A Classification System for Diabetic Retinopathy

Prerna Gupta[1], Dr. Hari Om Sharan[2], Dr. C. S. Raghuvanshi[2]
1- Research Scholar – Rama University Uttar Pradesh, Kanpur, India
2- Professor – Rama University Uttar Pradesh, Kanpur, India
drsharan.hariom@gmail.com

ABSTRACT:

Diabetic retinopathy is one of the most common side effects of diabetes and the primary cause of blindness (DR). The development of the sickness can be stopped if DR is discovered early. Due to differences in the distribution of medical conditions and poor labour productivity, the best window for diagnosis and treatment was lost, which causes eyesight deterioration. Neural network models can be used to categorise and diagnose DR, enhancing efficiency and reducing costs. Three hybrid model structures were created in this study—Hybrid-a, Hybrid-f, and Hybrid-c—to improve the performance of DR classification models together with an improved loss function. EfficientNetB4, EfficientNetB5, NASNetLarge, Xception, and InceptionResNetV2 CNNs were the main models. These basic models were trained using enhance cross-entropy loss and cross-entropy loss, respectively. The output of the basic models was used to train the hybrid model structures. By using enhanced cross-entropy loss, it is possible to greatly speed up the training of the basic models and enhance their performance when evaluated using various metrics. The recommended hybrid model architectures can also improve the performance of DR classification. The accuracy of DR categorization was improved by adopting hybrid model structures.

Keyword— Artificial Intelligence, Computer-aided diagnosis, Hybrid Model, Diabetic retinopathy, Convolutional neural network

DOI Number: 10.14704/nq.2022.20.13.NQ88438 **NeuroQuantology2022;20(13): 3570-3578**

1. INTRODUCTION:

Diabetic retinopathy (DR), which damages the retina, is a disorder brought on by diabetes. Longer-term diabetic patients are more susceptible to developing diabetic retinopathy. No DR, mild DR, moderate DR, severe DR, and proliferative DR are the five severity levels for DR. Mild, moderate, and severe Nonproliferative Diabetic Retinopathy (NPDR) stages are available. When a patient exhibits no overt symptoms, they are at the NPDR stage. To detect NPDR, a trained ophthalmologist must examine the fundus. As things go worse, the proliferative DR (PDR) stage of DR will take place. During the PDR stage, abnormal new blood vessels form in the rear of the eye. These fragile blood arteries are vulnerable to rupturing and bleeding, which can obstruct vision and ultimately lead to blindness. The NPDR phase of DR treatment has so far proven to be the most effective. Therefore, frequent fundus examination screening is the most effective method for identifying early abnormal symptoms of DR in diabetic individuals. Early identification and appropriate treatment are crucial to helping people avoid DR [1]. To test for diabetic retinopathy, ophthalmologists must have professional clinical knowledge, skill, and diagnosis time. Ophthalmologists often need to perform a direct examination of the patient's fundus and integrate the fundus retinal pictures taken by specialized equipment in order to assess the severity of the patient's diabetic retinopathy. This procedure will take some time to finish. In addition, the expected number of patients requires the services of far too few skilled ophthalmologists. The automated algorithm for classifying the severity of diabetic retinopathy is thus essential for improving the accuracy of DR diagnosis. Fundus images, which are the main images utilized to analyze DR [2,3], are a current research hotspot. For DR detection and categorization, machine learning and algorithms are applied in various studies [4–7]. Nevertheless, due to their success in numerous contests, deep learning algorithms are being applied for DR detection and classification in an increasing number of researches. In this work, CNNs were employed to categories the DR severity of fundus images from start to finish. In a study, Li et al. reported a novel cross-disease attention network (CANet) to simultaneously grade DR and DME. They created a disease-dependent attention module and a disease-specific attention module to extract useful features. Their network has an accuracy of 92.6% and an AUC of 96.3% when classifying DR in the Messidor database. Shanthi and Sabeenian [9] created a modified AlexNet architecture [10] with the incorporation of suitable Pooling, Softmax, and Rectified Linear Activation Unit (ReLU) layers to obtain a high level of accuracy for the categorization of DR fundus images according to the severity of the illness. To verify the effectiveness of the recommended algorithm, they also employed the Messidor database [11]. Finally, the proposed method achieved a classification accuracy of 96.6% on the Messidor database. In a study, Hosseinzadeh et al. [12] developed a new feature extraction technique using a modified Xception architecture for the detection of DR situation. When compared to the original Xception architecture, the updated Xception architecture they presented enhanced DR categorization with classification accuracy of 83.09% versus 79.59%, sensitivity of 88.24% versus 82.35%, and specificity of 87.00% versus 86.32%. The solution that is being suggested combines multilevel features from several convolutional layers of the Xception architecture via deep layer aggregation. Li et al. [13] created four convolutional networks with multistage inputs by extending a basic network. The suggested method evaluated the severity of diabetic retinopathy on the EyePACS dataset and produced a brand-new, cutting-edge kappa score. By employing a hierarchical pruning method, Hajabdollahi et al. modified the original VGG16-Net [15] in the study [14] to reduce the model's structural complexity for DR analysis. The proposed method was tested using the Messidor database and 35% of the feature maps of the VGG16-Net, with merely a loss of accuracy of 1.89%. The performance of the CNN for 2 classes and 5 classes of DR classification utilizing 3 distinct CNN architectures, including VGG16, VGG19, and InceptionV3 [17], was examined by Jain et al. [16]. They found that the CNN's performance directly correlates with the number of convolutional and pooling layers present. For the two classes of DR classification, VGG19 achieved an accuracy of 80.40%.

