2021 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)

2021 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS) | 978-1-6654-1327-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/HPBDIS53214.2021.9658352

A Dual Weighted Shared Capsule Network for Diabetic Retinopathy Fundus Classification

Grace Ugochi Nneji[1], Jingye Cai[1*], Jianhua Deng[1], Happy Nkanta Monday[2], Saifun Nahar[3], Goodness Temofe Mgbejime[2], Edidiong Christopher James[1], Surafel Kifetew Woldeyes[2]
[1] *School of Information and Software Engineering University of Electronic Science and Technology of China*
[2]*School of Computer Science and Engineering University of Electronic Science and Technology of China*
[3]*School of Computer and Information System University of Missouri St louis MO, USA*
ugochinneji@std.uestc.edu.cn, jycai@uestc.edu.cn, jianhua.deng@uestc.edu.cn, mh.nkanta@std.uestc.edu.cn, snnnm@umsl.com, temofeerics@gmail.com, edianajames@yahoo.com, surafelkifetew1@gmail.com

***Abstract*—Diabetic Retinopathy (DR) is as a result of prolonged high blood glucose level leading to blood vessel complications and loss of eyesight. Early detection could reduce the risk of non-proliferated diabetic retinopathy and death rate. Depending on domain experts for diagnostic results call for a collective judgment and the diagnostic procedure could be time-consuming, hence, there is need for an automatic diagnostic method for the prediction of referable DR or not. This paper proposes siamese capsule network for the classification of DR. At the preprocessing stage, we enhanced the quality of the image by using the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique. Then, the siamese capsule network is applied for distinct features extraction from the preprocessed image and sigmoid classifier is used for the predicted result. Using MESSIDOR dataset of 1200 scans, we achieved accuracy of 99.1%, specificity of 98.0%, and sensitivity of 98.5% as compared to other models.**

***Index Terms*—Diabetic retinopathy, image identification, deep learning, CLAHE, fundus scan, siamese network, capsule network**

I. Introduction

Diabetes occur as a result of excessive growth of glucose in the blood system [1]. A person diagnosed with diabetes is exposed to other ailments such as kidney malfunction, stroke, heart failure, and so on [2]. Diabetes can greatly affect the vascular system of the retina leading to a structural change. This change could provide physical cues for the disease, hence the domain experts recommend yearly retinal screening test using the dilated eye exams or retinal fundus scans [3]. Significantly, these scans of retina could be used as a means to identify diabetes but it demands for general judgment from the ophthalmologists which could be time consuming. To alleviate this challenge, there is need for an automated system which could reduce the stress of the ophthalmologist as well as diagnose patients within short range of time.

Numerous studies have reported the diagnosis of diabetes based on specific lesions or clinical markers such as Glucose, and HbAIs [2], [3] or even handcrafted features or variables measured by medical professionals. Despite the recent outcome of these developments, DR remains challenging due to some factors such as quality, noise, artifacts, and so on. These interference may influence the performance of the DR task. In replacing the handcrafted feature, deep learning provides a satisfactory solution to these issues. In recent years, the application of deep learning has proven to have the capability of extracting distinct features directly from medical dataset [4] - [7]. In the classification and detection of medical images, the application of Convolutional Neural Network (CNN) has been widely used and has achieved greater accuracy performance.

A dual weighted shared capsule neural network is proposed for the DR classification. Our article introduce the use of CLAHE for the image pre-processing and siamese capsule network with pre-trained VGG16 as the base model to automatically identify fundus scans as RDR or not. This model takes in paired binocular scans of both eyes after being pre-processed as input and produces a classification result of each eye as the output at the same time. The contribution of our work is in four-fold.

- A pair of CLAHE pre-processed images are used as input for the DR task.
- A siamese capsule network while using pre-trained VGG16 for the feature extraction is utilized.
- We replace the usual euclidean distance and replaced with the concatenated layer and fully connected layer.
- A comparative study was carried out to check the performance of a siamese model and single model for DR classification.

