the performance of our framework to various methods such as conventional DL, ensembles, lesion isolation, multi-sieving, modified *AlexNet*, circular Hough-CNN, and Hessian matrix with altered CNN features. Our proposed pipeline obtained accuracy, precision, sensitivity, and specificity based on the results with 0.992, 0.989, 0.994, and 0.989, respectively. The results had a marked average accuracy increase of 5.64% among existing methods.

TABLE III. COMPARATIVE PERFORMANCE AGAINST DIFFERENT TECHNIQUES

Approach	AC	PR	SE	SP
Conventional DL [27]	0.861	0.790	0.840	0.820
Ensembles [28]	0.869	0.803	0.881	0.855
Lesion isolation [29]	0.920	0.820	0.88	0.831
Multi-Sieving [30]	0.961	0.814	0.978	0.921
Modified AlexNet [31]	0.966	0.953	0.953	0.973
Circular Hough – CNN [17]	0.985	-	0.996	0.982
Hessian matrix + CNN with squeeze, excitation & bottleneck [18]	0.987	0.972	0.996	0.982
Proposed DR-UNet with ASPP + Attention-aware DCNN	**0.992**	**0.989**	0.994	**0.989**

A confusion matrix in Table IV shows that DR detection performs better than normal eye categorization, but with only one error difference.

TABLE IV. CLASSIFICATION CONFUSION MATRIX
(TRAINING = 1930, TESTING = 386 WITH HOLDOUT = 10)

	Normal	DR
Normal	191	2
DR	1	192

We also tested ten fundus images exhibiting early signs of DR to quantify if our model can empirically distinguish the abnormality. The machine identified nine out of ten instances of early DR, illustrated in Fig. 8.

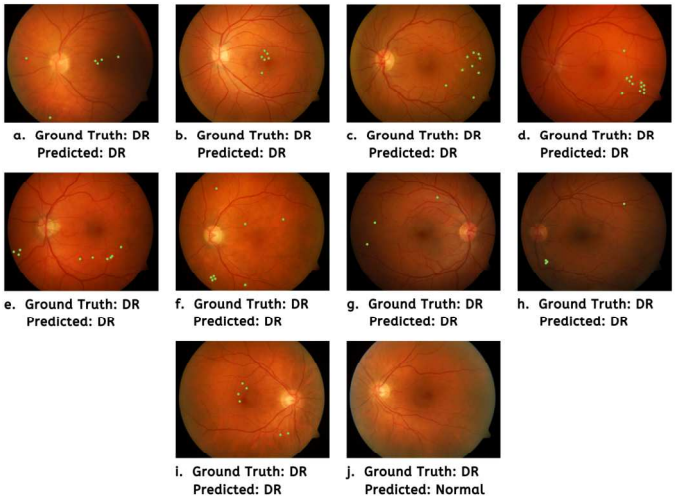

Fig. 8. The amplified segmentation technique and attention-aware DCNN recognized eary signs of DR (a-i) except for (j).

D. Model Training-Validation Loss and Accuracy

As shown in Figure 9(a), both the training and validation loss decreased steadily during the initial stages of the iteration process. These values eventually reached a stable point at the 85th epoch, yielding impressive results. It suggests that our model is effectively fitting the training and validation data. On the next graph, Figure 9(b) reveals an incremental increase in both training and validation accuracy for each epoch. This demonstrates the ability of our proposed DCCN to learn and extract important features from fundus images for diabetic retinopathy classification. Additionally, the convergence of the two lines in both graphs indicates that our modified neural network is neither under nor overfitting.

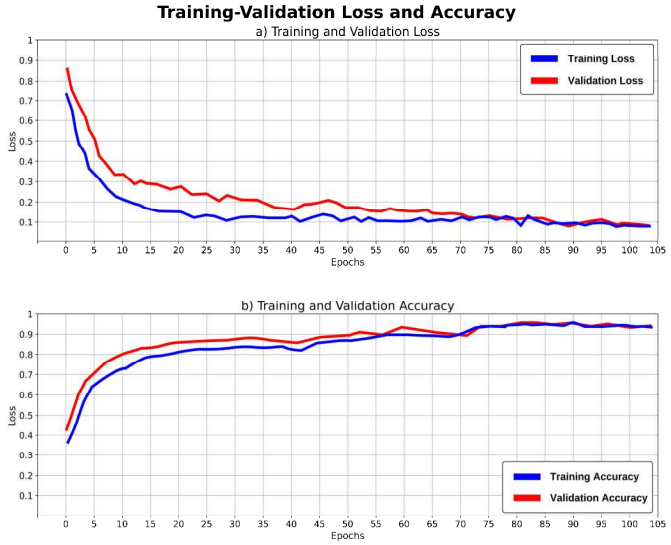

Fig. 9. A gradual convergence are achieved indicating no signs of under or overfitting.