The main contributions of this paper are as follows: An increased loss function, the enhance cross-entropy (E-CE) loss function, is to improve the performance of basic DR classification models. Three proposed hybrid model structures are to fuse various basic models for better DR classification performance. The fundus images used in this investigation underwent initial pre-processing.

During the training of the fundamental models, data augmentation approaches were used to increase the number of samples and the variety of samples for the DR fundus dataset.

Additionally, utilizing cross-entropy (CE) loss and E-CE loss, multiple fundamental models were trained. The output attributes of the superior basic models were then combined in a variety of ways to train the hybrid model structures. The outcomes showed that hybrid model architectures perform even better when compared to simple models.

2. MATERIALS AND METHOD:

The proposed algorithm graph for this work is shown in Figure 1. The graph is composed of fundus image preprocessing, fundamental CNN model prediction, hybrid model structures prediction, and DR grade output. Preprocessing would be done on the fundus pictures initially. Each fundamental CNN model then forecasted the preprocessed fundus pictures. Additionally, the outputs from each basic CNN model were input into the hybrid model structures. The fundus image's outcome is determined by the DR grade with the highest probability after the hybrid model outputs five anticipated values that are related to the likelihood of the five DR grades. The proposed algorithm graph for this work is shown in Figure 1. The graph is composed of fundus image preprocessing, fundamental CNN model prediction, hybrid model structures prediction, and DR grade output. Preprocessing would be done on the fundus pictures initially. Each fundamental CNN model then forecasted the preprocessed fundus pictures. Additionally, the outputs from each basic CNN model were input into the hybrid model structures. The fundus image's outcome is determined by the DR grade with the highest probability after the hybrid model outputs five anticipated values that are related to the likelihood of the five DR grades.

2.1 Dataset:

The three datasets that make up the dataset for this work are APTOS 2019 Blindness Detection organized by the 4th Asia Pacific Tele-Ophthalmology Society [19] and Deep DR Diabetic Retinopathy Image Dataset provided by the IEEE International Symposium on Biomedical Imaging (ISBI) 2020 [20]. All of them came from Eye PACS. The Eye PACS dataset consists of 35,126 training fundus photos and 53,576 test fundus images. The APTOS dataset consists of 3,662 training fundus photos and 1,928 test fundus images. The Deep DR dataset consists of 1,200 training fundus photos, 400 validation fundus images, and 400 test fundus images. Each fundus image from the three datasets was assessed for the presence of diabetic retinopathy on a scale of 0 to 4: 0 indicates no DR, 1 indicates mild DR, 2 indicates moderate DR, 3 indicates severe DR, and 4 indicates proliferative DR. Examples of DR fundus images with varying degrees of severity are shown in Figure 2. Each of the fundus photos in the three datasets has a high resolution [31].