The structure of this paper is organized as follows. The related work is showed in Section II. The methodology of our proposed model is described in Section III. The experimental results and evaluation of our model are explained in Section IV. Lastly, discussion and conclusion are written in Section V.

II. Literature Review

Deep learning (DL) frameworks have been reported to have the ability to learn distinctive features directly from fundus scans. Numerous works have provided different models for obtaining DR classification.

Gulshan et al. [8] proposed the detection of RDR using InceptionV3 and achieve 99.1% and 97.5% on two different dataset respectively. Shankar et al. [9] presented HPTI-v4

∗ Corresponding author.

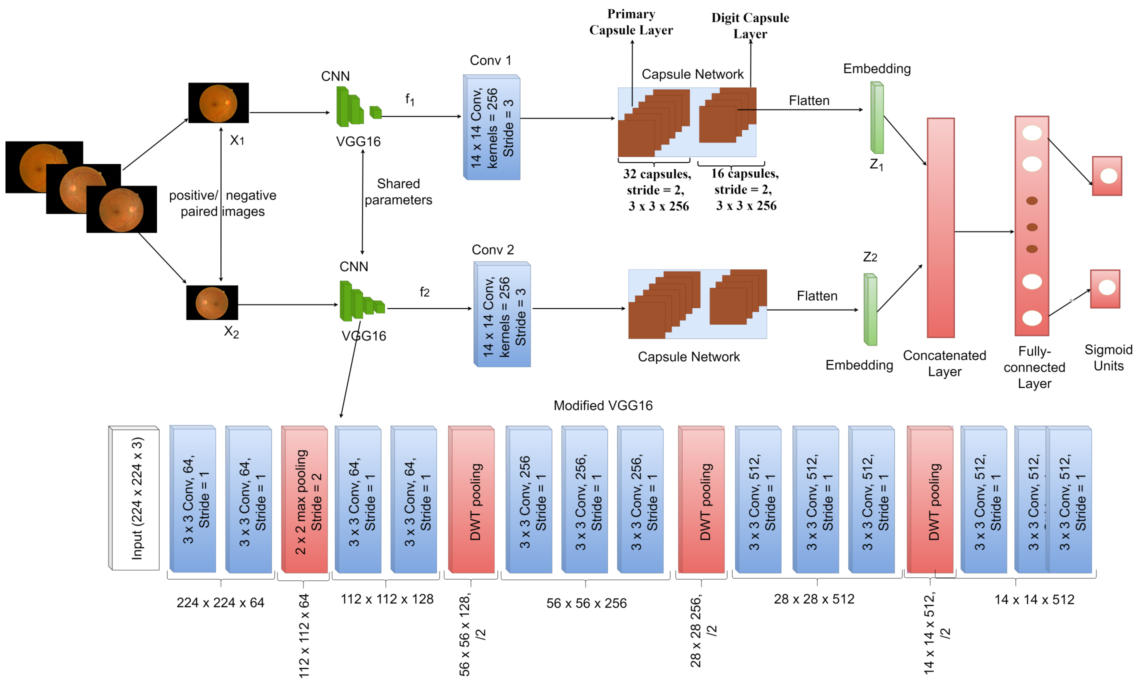

Fig. 1. Our proposed siamese capsule network with wavelet pooling for the classification of DR

model for DR classification achieved 99.49% in accuracy, 98.83% in sensitivity and 99.68% in specificity. A combination of machine learning (ML) and DCNN(ResNet) was reported in [10] and decision tree was used to classify between healthy fundus and fundus with RDR achieving sensitivity of 93.0%, specificity of 87.0% and AUC of 94.0%. Shanthi et al. [11] presented a modified AlexNet algorithm using Messidor dataset achieving average accuracy of 96.25%. Authors in [12] employed DCNN to segment and detect DR lesions. New biomarkers were discovered with the use of heatmaps. A good detection of Free-Response Receiver Operating Curve (FROC) of 95.4% was achieved with kaggle dataset. A report in [13] applied Inception-ResNet v2 for feature extraction with Moth Optimization DNN technique using the Messidor dataset and achieved 99.12% accuracy, sensitivity of 97.91% and specificity of 99.47%. Researchers in [14] designed an alternative dual classification process with a bootstrapped decision tree which was applied for the segmentation of fundus scans. The authors in [15] created a CNN architecture by extracting subtle local features with the technique of speed-up robust properties and Bag of Visual Words. A pretrained CNN framework with attention and crop network was used in [16] to detect patch sites also known as Zoom for DR detection while using the MESSIDOR dataset. Another study in [17] proposed SI2DRNet-v1 by rescaling the kernel size after subsequent pooling layer in CNN model. In conclusion, DL models for DR classification outperform the handcrafted features technique. However, only a few research works focus on siamese network and CLAHE for improving the quality of the images.

