International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 9s, 2025
https://www.theaspd.com/ijes.php

Prediction Of Diabetic Retinopathy With Pyramidal Structured Convolution And Rank-Based Lesion Candidate Estimation

Mrs. Sharmila EMN[1] , Dr. Prof. R. Suchitra[2]

[1]Research Scholar, Chirashree Institute of Research and Development & RPA First Grade College, (Approved by University of Mysore)
Email Id-emnsharmila@gmail.com
[2]Professor & Research Guide, Chirashree Institute of Research and Development & RPA First Grade College, (Approved by University of Mysore)
Email Id-suchithra.suriya@gmail.com

Abstract

This paper proposes a diabetic retinopathy (DR) grading approach that uses pyramidal structured convolution (Py-CNN) with a rank-based lesion candidate estimation process for extracting the deep lesion descriptors. Initially, the proposed approach segments the possible lesion candidates after preprocessing using the HSV transform-based adaptive histogram equalization approach. The top Ksegmented lesion candidates are ranked based on the edge and energy components present in the lesion candidates. The pyramidal structured convolution has different sections of the convolutional filter where each section processes the ranked lesion candidates. The lower section of the pyramidal structured convolution has higher length filters that process the top-ranked lesion candidates. The length of the convolutional filter reduces as the rank of the lesion candidate reduces. Two networks namely Network-Aand Network-B are utilized to extract the global fundus image and deep lesion candidature descriptors. The deep lesioncandidaturedescriptorsare implantedinthelesionregionofthe featuremap withthe use of a diffusion layer. The diffused descriptors are used by one of the dense networks, while the global descriptors and the deep lesion candidature descriptors are used by another dense network. Based on the predicted probability of the two dense networks the actual classification result can be computed. Datasets namelyAPTOS 2019 and Messidor-2 are utilized to evaluate thealgorithmperformancewithdifferentevaluationscales.Theproposedapproachresultsin a specificity, Mathew's correlation coefficient, F1-score, precision, and accuracy of 99.36%, 96.77%, 97.41%, 97.35%, and 97.47%when evaluated using the APTOS2019 datasets.

Keywords: *Diabetic retinopathy, Severity grading, Convolutional neural network, Dense network, Lesion candidates.*

1. INTRODUCTION

The primary cause of vision lossin diabetespatientsacrossthe globe is diabetic retinopathy (DR) [1]. This DR has early symptoms such as a dark area in vision, blurred vision, visionspots, and progressive loss of vision [2]. The stages of DR include non-proliferative phases (NPDR)suchasmild,moderate,andsevere.Thestageafterthenon-proliferationleadstothe proliferative phase (PDR). The mild NPDR is the early stage where small bulges occur in the blood vessel.At the moderate NPDR stage, the retina gets damaged with the formation of hard exudates, retinal hemorrhages, and microaneurysms. Severe NPDR is the advanced stage in the NPDR phase where abnormal blood vessel growth occurs with the formation of extensive microaneurysms, andhemorrhages [3].Theadvancedstage DRisPDR wherethereis extensive growth of blood vessels. Ophthalmologists access the DR grade on fundus images based on factors such as width, color, angle, and length. The manual detection of DR severity grades is prone to errors and can utilize more time [4]. Different authors proposed different schemes for grading the DR levels in which the deep learning approach plays a vital role. Even though optical coherence tomography (OCT) [5] imaging provides cross-sectional information, the fundus imaging-based diagnosis is low-cost and low time-consuming.

The authors Wan et al. [6] used convolutional neural networks (CNN) along with hyperparameter tuning, and transfer learning [7] approaches. Transfer learning approaches namely ResNet, GoogleNet, VggNet, and AlexNet are utilized in the classification model. An ensemble scheme was proposed [8] which uses five different deep models such as Dense169, Dense121, exception, inceptionV3, and ResNet50. This deep model improves the performance inDRclassificationbyextractingeffectivedescriptorsfromfundusimages.Capsulenetwork [9] is used along with convolutional filters to extract fundus image features. This approach also uses a soft-max layer, and class capsule layer to detect the severity class of the fundus image. The authors Kar et al. [10] detected the DR levels in four stages. Initially, the optic disc present in the fundus image is removed and the vessels are extracted. The resultant image is pre-processed,andthelesioncandidatesare detected.Finally,thefalsecandidatesareeliminated in the post-processing stage. To eliminate the regions that are poorly illuminated a curveletedge enhancement process isutilized.ToidentifytheDRlevel and suspicious regions, attention maps are generated to derive a zoom-in-net approach [11]. For classification, this approach not only uses the suspicious patches it also considers the complete retinal image. The computation of this zoom-in process is high, though it provides reasonable performance.