The total dataset for this study consists of 39,988 fundus photographs from the three datasets, which are the only ones that receive ratings. According to Table 2, the dataset's class distribution is quite unequal and the majority of the fundus images are non-DR grade.

2.2 Data Processing:

Data processing is divided into two steps. The first involves preprocessing fundus images before to training fundamental models, while the second involves enhancing fundus images while training. The black border around the fundus photos needs to be removed as part of the data processing process since it adds extraneous information, makes it more difficult to extract features from the underlying models, and makes it more challenging to resize the images to the right size for model inputs. The details are as follows:

(1) The border between the dark area and the fundus area was determined using binary processing on the fundus images. The additional black border around each fundus image was then eliminated. The processes are shown in Figure 3.

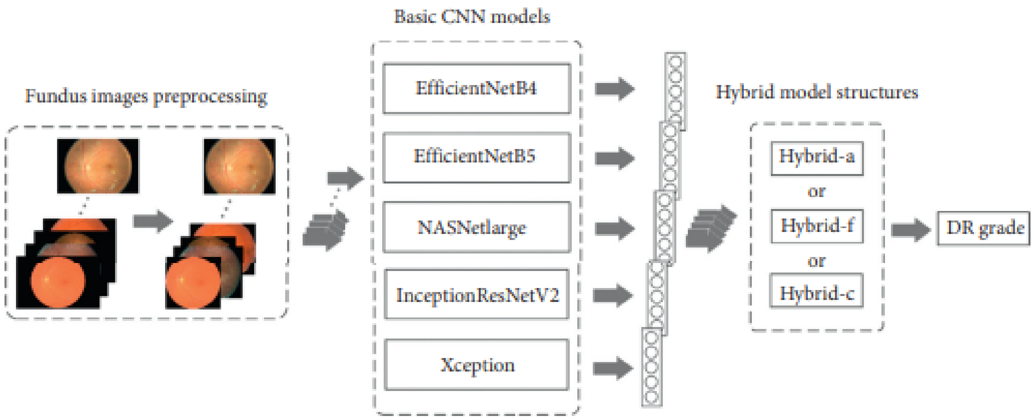

Figure-1: Proposed algorithm architecture

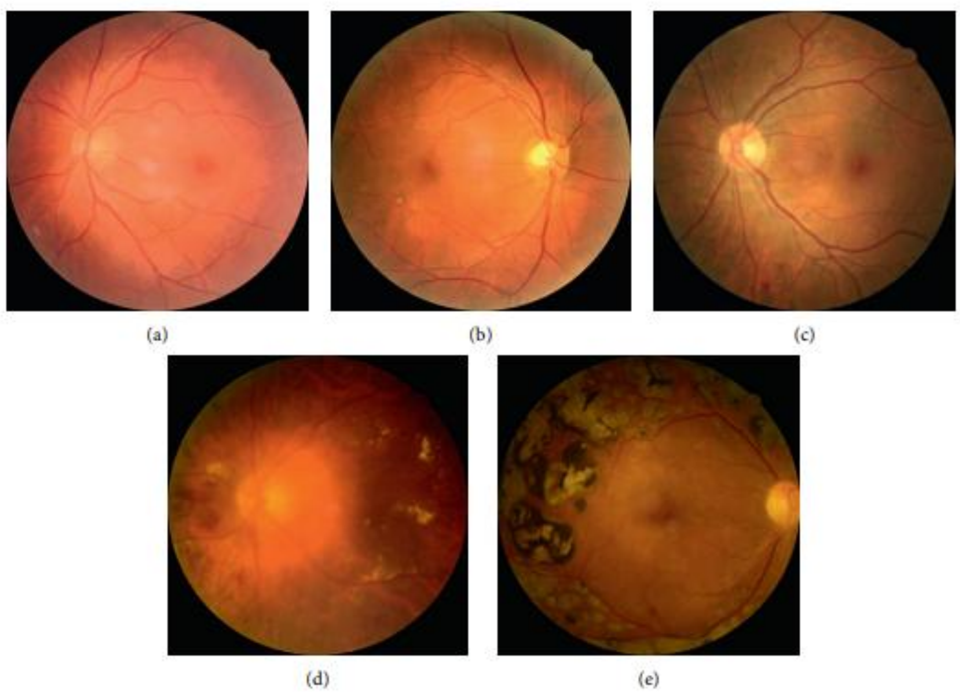

Figure-2: Examples of different severity of DR fundus images. (a) No DR. (b) Mild DR. (c) Moderate DR. (d) Severe DR. (e) Proliferative DR.

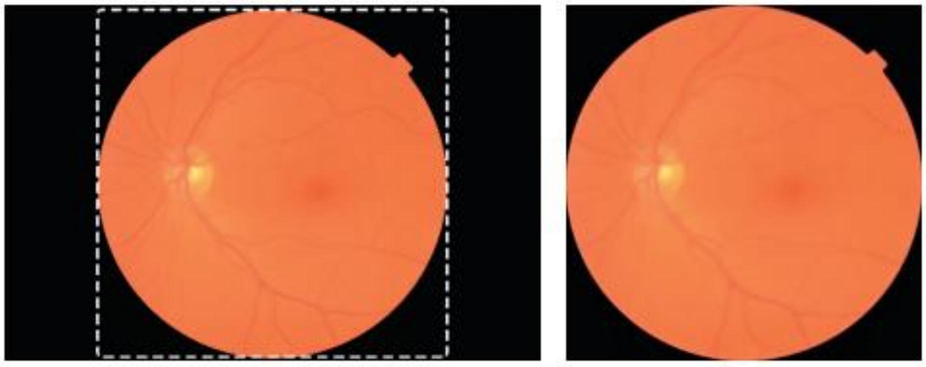

Figure-3: The procedure for eradicating photographs' black borders. The images of the fundus are (a) raw and (b) processed.

(2) Due to the higher resolution of each fundus image, which is inappropriate for the input of the fundamental models, all images were resized to 380 x 380 pixels for EfficientNetB4, 380 x 380 pixels for EfficientNetB4, 331 x 331 pixels for NASNetLarge, and 299 x 299 pixels for EfficientNetB5, Xception, and Inception-ResNetV2.