III. Methodology

In this phase, we explain in detail the network architecture and the image enhancement technique follow by the implementation details in the subsections.

A. Network Architecture and Implementation

A siamese capsule network is proposed in this paper as seen in Figure 1. Primarily, the model accepts two fundus scans corresponding to the left eye and right eye as inputs and then transmits them into a dual framework. The information gathered are fed into the fully-connected layer and lastly the architecture predict the result as with or without proliferative DR. The approach of our proposed model is based on a real-life clinical diagnose of DR. The model layers and the illustration of the architectural framework is given in the subsequent subsections.

1) Capsule Network: For classification problems, the authors in [18] were the first to propose capsule network. A capsule network is made up of entity-oriented vectorial capsules, unlike traditional CNNs that employ scalar neurons to represent the probabilities of the presence of specific features. In retrieving intrinsic and distinguishing properties of entities, capsule networks are far more stronger and robust than ordinary CNNs [18], [19].

A key characteristic of this capsule formulation which is impossible to achieve with CNN scalar neuron models is that the vectorial representation permits a capsule to learn and detect variants of a feature. The overall input in the convolutional capsule layers is the weighted summation of all predictions gotten from the capsules inside the convolutional kernel as seen in Equation (1).

$$C_j = \sum_i a_{ij} \ . \ U_{j/i} \tag{1}$$

Where Cj represents the sum total input to capsule j; the indicator that depicts the degree of which capsule i activates capsule j is called the coupling coefficient a_{ij}, $U_{j/i}$. The prediction Uj/i from capsule i to capsule j is illustrated below in Equation (2).

$$U_{j/i} = W_{ij} \ . \ U_i \tag{2}$$

Ui denotes the capsule $i's$ output and Wij represents the weight network on the edge linking capsules i and j. A robust routing method [19] decide the coefficient between capsule i and other linked capsules in the above layer summing to 1. To ignite another capsule, the robust routing method known as routing by agreement considers both the length of a capsule and its instantiation parameters. This differs from traditional CNN models, which simply evaluate probability. As a result, capsule networks are reliable and more powerful at abstracting inherent object characteristics. It is worth mentioning that the capsule's length is used to determine the likelihood of an entity's presence. Non-linear "squashing' function is used as the activation function for the convolutional capsule layers in ensuring a perfect probability estimation where capsules with short vectors have low probability, otherwise high probability all at maintaining a constant orientation. Squashing function is given in Equation (3).

$$U_j = \frac{||C_j||^2}{1 + ||C_j||^2} \cdot \frac{C_j}{||C_j||} \tag{3}$$

2) Proposed Siamese Capsule Network: In a real-life scenario, when a patient takes fundus scans in the medical center, both eyes will be captured. Then, comparison of both eyes scans will be checkmated in order to know the condition of the eyes as one eye becomes a vital guide to the screening of the other eye. Hence, this becomes the inspiration of using a siamese capsule network that accepts both fundus images, access their features and predict if with RDR or not. Generally, siamese network is built for the comparison of two inputs images sharing the same weight and parameters. Then, the extracted feature vectors of the two inputs are gotten from the last layer and euclidean distance will be calculated. If the distance between both images are large, then they are dissimilar, otherwise similar.