IV. DISCUSSIONS

This study introduced an enhanced segmentation technique and attention-aware DCNN for identifying DR. Based on benchmarks, our DR-UNet with ASPP mechanism outperforms standard U-Net and Res-UNet architecture with IoU and DSC values of 0.871 and 0.845 against the ground truths - a mean improvement of 0.082 and 0.043, respectively. We also verified the robustness of our classification model based on combined training and highly augmented datasets. Compared with different techniques, our proposed network attained higher performance on various metrics by introducing multiple sub-networks - a proven effective strategy for high-level attribute abstractions (accuracy = 99.20%). Furthermore, the model's exceptional performance was validated by its capability to detect DR's early signs according to the validation sets. Lastly, training/validation loss and accuracy plots exhibited a slow but gradual and smooth converging learning rate. Our research has made significant progress in automated medical image processing, building on the work of the advancement in data science [32 – 45].

While conducting this research, we encountered a common challenge of working with noisy and low-quality fundus images that can lead to segmentation and classification errors.

Despite these obstacles, we did not examine methods for improving image quality, such as 'image dehazing' and spatial filtering.

V. Conclusions and Future Works

DR is a severe consequence of diabetes that can cause significant vision loss or even blindness if not detected and treated properly. One of the main reasons the eye's abnormality is hard to diagnose is that it does not often indicate noticeable symptoms. Consequently, affected individuals become aware only once it progresses to the advanced stage. Multiple attempts have been conducted to automate the prognosis using various artificial intelligence-based approaches. However, there is still room for improvement. We developed two strategies to boost lesion detection in DR. Primarily, an enhanced segmentation framework (DR-UNet) with the integration of an ASPP block that extracts more relevant information and discards irrelevant regions from the fundus images. Nevertheless, we observed that it is prone to over-segmentation. To compensate for this dilemma, we incorporated a sub-network in our attention-aware DCCN classifier to contract the lesion segments before classification for adjusted fitting. More importantly, the stacks of *RNets* with pre-trained base, segment contraction, lesion-conscious (with *MaskRCNN*), and concatenation networks proved to be robust in DR's identification, including on its onset.

Our research significantly contributed to the medical image prognosis to identify DR automatically and non-invasively. Artificial intelligence can be a valuable tool in the healthcare industry. However, it should be considered as a supplement to human expertise. Instead, it is most effective when combined with traditional healthcare practices to provide individuals with the most comprehensive care possible. By consolidating the strengths of both approaches, we can create a healthcare system that is more efficient and effective at meeting the patients' needs. The proponents intend to classify DR based on severity and design a better architecture in the future.

Datasets Links

DIARETDB0 (https://bit.ly/3H4guBE), DIARETDB1 (https://bit.ly/3J9O8bO), FIRE (https://bit.ly/3Ja3oFL), and DRIVE (https://bit.ly/3J4PiFI).

References

[1] New England Retina Associates, "Diabetic retinophathy and its kinds", Available at: https://www.retinamd.com/diseases-and-treatments/retinal-conditions-and-diseases/diabetic-retinopathy/, 2022.

[2] N. Santosh and Y. Radhika, "Optimized maximum principal curvatures based segmentation of blood vessels from retinal images", Biomedical Research, 2019.

[3] G. Hassan, Nashwa El-Bendary, A. Hassanien, A. Fahmy, M. Shoeb and V. Snasel, "Retinal blood vessel segmentation approach based on mathematical morphology," Procedia Computer Science, vol. 65, pp. 612-622, 2015.

[4] J. Orlando, E. Prokofyeva and M. Blaschko, "A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images," IEEE Transactions of Biomedical Engineering, 64(1), pp. 16-27, 2017.

[5] R. Manjula Sri, J. Jyothirmai and D. Swetha, "Analysis of retinal blood vessel segmentation in different types of diabetic retinopathy," International Journal of Engineering and Advanced Technology, 8(2), pp. 1-4, 2019.

[6] N. Memari, A. Ramli, M. Saripan, S. Mashohor and M. Moghbel, "Retinal blood vessel segmentation by using matched filtering and fuzzy C-means clustering with integrated level set method for diabetic retinopathy assessment," Journal of Medical and Biological Engineering, pp. 713-731, 2019.