During the training phase, the fundus images underwent rotation, width shift, height shift, shear range, zoom, horizontal flip, and vertical flip operations. The images were later produced with RandAugment. Cubuk and co. An improved method of data augmentation is RandAugment[21]. .'s Cubuk et al. improved the accuracy of their RandAugment method by 1.0% over baseline augmentation and 0.6% over the previous state-of-the-art on the ImageNet dataset.

2.3 Basic Model Structure:

Hybrid model structures were introduced in this work to improve the fundamental models' potential for categorisation. Xception, InceptionResNetV2, EfficientNetB4, EfficientNetB5, NASNetLarge, and EfficientNetB5 were the primary models. The hybrid model structure was then implemented using three ways. Finally, experiments were carried out to validate the performance of the core models both on their own and in combination with hybrid model structures. Section 3 presents the output. The fundamental models have the following structures:

(1) EfficientNet: Tan et al. created the EfficientNet family of models [22]. They proposed a scaling method [23] that consistently scales the depth, breadth, and resolution of CNNs using a simple yet very effective compound coefficient. In order to construct EfficientNet, which performs better than earlier ConvNets in terms of accuracy and efficiency, the researchers first created a new baseline network using a neural architecture search. The basic models in this study were chosen as EfficientNetB4 and EfficientNetB5. EfficientNetB5's input size was changed to 299 299 pixels whereas EfficientNetB4 kept the previous input resolution. To each of them, a fully linked layer with 5 units and a dropout layer with a 0.4 drop rate were added. The softmax function served as the activation function for the completely connected layer.