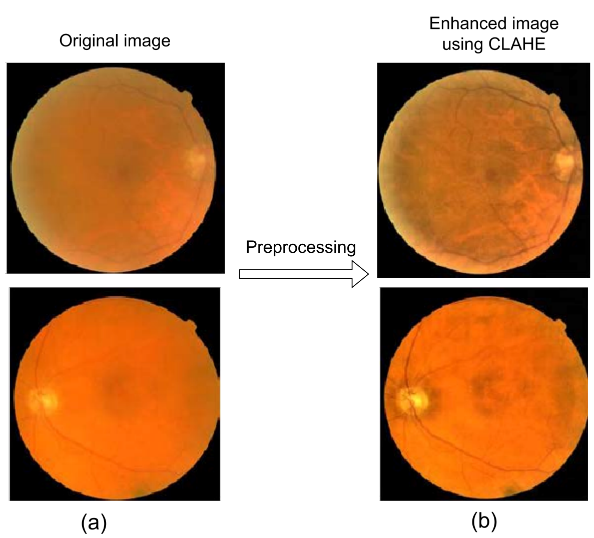

Fig. 2. The result of the retinal image enhancement using CLAHE. (a) depicts sample of the original images and (b) depicts the CLAHE images

TABLE I
DESCRIPTION OF MESSIDOR DATASET

DR stages	Details	Number	Label
Healthy	Zero abnormalities	548	Normal
Mild NPDR	Microaneurysms	152	Stage 1
Moderate NPDR	Few microaneurysms	246	Stage 2
Severe NPDR	Venorous beading + Intraretinal microvascular abnormality		
PDR	Vitreous/Pre-retinal hemorrhage	254	Stage 3

In our study, calculating the similarity of the fundus scans will be eliminated, instead the relationship between both images will be used for the prediction process. Therefore, in our work the calculation part for the distance is taken away as seen in our Figure 1. Two dual-weighted capsule network with VGG16 as the backbone is adopted for the feature extraction, then the extracted features will be merged together and fed into the fully-connected layer and finally using sigmoid classifier for the prediction of both fundus images. The prediction result is between 0 to 1 which represents the probability of which eye has RDR or not.

In this study, the input size of the preprocessed images are resized to $224 \ x \ 224 \ x \ 3$ and fed into network. We utilized VGG16 model as the feature extractor due to its reasonable network depth without performance degradation and better results on image classification task. VGG16 has 13 convolutional layers arranged in blocks, 3 fully connected layers and 5 max-pooling layers. To achieve better classification performance and maintain the integrity of high-level features, it is necessary to extract the high-level features of the fundus images by removing the max-pooling layers in the second, third, fourth and replace it with discrete wavelet transform (DWT) pooling layers in order to reduce the loss of spatial details. More so, the fifth block max-pooling and 3 FC layers were removed and arrive at an output of $14 \times 14 \times 512$ feature vector as seen in Figure 1. Preserving the spatial details of the extracted features, f_1 and to reduce the vector dimension as well as to maintain dimensionality match with the primary capsule layer, we added a convolutional layer of 14×14 with kernel size of 256 and stride of 3. In our proposed model, the spatial details

TABLE II
THE DISTRIBUTION OF RELABELED DATA

Label	Class	Number	Percentage
0	No RDR	548	45.6%
1	RDR	652	54.4%

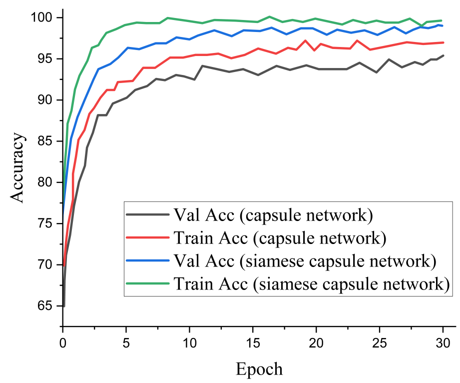

Fig. 3. Training accuracy and validation accuracy performance of two models- the single capsule network and the siamese capsule network