[7] A. Budai, R. Bock, A. Maier, J. Hornegger and G. Michelson, "Robust vessel segmentation in fundus images," International Journal of Biomedical Imaging, 154860, pp. 1-11, 2013.

[8] A. Garifullin, L. Lensu and H. Uusitalo, "Deep Bayesian baseline for segmenting diabetic retinopathy lesions: Advances and challenges," Computers in Biology and Medicine, 136, 104725, 2021.

[9] Z. Tian, L. Liu, Z. Zhang and B. Fei, "Superpixel-based segmentation for 3D prostate MR images," IEEE Transactions of Medical Imaging, 35 (3), pp. 791–801, 2015.

[10] D. Nguyen, S. Benameur, M. Mignotte and F. Lavoie, "Superpixel and multi-atlas based fusin entropic model for segmentation of X-ray images," Medical Image Analysis, vol. 48, pp. 58-74, 2018.

[11] S. Selvaperumal and R. Bhoopalan, "An efficient approach for the automatic detection of hemorrhages in colour retinal images," IET Image Processing, pp. 1550-1554, 2018.

[12] R. Chalakkal, W. Abdulla and S. Thulaseedharan, "Automatic detection and segmentation of optic disc and fovea in retinal images," IET Image Processing, pp. 2100-2110, 2018.

[13] A. Benzamin and C. Chakraborty, "Detection of hard exudates in retinal fundus images using deep learning," Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Kitakyushu, Japan, 2018, pp. 465–469, 2018.

[14] S. Kumar, A. Adarsh, B. Kumar and A. Singh, "An automated early diabetic retinopathy detection through improved blood vessel and optic disc segmentation," Optics & Laser Technology, vol. 121, pp. 1-11, 2020.

[15] P. Shah, D. Mishra, M. Shanmugam, B. Doshi, H. Jayaraj and R. Ramanjulu, "Validation of deep convolutional neural network-based algorithm for detection of diabetic retinopathy – artificial intelligence versus clinician for screening," Indian Journal of Opthalmology, 68(2), pp. 398-405, 2020.

[16] M. Andonova, J. Pavlovicova, S. Kajan, M. Oravec and V. Kurilova, "Diabetic retinopathy screening based on CNN," International Symposium ELMAR, pp. 51-54, 2017.

[17] Kemal Adem, "Exudate detection for diabetic retinopathy with circular hough transformation and convolutional neural networks", Expert Systems with Applications, vol. 114, pp. 289-295, 2018.

[18] S. Das, K. Kharbanda, Scucheta M, R. Raman and E. Das, "Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy," Biomedical Signal Processing and Control, vol. 68, 102600, 2021.

[19] A. Foo, W. Hsu, M. Lee, G. Lim and T. Wong, "Multi-task learning for diabetic retinopathy grading and lesion segmentation", Proceedings of the AAAI Conference on Artificial Intelligence, 2020.

[20] O. Daanouni, B. Cherradi and Amal Tmiri, "Self-attention mechanism for diabetic retinopathy detection," Advances in Science, Technology and Innovation, pp. 79-88, 2021.

[21] Y. Li, Z. Song, S. Kang, S. Jung and W. Kang, "Semi-supervised auto-encoder graph network for diabetic retinopathy grading," IEEE Access, pp. 140759-140767, 2021.

[22] B. Bansode, Bakwad K. M., A. Dildar and Sable G.S., "Deep CNN-based feature extraction with optimized LSTM for enhanced diabetic retinopathy detection," 2022.

[23] J. Gu, X. Sun, Y. Zhang, K. Fu and L. Wang, "Deep residual squeeze and excitation network for remote sensing image super-resolution", Remote Sensing, vol. 11, 2019.

[24] L. Qiao, Y. Zhu and H. Zhou, "Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algoritms," IEEE Access, 2020.

[25] A. Skouta, A. Elmoufidi, S. Jai-Andaloussi and O. Ouchetto, "Hemorrage semantic segmentation in fundus images for the diagnosis of diabetic retinopathy by using a convolutional neural network," Journal of Big Data, 9(78), 2022.

[26] G. Suardika, M. Maysanjaya and M. Kesiman, "Optic disc segmentation based on mask R-CNN in retinal fundus images," 4th International Conference on Biomedical Engineering (IBIOMED), 2022.

[27] T. Kauppi, V. Kalesnykiene, J. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kalviainen and J. Pietla, "DIARETDB1 diabetic retinopathy database and evaluation protocol," Technical report, 2021.