The CNN algorithm [12] is used to detect the occurrence of microaneurysms [13], in which the affected and normal regions are categorized using semantic segmentation. Based on the semantic segmentation results, the segmented region is trained to classify the NPDR severity levels.To improvethe ability of feature extraction Calvin et al. [14] modify theVGG-16 model to derive attention-enhanced VGG-16 architecture. Spatial and channel attention are also utilizedtoobtaintherelationaldescriptors.CNNwithsqueezeandexactionframeworkis utilized to identify the dominant descriptors and channels. This approach also uses a DeepRetiNet for separating the hard exudates. Crossover grasshopper scheme [15] was used to improve feature extraction. This approach combines the optimization algorithm, and the deep learning model to collect essential descriptors from the fundus pictures. Further, a generative adversarial network (GAN) is used for feature augmentation which uses a multi-class classifier instead of a binary classifier. Due to the deployment of optimization algorithms in the deep learning model, the complexity is higher.

Roberto et al. [16] used attention to separate the bright and dark structures of the retina.This attention focuses on hard exudates, hemorrhages, and microaneurysms. To handle the data imbalance problem a focal loss function is used while the descriptors are extracted with Xception architecture. This scheme provides a Kappa of0.78, and an accuracy of83.7% which can be further improved.Acapsule network [17] is used along with a vision transformer for predicting the DR severity. Initially, contrast limited adaptive histogram equalization (CLAHE) is used along with power law transform as pre-processing. The usage of capsule networks in the tuned vision transformer yields an accuracy of 87.78% in the Messidor-2 dataset. The principal component analysis [18] is used to extract the descriptors from the pre- processed fundus images. Deep features are also extracted utilizing the pre-trained CNN structures namely SqueezeNet, ResNet152, and ResNet150. Usage of the pre-trained model ResNet152 yields a maximum accuracy of94.4%in the APTOS 2019 dataset.

For segmenting the blood vessels, and optic disc the author Zhu et al. [19] proposed the independent U-Net structures. The inceptionV3 model is used to extract the feature, while the features are classified using the symmetric CNN with singular value decomposition (CNN- SVD) structure. During the backpropagation stage of CNN, a synaptic metaplasticity [20] process is used that also uses inceptionV3 structures for feature extraction. This approach providesanF1-scoreof94.24%withreducedcomplexity.Aradialbasisfunctionclassifier
[21] is used to classify DR and non-DR in which features are extracted from microaneurysms, exudates, and blood vessels which are detected by the Histon-based segmentation process. Bi-channelCNNwasproposedbyLin etal. [22],wherethegreencomponent isutilizedtoestimate the entropy, and unsharp masking for enhancing the fundus image. The bi-channel CNN classifies the fundus image using the green component, and grey level image. To compensatefor the non-uniform brightness in the fundus image, the authors Tamoor et al. [23] used modified gamma correction and a Gaussian match filter. Further, mathematical morphologyandentropythresholdingareusedtodetectthelesioncandidatesfromwhichthecandidate descriptors are extracted by shallow models like AlexNet, and LeNet-5. The red lesions are detected [24] for classifying the DR, and non-DR in fundus images. Two CNN models are utilized to extract the features and to perform DR classification.

The DR grading approaches that are discussed above do not use a rank-based lesion candidate deep descriptor extraction process. The rank-based lesion candidate extractionprocess will extract more deep descriptors from the lesion candidates that contribute more toDR grading. Moreover, the descriptors are extracted from these lesion candidates based on the estimated rank. This deep descriptor which is extracted from the lesion candidates along with the global descriptor is utilized to classify the DR types. The key contribution of the work is highlighted below.