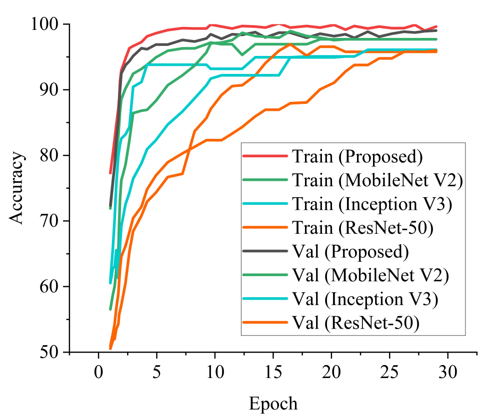

Fig. 4. Training and validation curves showing the performance of our proposed siamese model and the pre-trained models.

are transformed into Primary Capsule layers (PrimaryCaps) in form of capsules after the feature extraction stage. Also, the max-pooling layer in capsule performs feature down-sampling in order to reduce the network size. However, we adopted DWT pooling of K_m x K_m to replace max-pooling layer and stride of Km direct to the feature maps of the last convolutional capsule layer. Only capsule with the longest vector are kept while the rest are discarded in the K_m x K_m kernel. To this effect, the size of the network and capsule number are reduced and leaving a selection of the essential capsules. Routing by agreement is introduced to learn the spatial details in form of transformation matrix. The connection strength of the digit capsule is controlled by the routing coefficient. A high-level entity abstraction with a global view is flatten to arrive at the feature embeddings. More so, dynamic routing between two connected capsule layers is utilized and the squashing function is used to normalize capsule outputs. The probability that the capsule instance exists is represented by the length of the feature vector which is compressed to 1 with the help of a non-linear function. We added a convolutional layer to each branch of the network in order to reduce the dimension.

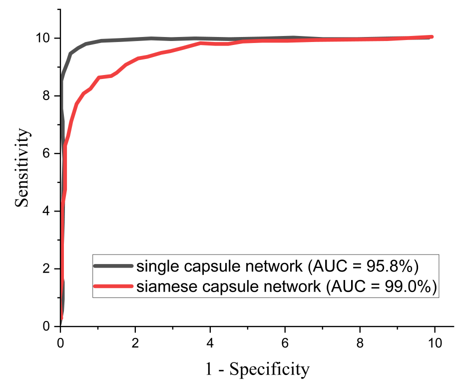

Fig. 5. ROC of two models - the single capsule network having an AUC of 95.8% and the siamese capsule network having an AUC of 99.0%

We introduce a regularization term to enhance the robustness of the model. Then, these embedding features will be merged together and connected with a fully-connected (FC) layer for feature integration and the network will decide and predict the result of each eye. The last FC layer is a sigmoid activation function for the prediction and classification of the two eyes respectively which will be continuous between the probability value of 0 to 1 denoting if the eye is with referable DR or not.

B. Image Enhancement Technique

The enhancement of image is an important technique as it highlights the important details of an image and eliminates certain tertiary information in order to improve the classification quality of the image. Contrast-Limited Adaptive Histogram Equalization (CLAHE) is more natural in appearance and helpful in the elimination of noise amplification among the histogram equalization family and we have considered CLAHE and applied it on our dataset. To display the effectiveness of CLAHE technique, a detailed explanation is given below.

The first stage of CLAHE technique involves the development of the image transformation by employing bin value of histogram. Thereafter, the contrast is limited to a binary count from 0 to 1 using clip boundary. The clip boundary is added to the image before the image segment is processed. Specific bin value of histogram region is applied to produce the whole image region in order to avoid mapping background area to gray scale. To achieve better mapping, clip boundary is applied with the help of histogram clip. Finally, the finished CLAHE image is produced by computing the regions of the image after which all the image pixels are extracted, mapped and interpolated to achieve optimal efficiency from the image. The result of the retina image enhancement using CLAHE is presented in Figure 2

IV. Experimental Results and Evaluation

A. Model Hyper-Parameters

In this study, we trained our model for DR classification using a siamese capsule network while using Adam optimizer

TABLE III
COMPARISON OF OUR PROPOSED MODEL WITH OTHER METHODS USING FUNDUS CLASSIFICATION (FC) ON MESSIDOR DATASET.