(2) NASNetLarge: According to the neural architecture search (NAS) framework, Zoph et alNASNet.'s architecture was the best architecture on CIFAR-10 [25]. Several variants of NASNets with different computing requirements can be constructed by simply

altering the number of convolutional cells and the number of filters inside the convolutional cells. We chose the enormous NASNet-A, which had the best outcomes for categorizing ImageNet picture data, as our primary model. The softmax function was used to replace the original model output with a dropout layer with a 0.4 drop rate and a fully connected layer with 5 units.

(3) Szegedy et al., InceptionResNetV2 The residual connections [28] proposed by He et al., which are based on the inception network design [27], replace the filter concatenation stage in the InceptionResNetV2 model [26]. Training utilizing residual connections greatly speeds up inception network training. Relatively less expensive inception networks with residual links perform better, however by a little margin. In this work, the softmax function was utilized to replace only the final fully linked layer with 1000 units with a fully connected layer with 5 units.

(4) Xception: Inspired by the movie Inception, Chollet [29] developed the Xception architecture, a convolutional neural network architecture that is entirely composed of depthwise separable convolution layers. The 36 depthwise separable convolutional layers that make up the feature extraction basis of the network in the Xception architecture make it relatively easy to define and modify the architecture. The 14 modules made up of the 36 convolutional layers all have linear residual connections surrounding them, with the exception of the first and end modules. In this work, the fully-connected layers and the logistic regression layer of the Xception architecture were replaced with a dropout layer and a fully-connected layer with 5 units.

2.4 Hybrid Model Structure:

To implement the hybrid model structure, three methods—Hybrid-a, Hybrid-f, and Hybrid-c—were proposed. The details are as follows:

Hybrid-a: The hybrid model's ultimate output in Hybrid-a is calculated as the average value of each DR grade that the basic model outputs. The formula is

$$Y_{\text{grade}} = \frac{1}{N}\sum_{n=1}^{N} y_n^{\text{grade}}, \quad (\text{grade} = 0, 1, 2, 3, 4), \qquad (1)$$

where N is the total number of fundamental models. Ygrade n stands for the DR grade of the output from the nth model, and ygrade stands for the DR grade of the output from the hybrid-final a model.

Hybrid-f: Fully connected layers make up the majority of the hybrid-f model. A 25 by 1 column vector is produced by stacking each basic model's output, which is a 5 by 1 column vector, vertically. This vector is used as the input for the Hybrid-f model structure. Figure 4 depicts the f's of the hybrid-structure. Hybrid-f consists of 2 fully connected layers. A softmax activation function is utilised in both the 2048 units of the hidden layer and the 5 units of the output layer, which has 5 units.

Hybrid-c: The hybrid-c algorithm is mostly composed of 2D convolution layers. A 5x5 matrix is produced by horizontally stacking the 5x1 column vector output of each basic model, which is then utilised as the input for the Hybrid-c model structure. Three 2D convolution layers that serve as feature extraction layers make up the Hybrid-c structure. The last part of the Hybrid-c structure is the Hybrid-f structure.

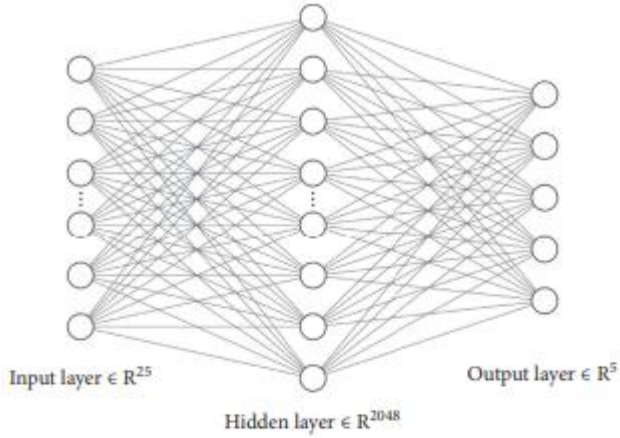

Figure-4: The structure of Hybrid-f

2.5 Loss Function:

Varied loss functions have different effects on the training process and results of network models. In this work, an improved loss function known as the E-CE loss function was proposed for the training of the fundamental models. Comparative analyses utilizing the CE loss function were also carried out. The formula for the CE loss function is