[28] S. Qummar, F. Khan, S. Shah, A. Khan, S. Shamshirband, Z. Rehman, I. Khan and W. Jadoon, "A deep learning ensemble approach for diabetic retinopathy detection," IEEE Access, pp. 150530-150539, 2019.

[29] K. Adal, P. Van Etten, J. Martinez, K. Rouwen, K. Vermeer and L. van Vliet, "An automated system for the detection and classification of retinal changes due to red lesions in longitudinal fundus images," IEEE Transactions of Biomedical Engineering, 65(6), pp. 1382-1390, 2017.

[30] L. Dai, R. Fang, H. Li, X. Hou, B. Sheng, Q. Wu and W. Jia, "Clinical report guided retinal microaneurysms detection with multi-sieving deep learning, IEEE Transaction of Medical Imaging, 37(5), pp. 1149-1161, 2018.

[31] T. Shanti and R. Sabeenian, "Modified Alexnet architecture for classification of diabetic retinopathy images," Computers & Electrical Engineering, vol. 76, pp. 56-64, 2019.

[32] R. Maaliw, A. Alon, A. Lagman, M. Garcia, J. Susa, R. Reyes, M. Fernando-Raguro and A. Hernandez, "A Multistage transfer learning approach for acute lymphoblastic leukemia classification," 13th IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, USA, 2022.

[33] J. Wang, Peiqing LV, H. Wang and C. Shi, "SAR-U-Net: Squeeze-and-excitation block and atrous spatial pyramid pooling based residual U-Net for automatic liver segmentation in computed tomography," Computer Methods and Programs in Biomedicine, 106268, 2021.

[34] X. Zhang, L. Li, D. Di, J. Wang, G. Chen, W. Jing and M. Emam, "SERNet: Squeeze and excitation residual network for semantic segmentation of high-resolution remote sensing images," MDPI Remote Sensing, vol. 14, pp. 1-19, 2022.

[35] R. Maaliw, A. Alon, A. Lagman, M. Garcia, M. Abante, R. Belleza, J. Tan and R. Maaño, "Cataract detection and grading using ensemble neural networks and transfer learning," IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, Canada, 2022.

[36] B. Al-Bander, B. Williams, W. Al-Nuaimy, M. Al-Taee, H. Pratt and Y. Zheng, "Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis," Symmetry, 2018.

[37] V. Badrinarayanan, A. Kendall and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), pp. 2481-2495, 2017.

[38] R. Maaliw, J. Susa, A. Alon, A. Lagman, S. Ambat, M. Garcia, K. Piad and M. Fernando-Raguro, "A deep learning approach for automatic scoliosis cobb angle identification," IEEE World Artificial Intelligence & Internet of Things Congress (AIIoT), Seattle, Washington, USA, 2022.

[39] K. Cha, L. Hadjiiski, R. Samala, H. Chan, R. Cohan, E. Caoili, C. Paramagul, A. Alva and A. Weizer, "Bladder cancer segmentation in CT for treatment response assessment: Application of deep-learning convolution neural networks – A pilot study," MDPI Tomography, 2(4), pp. 421-429, 2016.

[40] P. Coupe, B. Mansencal, M. Clement, R. Giraud, B. D. de Senneville, V. Ta, V. Lepetit and J. Manjon, "AssemblyNet: A large ensemble of CNNs for 3D whole brain MRI segmentation," NeuroImage, 219, 117026, 2020.

[41] M. Goyal, A. Oakley, P. Bansal, D. Dancey and M. Yap, "Skin lesion segmentation in dermoscopic images with ensemble deep learning methods," IEEE Access, vol. 8, pp. 4171-4181, 2020.

[42] S. Guo, K. Wang, H. Kang, Y. Zhang, Y. Gao and T. Li, "BTS-DSN: Deeply supervised neural network with short connections for retinal vessel segmentation," International Journal of Medical Informatics, vol. 126, pp. 105-113, 2019.

[43] T Hofmanninger, F. Prayer, J. Pan, S. Röhrich, H. Prosch and G. Langs, "Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem," European Radiology Experimental," 4(1), pp. 1-13, 2020.

[44] Q. Jin, Z. Meng, T. Pham, Q. Chen, L. Wei and R. Su, "DUNet: A deformable network for retinal vessel segmentation," Knowledge-Based Systems, vol. 178, pp. 149-162, 2019.

[45] F. Yuan, Z. Zhang and Z. Fang, "An effective CNN and transformer complementary network for medical image segmentation," Pattern Recognition, vol. 136, 109228, 2023.