Authors	Training type	Process type	ACC (%)	AUC (%)	SEN (%)
Chen et al. [13]	SI2DRNet	FC	91.2	96.5	-
Costa et al. [11]	SURF + CNN	FC	-	90.0	-
Gargeya et al. [6]	CNN	FC	-	94.0	-
Gulshan et al. [4]	CNN	FC	-	99.0	87.0
Wang et al. [12]	CNN	FC	91.1	95.7	-
Ours	Siamese CapsNet	FC	99.1	99.0	98.5

TABLE IV
THE EVALUATION RESULT OF THE SINGLE CAPSULE NETWORK AND SIAMESE CAPSULE NETWORK.

Model	ACC (%)	SEN (%)	SPE (%)	AUC (%)	Time (min)
Single capsule network	95.3	95.0	96.0	95.8	20.9
Proposed siamese capsule network	99.1	98.5	98.0	99.0	21.7

TABLE V
COMPARISON OF OUR MODEL WITH SELECTED PRE-TRAINED MODEL ON THE SAME MESSIDOR DATASET.

Model	ACC (%)	SEN (%)	SPE (%)	AUC (%)	Time (min)
ResNet-50	95.2	94.8	94.1	95.9	19.8
Inception V3	96.8	95.3	95.7	96.5	19.3
MobileNet V2	97.8	96.9	96.2	97.6	18.8
Ours	99.1	98.5	98.0	99.0	21.7

TABLE VI
PERFORMANCE EVALUATION OF OUR PROPOSED SIAMESE MODEL BASED ON DIFFERENT HYPERPARAMETER TUNING ON MESSIDOR DATASET.

Hyperparameter	Proposed model + Wavelet pool + BN + Adam	Proposed model + Max-pool + BN + Adam
	Accuracy (%)	Accuracy (%)
LR (0.0001) + Dropout (0.25)	98.0	95.2
LR (0.0001) + Dropout (0.50)	99.1	96.9
LR (0.0001) + Dropout (0.75)	98.8	96.3

and dropout to avoid over-fitting. A learning rate of 0.0001 is utilized for optimum performance and early stop technique is introduced to monitor the training progress of the model and to terminate the process when the model repeatedly outputs constant values for a prolong number of epochs. In this work, we have examine how the model performance can be controlled by tuning some hyper-parameters such as dropout, pooling method, learning rate (LR), batch normalization (BN) and optimizer (Adam). Also, we applied some evaluation metrics in terms of accuracy, sensitivity, specificity and AUC on our proposed model.

B. Dataset

In this study, we experimented using MESSIDOR dataset [20]. The dataset presented in Table I has been classified originally into four stages. The 1200 fundus scans were gathered by 3 ophthalmologic departments and ranges in different pixels. The scans with microaneurysms belong to Stage 1. The scans found in both microaneurysms and hemorrhages belong to Stage 2 and finally high microaneurysms and hemorrhages belong to Stage 3. For this study, the dataset is reconstructed as presented in Table II, where the healthy scans are label as 0 under No RDR class with a total of 548 images whereas the Stage 1, Stage 2, and Stage 3 in Table I are merged together and label as 1 under RDR class with 652 images as presented in Table II. It is obvious that the dataset is not perfectly balance so we did few augmentation to balance the dataset. Then for the data split, we used 80% for training and 20% for testing set with 30 epoch. All implementations were carried out in Keras framework.

As illustrated in Figure 3 and Figure 4, our proposed model achieves satisfactory performance. Figure 3 displays the validation and training accuracy of the single capsule network and the siamese capsule network. Thus, our proposed model performs better than the single capsule network with a higher margin. More so, our proposed siamese model shows great potential in clinical application as depicted in Figure 4. The ROC-AUC curve shows that our proposed model achieves AUC of 99.0% which is higher than the single capsule network with AUC of 95.8% as shown in Figure 5.

C. Result Analyses

Table III shows a comparative study of other researches using the MESSIDOR dataset, and illustrates that our method outperforms other models. Table IV presents the results in terms of accuracy, sensitivity, specificity and AUC of our proposed siamese capsule network and single capsule network. We also conducted another experiment to showcase the efficacy of our proposed model by running some DL pre-trained models on the same dataset using their source codes available online. Table V shows that our model outweighs the other models across all the metrics. Figure 4 shows the convergence behaviour of our proposed model in comparison with the pre-trained models. The result indicates that our proposed algorithm is effective.