$$L(\hat{y}, y) = -\frac{1}{N}\sum_{n=1}^{N}\left[y_n \log \hat{y}_n + (1-y_n)\log(1-\hat{y}_n)\right], \quad (2$$

y is the true value, y is the predicted value, and y is the true value, where N is the total number of DR grades. Based on the CE loss function, the E-CE loss function is depicted as follows:

$$L(\hat{y}, y) = -\frac{1}{N}\sum_{n=1}^{N}\left[y_n \log \hat{y}_n + (1-y_n)\log(1-\hat{y}_n) + \left|\frac{G_y - G_{\hat{y}}}{N-1}\right|\right] \quad (3$$

Gy stands for the DR grade of truth and Gy for the DR grade of prediction, respectively. The DR grade ranges from 0 to 4. A fraction of the loss is supplied to calculate the impact of misclassifying the core models during training. The greater the deviation between the output value of the model and its true value during the model training phase, the greater the excess loss. Experiments show that the E-CE loss function will speed up the training of the basic models and improve their precision.

CONCLUSION:

In this study, we proposed three hybrid model structures: Hybrid-a, Hybrid-f, and Hybrid-c, to improve the performance of a single model. The E-CE loss function, which we also proposed, is an improved loss function. The findings demonstrate that, in comparison to the CE loss function, the E-CE loss function may efficiently speed up the training of a single basic model and enhance the performance of a single model. The three different hybrid model architectures can improve the performance of the base models in all areas. Although the performance of the hybrid model does not increase linearly with complexity, in some evaluation criteria the more complex Hybrid-c and Hybrid-f surpass the simpler Hybrid-a. The proposed algorithm in this paper generated five precise classifications.

References:

[1] L. Hill and L. E. Makaroff, “Early detection and timely treatment can prevent or delay diabetic retinopathy,” Diabetes Research and Clinical Practice, vol. 120, pp. 241–243, 2016.

[2] Z. Yavuz and C. K¨ose, “Blood vessel extraction in color retinal fundus images with enhancement filtering and unsupervised classification,” Journal of Healthcare

Engineering, vol. 2017, Article ID 4897258, 12 pages, 2017.

[3] X. Feng, G. Cai, X. Gou, Z. Yun, W. Wang, and W. Yang, "Retinal mosaicking with vascular bifurcations detected on vessel mask by a convolutional network," Journal of Healthcare Engineering, vol. 2020, Article ID 7156408, 13 pages, 2020.

[4] H. Narasimha-Iyer, A. Can, B. Roysam et al., "Robust detection and classification of longitudinal changes in color retinal fundus images for monitoring diabetic retinopathy," IEEE Transactions on Biomedical Engineering, vol. 53, no. 6, pp. 1084–1098, 2006.

[5] B. Antal and A. Hajdu, "An ensemble-based system for microaneurysm detection and diabetic retinopathy grading," IEEE Transactions on Biomedical Engineering, vol. 59, no. 6, pp. 1720–1726, 2012.

[6] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "DREAM: diabetic retinopathy analysis using machine learning," IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 5, pp. 1717–1728, 2014.

[7] K. M. Adal, P. G. van Etten, J. P. Martinez, K. W. Rouwen, K. A. Vermeer, and L. J. van Vliet, "An automated system for the detection and classification of retinal changes due to red lesions in longitudinal fundus images," IEEE Transactions on Biomedical Engineering, vol. 65, no. 6, pp. 1382–1390, 2018.

[8] X. Li, X. Hu, L. Yu, L. Zhu, C.-W. Fu, and P.-A. Heng, "CANet: cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading," IEEE Transactions on Medical Imaging, vol. 39, no. 5, pp. 1483– 1493, 2019.

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

[10] A. Krizhevsky, I. Sutskever, and E. H. Geoffrey, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.