(i) The work proposes a pyramidal structured convolution-based DR gradingapproachthat estimates the lesion candidates from which the candidate rank is computed. The pyramidal structural convolution will extract the descriptor based on the rank of the lesion candidate.

(ii) The work also proposes a feature diffusion process, which implants the feature maps of the lesion candidates on the feature map that represents the global descriptor.

(iii) The work uses two dense networks for classifying the feature where the first network predictstheclass probabilityusingthe globaland the deepdescriptor extractedfrom the lesion candidates, while the second network predicts the class probability using the diffused feature map descriptor.

(iv) Finally, the evaluation of the suggested Py-CNN-based DR grading was done based on the average of the predicted probability result obtained from the two dense networks. Datasets Messidor-2 and APTOS 2019 are utilized in the evaluation of the Py-CNN model-based DR grading in classifying 5 different DR severity grade levels.

The description of the forthcoming sections is summarized as follows. Section 2 explains the proposed Py-CNN-based DR grading approach along with its architecture. Section 3 provides the analysis of the Py-CNN approach and lastly, the conclusion is presented in Section 4.

2. PROPOSEDMETHOD

The suggested Py-CNN-based DR grading approach involves major steps such as Fundus image preprocessing, detection of lesion candidate, Edge, and energy-based DR ranking, and Py-CNN classifier as provided in Fig. 1.

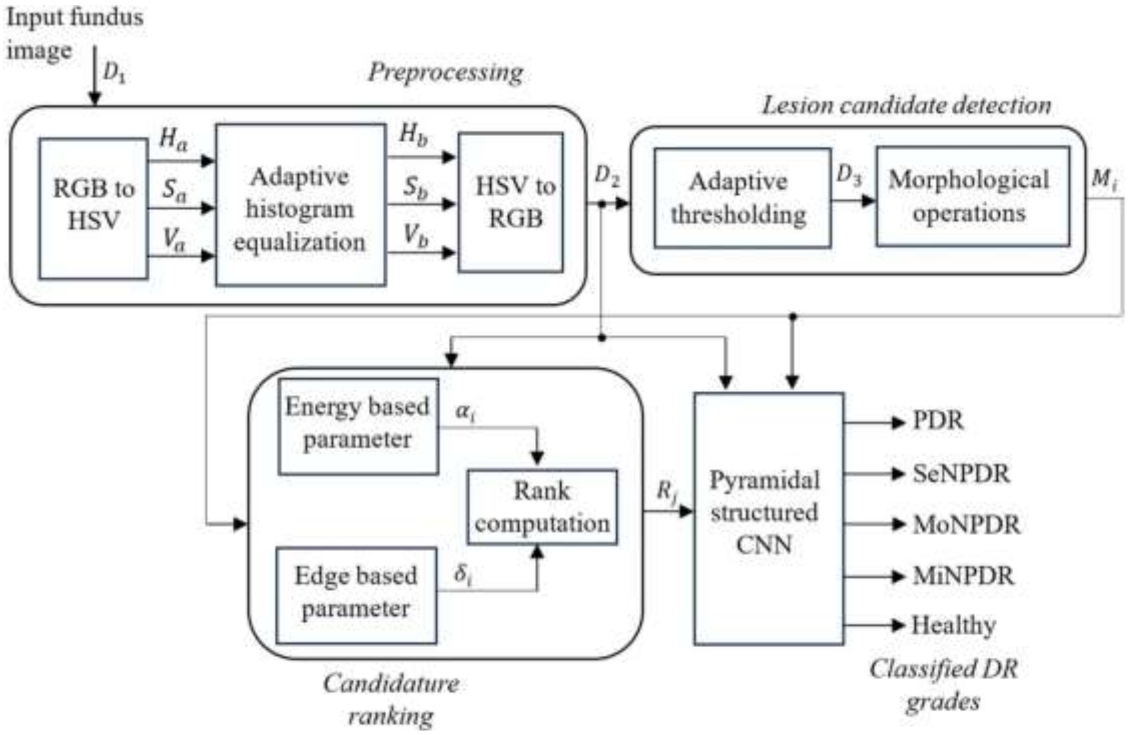


Fig.1:StructureofthesuggestedPy-CNN-basedDRseveritygradingapproach

(a) Fundusimagepreprocessing

ThepreprocessinginthesuggestedPy-CNN-basedDRgradingapproachaimstoenhancethe lesion candidate region suitable for segmentation since the lesion candidate regions are utilized to extract the deep candidate descriptors. LetD_1resembles the input fundus picture. This includesresizingtheimagetoadimensionof256×256,medianfiltering,andenhancingthe image by HSV and CLAHE-based approach [25].The approach initially converts the RGB funduspicturetoHSVcoefficients(H_{a},S_{a},V_{a}),andeachcoefficientisenhancedusingthe CLAHEapproach.Aftertheenhancementofeachchannel,theenhancedHSVcoefficients