It is worth mentioning that our proposed model responds positively in terms of performance to hyper-parameter tuning as depicted in Table VI. More so, on the verge to further evaluate the wide application of our proposed network on other medical diseases as well as the impact of hyper-parameter tuning on the network performance, we considered a common medical disease known as pneumonia disease [21], [22] as ablation study. In this case, we considered a two-category of our dataset which is pneumonia and non-pneumonia. We merged COVID-19 pneumonia, viral pneumonia and bacterial pneumonia as the pneumonia class verses non-pneumonia

TABLE VII
COMPARISON OF OUR MODEL WITH SELECTED PRE-TRAINED MODEL ON THE PNEUMONIA DATASET [21], [22].

Model	ACC (%)	SEN (%)	SPE (%)	AUC (%)	Time (min)
ResNet-50	96.0	95.1	96.2	96.3	21.4
Inception V3	95.2	95.9	94.8	95.2	22.8
MobileNet V2	96.8	95.4	95.9	96.8	20.3
Ours	98.1	98.7	97.9	98.5	24.7

class (labelled as healthy). Table VII shows the effectiveness of our proposed network in comparison with selected pre-trained models when applied to pneumonia disease.

V. Discussion and Conclusion

An effective DR classification using a dual-weighted siamese capsule network has been proposed. This involved the pre-processing and augmentation of MESSIDOR dataset and using pre-trained VGG16 as the backbone model for the extraction of features. The proposed siamese capsule network architecture accepts a pair of fundus scan as inputs and predict the possibility of the presence of RDR or not in each eye. When we introduced batch normalization, 50% dropout, and wavelet pooling as the pooling method, the evaluation results show that the proposed model achieves better performance with 99.1% accuracy, sensitivity of 98.5%, specificity of 98%, and AUC of 99.0% than just using a straight/single capsule network for diabetic retinopathy classification. These results could effectively assist the ophthalmologists to diagnose the presence of RDR or not while saving screening time. For the same hyper-parameter tuning, our proposed model achieved satisfactory performance when applied to pneumonia disease as seen in Table VII.

Although, capsule network requires more computational time as compared to traditional CNN models. In order to reduce the computational time, we introduce wavelet pooling to replace max-pooling for dimensionality reduction while avoiding loss of spatial information. The proposed model achieved superior performance in terms of classification accuracy, although the computational time of our proposed siamese model is slightly higher as compared to the single model and the other CNN models adopted for comparison in this study. Despite the fact that our proposed model is slightly higher in computational time as depicted in Table IV, Table V, and Table VII our proposed model still outweighs the single model and the other CNN models across all the evaluation metrics. This can be considered as a trade-off between accuracy and computational time. Utilization of batch normalization and dropout to normalize and prevent over-fitting resulted to the model's steady convergence as depicted in Figure 3.

In our future work, we plan to experiment more by considering:

- The application of the same technique on various medical diseases.
- Increasing the dataset size and perform more hyper-parameter tuning.

References

[1] S. D. Solomon, E. Chew, E. J. Duh, L. Sobrin, J. K. Sun, B. L. VanderBeek, C. C. Wykoff, and T. W. Gardner, "Diabetic retinopathy: A position statement by the American diabetes association," Diabetes Care, vol. 40, no. 3, pp. 412-418, Mar. 2017.

[2] A. Bora, S. Balasubramanian, B. Babenko, S. Virmani, S. Venugopalan, A. Mitani, G. de Oliveira Marinho, J. Cuadros, P. Ruamviboonsuk, G. S. Corrado, L. Peng, D. R. Webster, A. V. Varadarajan, N. Hammel, Y. Liu, and P. Bavishi, "Predicting the risk of developing diabetic retinopathy using deep learning," Lancet Digit. Health, vol. 3, no. 1, pp. e10-e19, Jan. 2021.