[11] E. Decenci`ere, X. Zhing, G. Cazuguel et al., "Feedback on a publicly distributed image database: the messidor database," Image Analysis & Stereology, vol. 33, no. 3, pp. 231–234, 2014.

[12] S. K. Hosseinzadeh, K. H. Peyman, R. Khazaeinezhad, J. W. Michal, A. S. Kevin, and D. Ralph, "Diabetic retinopathy classification using a modified xception architecture," in Proceedings of the IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 1–6, Ajman, UAE, December 2019.

[13] F. Li, D. Yuan, M. Zhang, C. Liang, X. zhou, and H. Zhang, "Multi-scale stepwise training strategy of convolutional neural networks for diabetic retinopathy severity assessment," in Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), pp. 1–5, Budapest, Hungary, July 2019.

[14] M. Hajabdollahi, R. Esfandiarpoor, K. Najarian, N. Karimi, S. Samavi, and S. M. R. Soroushmehr, "Hierarchical pruning for simplification of convolutional neural networks in diabetic retinopathy classification," in Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 970–973, Berlin, Germany, July 2019.

[15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," Computing Research Repository, vol. 14091556 pages, 2014.

[16] A. Jain, A. Jalui, J. Jasani, Y. Lahoti, and R. Karani, "Deep learning for detection and severity classification of diabetic retinopathy," in Proceedings of the 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), pp. 1–6, Chennai, TN, India, March 2019.

[17] C. Szegedy, V. Vincent, L. Sergey, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," 8 Journal of Healthcare Engineering in Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2818–2826, Las Vegas, NV, USA, July 2016.

[18] Kaggle, "Diabetic retinopathy detection," 2015, https://www. kaggle.com/c/diabetic-retinopathy-detection.

[19] Kaggle, APTOS 2019 Blindness Detection, https://www. kaggle.com/c/aptos2019-blindness-detection.

[20] DeepDR diabetic retinopathy image dataset, https://isbi. deepdr.org/index.html.

[21] E. D. Cubuk, B. Zoph, J. Shlens, and V. L. Quoc, "Randaugment: practical automated data augmentation with a reduced search space," in Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 3008–3017, Seattle, WA, USA, June 2020.

[22] M. Tan and V. L. Quoc, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," 2019, https:// arxiv.org/abs/1905.11946.

[23] M. Tan, Bo Chen, R. Pang, V. Vasudevan, and V. L. Quoc, "MnasNet: platform-aware neural architecture search for mobile," in Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) , pp. 2815–2823, Seattle, WA, USA, June 2019.

[24] B. Zoph, V. Vasudevan, J. Shlens, and V. L. Quoc, "Learning Transferable Architectures for Scalable Image Recognition," in Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8697–8710, Salt Lake, Utah, USA, June 2018.

[25] B. Zoph and V.L. Quoc, "Neural Architecture Search with Reinforcement Learning," 2017, https://arxiv.org/abs/1611.01578.

[26] C. Szegedy, S. Ioffe, V. Vincent, and A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in Association for the Advancement of Artificial Intelligence (AAAI), Palo Alto, CA, USA, February 2017.

[27] C. Szegedy, W. Liu, Y. Jia et al., "Going Deeper with convolutions," in Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9, Boston, MA, USA, June 2015.

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, Las Vegas, NV, USA, June 2016.

[29] F. Chollet, "Xception: deep learning with depthwise separable convolutions," in Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) , pp. 1800–1807, Honolulu, HI, USA, July 2017.

[30] L. Liu, H. Jiang, P. He et al., "On the variance of the adaptive learning rate and beyond," 2020, https://arxiv.org/abs/1908.03265.

[31] (Hari Om Sharan, 2022)

[32] (Hari Om Sharan Umesh Kumar Gera, 2021)

[33] (Hari Om Sharan C. R., 2014)

[34] (Dikchha Dwivedi, 2021)

[35] (Kumar, 2020)

[36] (Bhupendra Dhakrey, 2022)

[37] (Dr. Hari Om Sharan, 2022)

[38] (Vinod Kumar Singh, 2022)