(H_{b},b,V_{b}),areagainconvertedbacktoRGBformtoobtaintheenhancedfunduspictureD_2.

(b) Detectionoflesioncandidate

Let D_2resemble the pre-processed image in which the DR lesion candidates are enhanced. The lesion candidates are then detected from this image $D_2$by the processes such as adaptive thresholding and morphological operations. Since the DR lesion candidates have darker intensity than other regions, the adaptive thresholding [26] will segment the lesion candidates whose adaptive threshold is less than (x, y) which varies throughout the fundus picture. The lesion candidates can be detected by the relation

$$D_3(x,y)=\begin{cases}1 & if\quad (\underset{2}{x},y)<\lambda(x,y) \\ 0 & otherwise\end{cases} \qquad (1)$$

The segmented picture $D_3(x,y)$ also contains small noisy structures. Thus, small objects (connected regions) whose area is less than 30 pixels are not considered. Further erosion is performed so that the blood vessels further shrink. For erosion, a disk structuring elementwitha radius of 1 is used.Again, the small objects whose area is less than 30 are further eliminated which removesthe bloodvessels.Fromtheresultantobjects,dilationisperformedtoobtain the boundaries of the lesion region. For dilation, a disk structuring element with a radius of 2 is used. Let the segmented lesion candidate picture contain β number of lesion candidates represented as M_i, where i= 1,2 … . β.

(c) Edgeandenergy-basedDRranking

In the edge and energy-based DR ranking approach, the lesion candidates are ranked by estimating the energy and edge parameters. The lesion region usually has higher energy and edge information. Thus top K lesion candidates that have higher ranks are utilized to extract deeper descriptors. This deep descriptor is extracted by the parallel sections of the Pyramidal structuredconvolutionalfilter.FromβnumberoflesioncandidatesrepresentedasM_i,top$K$ lesioncandidatesareselectedwhichisrepresentedas$\hat{M}_j$where$j$=1,2,….K.Lettheenergy variationofeachcandidatebe$\alpha_i$whichcanbeestimatedbytheexpression,

$$\alpha_i=\rho_i-\mu \qquad (2)$$

Here $\rho_i$and μ resembles the energy of the lesion candidate and background respectively. The energy parameter $\alpha_i$gives the energy variation between the lesion candidate and background respectively.Let$s_1$resemblesthenumberofpixelsinthefundusimage. Theaverageenergy of the background region can be expressed as

$$\mu=\frac{1}{s1}\sum_{(x,y)\in R}\left((x,y)-\ 2\right)^2 \qquad (3)$$

Here R resemblesthecompletefundus imageregion. Similarly,theaverageenergyofthelesion candidate region $M_i$can be estimated as

$$\rho_i=\frac{1}{\gamma_i}\sum_{(x,y)\in M_i}\left((x,y)-\ 2\right)^2 \qquad (4)$$

Here $\gamma_i$resemblesthe numberof pixels enclosedinthelesioncandidate regionM_i.Substituting the average energy of the lesion candidate and the background region in equation (2), theenergy parameter can be expressed as,

$\alpha=^{1}\sum$ $((x,y))^{2}-^{1}\sum$ $-$ $(D(x,y)^{2}$ (5)

(x,y)∈M_i s1 (x,y)∈R 2)

$\overline{i}$ γ_i 2

The edge parameter $\delta_i$can be computed for each lesion candidate region $M_i$by subtracting the smoothened components of the lesion candidates. Thus, the edge information can be estimated as