[3] S. Musleh, T. Alam, A. Bouzerdoum, S. B. Belhaouari, and H. Baali, "Identification of potential risk factors of diabetes for the qatari population," in Proc. IEEE Int. Conf. Informat., IoT, Enabling Technol. (ICIoT), Feb. 2020, pp. 243-246.

[4] G. U. Nneji et al., "A Super-Resolution Generative Adversarial Network with Siamese CNN Based on Low Quality for Breast Cancer Identification," in 2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), 2021, pp. 218–223, doi: 10.1109/PRAI53619.2021.9551033.

[5] H. N. Monday et al., "The Capability of Multi Resolution Analysis: A Case Study of COVID-19 Diagnosis," in 2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), 2021, pp. 236–242, doi: 10.1109/PRAI53619.2021.9550802.

[6] G. U. Nneji et al., "Enhancing Low Quality in Radiograph Datasets Using Wavelet Transform Convolutional Neural Network and Generative Adversarial Network for COVID-19 Identification," in 2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), 2021, pp. 146–151, doi: 10.1109/PRAI53619.2021.9551043.

[7] H. N. Monday et al., "Improved Convolutional Neural Multi-Resolution Wavelet Network for COVID-19 Pneumonia Classification," in 2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), 2021, pp. 267–273, doi: 10.1109/PRAI53619.2021.9551095.

[8] V. Gulshan et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," Jama, vol. 316, no. 22, pp. 2402-2410, 2016.

[9] Shankar et al. "Hyper-parameter tuning deep learning for diabetic retinopathy fundus image classification". IEEE Access. vol. 8, pp.118164-118173, Jun. 2020.

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

[11] T. Shanthi and R. S. Sabeenian, "Modified alexnet architecture for classification of diabetic retinopathy images," Comput. Electr. Eng., vol. 76, pp. 56 - 64, Jun. 2019.

[12] G. Quellec, K. Charrière, Y. Boudi, B. Cochener, and M. Lamard, "Deep image mining for diabetic retinopathy screening," Med. Image Anal., vol. 39, pp. 178-193, Jul. 2017.

[13] K. Shankar, E. Perumal, and R. M. Vidhyavathi, "Deep neural network with moth search optimization algorithm based detection and classification of diabetic retinopathy images," Social Netw. Appl. Sci., vol. 2, no. 4, p. 748, Mar. 2020.

[14] M. Omar, F. Kheli, and M. A. Tahir, "Detection and classification of retinal fundus images exudates using region based multiscale LBP texture approach," in Proc. Int. Conf. Control, Decis. Inf. Technol. (CoDIT), Apr. 2016, pp. 6-8.

[15] P. Costa and A. Campilho, "Convolutional bag of words for diabetic retinopathy detection from eye fundus images," IPSJ Transactions on Computer Vision and Applications, vol. 9, no. 1, p. 10, 2017.

[16] Z. Wang, Y. Yin, J. Shi, W. Fang, H. Li, and X. Wang, Zoom-in-Net: Deep Mining Lesions for Diabetic Retinopathy Detection, Springer, International Conference on Medical Image Computing and Computer-Assisted Intervention, 2017.

[17] Y.-W. Chen, T.-Y. Wu, W.-H. Wong, and C.-Y. Lee, "Diabetic retinopathy detection based on deep convolutional neural networks," in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1030–1034, Calgary, AB, Canada, April 2018.

[18] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," arXiv Prepr. arXiv1710.09829, 2017.

[19] M. E. Paoletti et al., "Capsule networks for hyperspectral image classification," IEEE Trans. Geosci. Remote Sens., vol. 57, no. 4, pp. 2145–2160, 2018.

[20] Messidor-ADCIS.[Online]. Available: http://www.adcis.net/en/third-party/messidor/

[21] RSNA Pneumonia Detection Challenge — Kaggle [Internet]. [cited 2021 October 20]. Available from: https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data

[22] T. Rahman et al., "Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images," Comput. Biol. Med., vol. 132, p. 104319, May 2021, doi: 10.1016/j.compbiomed.2021.104319.