(x,y) =$D(x,y)-^{1}\sum$ $D(x,y)$ (6)

i 2 $\overline{\gamma_i}$ (x,y)∈M_i2

Here$(x,y)\in M_i$.Fromtheedgeinformation,theedgeparametercanbecomputedas

$\delta=^{1}\sum$ $((x,y))^{2}$ (7)

i $\overline{\gamma_i}$ (x,y)∈M_i i

The edge parameter gives the average energy of the edges. The energy parameter $\alpha_i$and the edge parameter $\delta_i$are used to compute the ranking parameter r_i, which can be estimated from the sum of the edge parameter and the energy parameter

$r_i=\delta_i+\alpha_i$ (8)

The ranking parameter is sorted in descending order and the lesion candidates that have top K ranking parametervalues areselected for extraction ofdeep features. Letthe indexof the lesion candidates that havetop K ranksbe R_j.The selectedtop-K-rankedlesion candidates are utilized to extract the deep features by the pyramidal-structured convolutional filters.

(d) Py-CNNclassifier

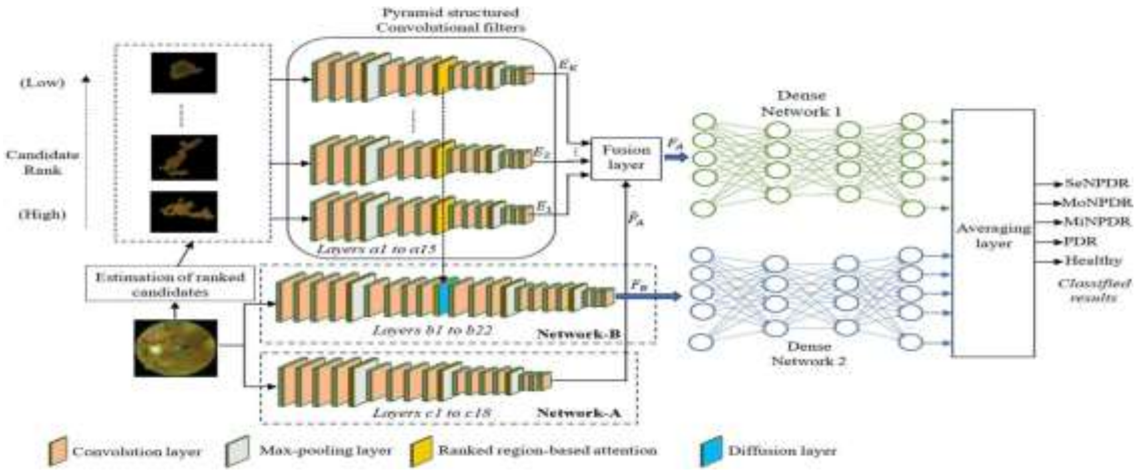
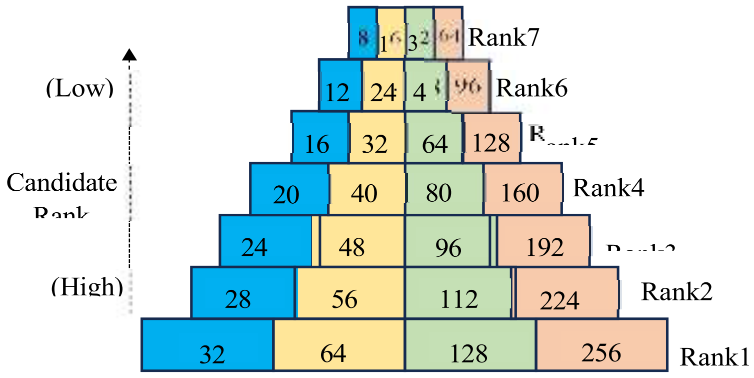


Fig.2:ArchitectureofsuggestedPy-CNNclassifierusedinproposedDRgradingapproach

The structure of the Py-CNN classifier in detecting the DR grade levels is provided in Fig. 2. Three different features are extracted by the Py-CNN classifier namely the global feature, deep lesion candidate features, and the diffused feature. These three features are utilized by twodense networks to classify the DR grade levels. Two subnetworks namely Network-A and Network-B are utilized in the suggested Py-CNN. The pyramidal structured CNN has different sections of convolutional filters, where the lesion candidates corresponding to different ranksare processed by each section. The bottom section corresponds to top-ranked lesion candidates, while the top section corresponds to the low-ranked lesion candidates.Thus, the bottom section of the pyramidal structured filter will have convolutional filters that have higher lengths (more channels), whilethe top sectionof the pyramidalstructured filter will have convolutional filters with a lower length (few channels). The filter length used by the pyramidal structured filterwith7 sectionsareillustrated in Fig. 3.The pyramidalstructured filterwith 7 sectionsresemble that, top-7 lesion candidates are utilized to extract the deep lesion descriptors. In this pyramidal filter, the filter length (channel) reduces as rank reduces.



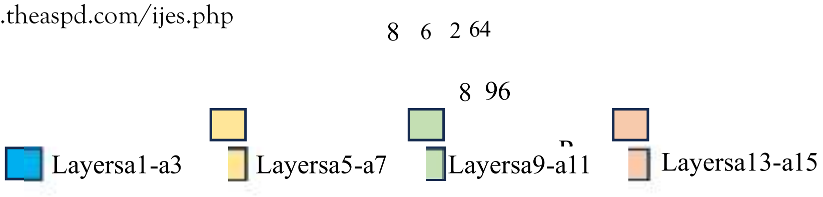


Fig.3:RepresentationofpyramidalstructureconvolutionallengthusedinPy-CNN

The input to the Py-CNN includes the mask of lesion candidates M_{j}that have top K ranks R_{j}along with the fundus image that was preprocessed i.e.$D_2(x,y)$. The lesion candidates are segmented using the mask and resized with a fixed dimension of 64 × 64. The description of variouslayersusedinthePy-CNNisprovidedinTable.I.Thelayers$a1toa3$,$5toa7$, $a9toa11$,and$a13toa15$inPy-CNNarchitecturehaveasizeof64×64,32×32,16×16, and8×8respectively. The number of channels used by the filters in the pyramidal convolutional section is provided in Fig. 3.

TableI:Specificationoflayers(Network-AandNetwork-B)usedinPy-CNN

Layers	Description	Outputsize
$b1tob4$	5×5,32,stride-1	256×256
$b6tob13$	5×5,64,stride-1	128×128
$b15tob18$	3×3,128,stride-1	64 × 64
$b20tob22$	3×3,256,stride-1	32 × 32
$c1toc4$	3×3,32,stride-1	256×256
$c6toc9$	3×3,64,stride-1	128×128
$c11toc14$	3×3,128,stride-1	64 × 64
$c16toc18$	3×3,256,stride-1	32 × 32

The max-pooling layer uses a filter size of 2 × 2, and stride-1.Apart from convolutional layers and max-pooling layers, the Py-CNN also has ranked region-based attention and diffusion layers. The ranked region-based attention will eliminate the feature outside the region of interestrepresentedbythescaledmask(32×32)whichhastwooutputs.Oneoftheoutputs is used by the succeeding layers of the same section, while the other output is used by the diffusionlayerafterresizingtothesizeofthescaledmaskinNetworkB.Therankedregion- based attention performs functions such as eliminating the features outside the region of interest,multiplicationbytherankingweight,andmax-poolingasillustratedinFig.4.

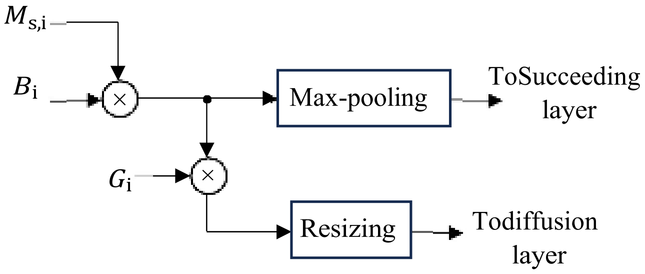


Fig.4:Representationofrank-basedattentionlayerusedinPy-CNN

Let the scaled mask image be represented as $M_{s,i}$and the feature map obtained from the previous layer be B_i. Let $G_i$be the weight that is used by the rank-based attention. The weights oftherank-1torank-7lesioncandidatesareassignedasG_i=[1,0.9,0.8,0.7,0.6,0.5,0.4]. Thesucceedinglayersofthepyramidalsectionusethemax-pooledoutputof$M_{s,i}B_i$,whilethe

network-B or diffusion layer uses the resized lesion candidate region of $M_{s,i}B_iG_i$. The diffusion layer implants the attention layer output in its corresponding lesion candidate region using its scaled mask$M_{s,i}$.Beforeimplanting thefeature,the attentionoutput$M_{s,i}B_iG_i$isresized toa size suitable for diffusion. During the diffusion, the feature maps of the diffusion layer has a larger filter length than the rank-based attention layer. For example, the layer before the diffusion layer in network B has a length of 64, but the output from the pyramidal filter section (rank-2) has a length of 56. Thus, the first 56 feature map of the diffusion layer correspondingto the region of the lesion candidate is replaced by the feature map from the first 56 attention layers, while the remaining feature maps (56to 64) of network-B are unchanged.

Theoutputofnetwork-Aandnetwork-Bisrepresentedas$\hat{F}_A$andF_B.Thefeature$\hat{F}_A$is termed to be a global descriptor obtained from a complete fundus image and feature $F_B$is termed to be a diffused descriptor. The features obtained from each section of the lesion candidates(Top-K)berepresentedasE_1,E_2.....E_K.Theglobaldescriptor$\hat{F}_A$andthelesion candidate descriptor E_1,E_2..... $E_K$can be combined to obtain the feature to be trained by dense layer-1. Thus, the input feature to dense layer-1 can be estimated as

$$F_A=[\hat{F}_A,E_1,E_2.....E_K] \tag{9}$$

The dense network 1 and dense network 2 use the flattened version of the feature $F_A$and $F_B$respectively. The dense network 1 and convolutional filters of network A use the following (equation (10)) cross-entropy loss function to update the kernels and weights in the dense network 1.

$$Q_C=-\sum_{z=1}^{z} a_z\log\hat{a}_z \tag{10}$$

Here, $\hat{a}_z$resembles the predicted probability for class z from dense network 1, while $a_z$is the actual probability and z is the number of classes. The dense network 2, pyramidal filter section, and the convolutional filters in network-B use the loss function expressed as

$$Q_m = \frac{1}{L_{tr}} \sum_{l=1}^{L_{tr}} (b_l - \hat{b}_l)^2 \qquad (11)$$

Here,$_{tr}$resemblesthenumberofimagesusedintraining,$b_l$resemblestheactualprobability

of l^{th} training image, $\hat{b}_l$resembles the predicted probabilityof the l^{th} image.Theconvolutional filtersandthedensenetworksupdatetheweightsusingtherespectivelossfunction$Q \in \{Q_{c,m}\}$withthefollowingweightupdateequation

$$h(t+1) = h(t) - 5\frac{dQ}{dh} \qquad (12)$$

Here $h(t)$ and $h(t+1)$resembles the past weight and the updated weight respectively at iteration t with learning rate 5. The proposed approach uses ReLu activation after each convolutional filter.Fromthepredictedprobability of thetwodense networks(dense network 1 and dense network 2), the actual output can be computed from the average of the two dense network probabilities.

3. EXPERIMENTALRESULTS

The fundus image datasets namely Messidor-2 [27] and APTOS 2019 [28] are utilized to evaluate the Py-CNN-based approach. The evaluation measures namely, F1-Score, Recall, Specificity, Precision, Accuracy, and Mathew's correlation coefficient (MCC) are utilized to evaluate the Py-CNN-based approach. The evaluation parameter can be evaluated using the relations provided below,

$$F1-Score = \frac{\Delta_{tp}}{\frac{1}{2}(\Delta_{fn}+\Delta_{fp})+\Delta_{tp}} \times 100\% \qquad (13)$$

$$Recall = \frac{\Delta_{tp}}{\Delta_{fn}+\Delta_{tp}} \times 100\% \qquad (14)$$

$$Specificity = \frac{\Delta_{tn}}{\Delta_{tn}+\Delta_{fp}} \times 100\% \qquad (15)$$

$$MCC = \frac{\Delta_{tp}\times\Delta_{tn}-\Delta_{fp}\times\Delta_{fn}}{\sqrt{(\Delta_{tp}+\Delta_{fp})\times(\Delta_{tp}+\Delta_{fn})\times(\Delta_{tn}+\Delta_{fp})\times(\Delta_{tn}+\Delta_{fn})}} \times 100\% \qquad (16)$$

$$Precision = \frac{\Delta_{tp}}{\Delta_{tp}+\Delta_{fp}} \times 100\% \qquad (17)$$

$$Accuracy=\frac{\Delta_{tp}+\Delta_{tn}}{\Delta_{tp}+\Delta_{tn}+\Delta_{fp}+\Delta_{fn}}\times 100\% \qquad (18)$$

Here,Δ_{tn},Δ_{fp},Δ_{fn},andΔ_{tp}bethetruenegative,falsepositive,falsenegative,andtruepositive result estimated in the testing phase of the proposed Py-CNN-based approach. DR grades namely severe NPDR (SeNPDR), mild NPDR (MiNPDR), moderate NPDR (MoNPDR), PDR, and Healthy classes are utilizedfor evaluating theseverity level.The number of imagesutilized fromtheMessidor-2andAPTOS-2019datasetsarelistedinTableII.Outof999 images,only 500imagesarerandomlyselectedfortheclassMoNPDRfromtheAPTOS2019dataset.

Table II: Number of images utilized for analyzing the suggested Py-CNN-based DR classificationapproach(Here*resemblestheclassinwhichaugmentationisnotperformed)

Augmentation	Dataset	**SeNPDR**	**MoNPDR**	**MiNPDR**	**PDR**	**Healthy***
Before	APTOS2019	193	500	370	295	1805
	Messidor-2	75	347	270	35	1017
After	APTOS2019	1158	3000	2220	1770	1805
	Messidor-2	450	2082	1620	210	1017

Table II also shows the number of images after augmentation. Five different augmentations namely 3 rotations (90°, 180°, 270°), brightening, and darkening by 50 are used. Using the augmented images, 30% of images are randomly chosen for classification, while the remaining 70% of images are utilized for training.The distribution of training and testing images in each class is provided in Fig. 5.

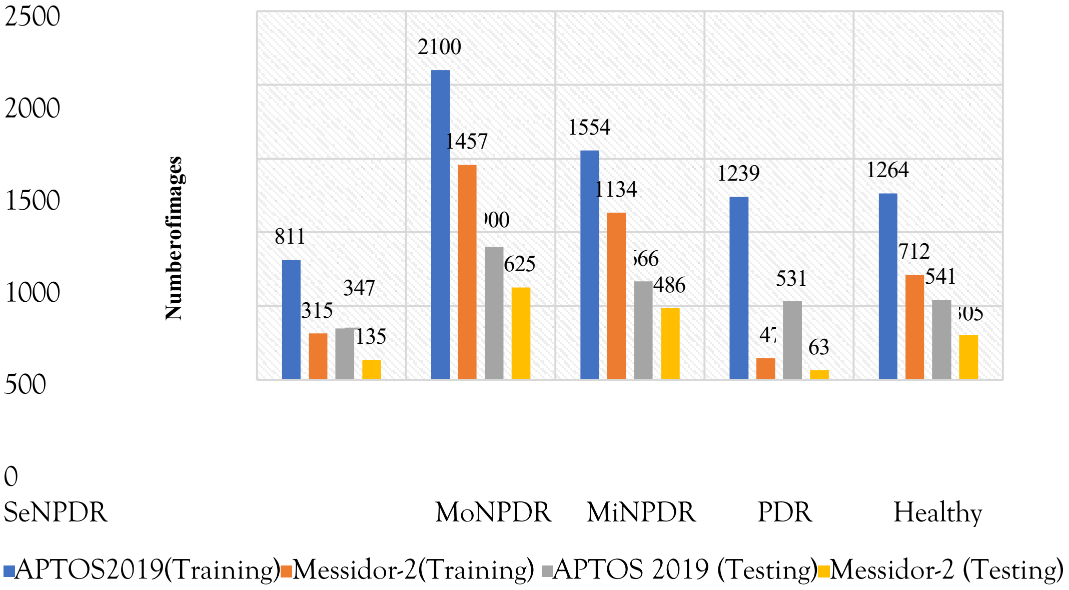


Fig. 5: Distribution of training and testing data used for the analysis of the suggested Py- CNN-based

approach

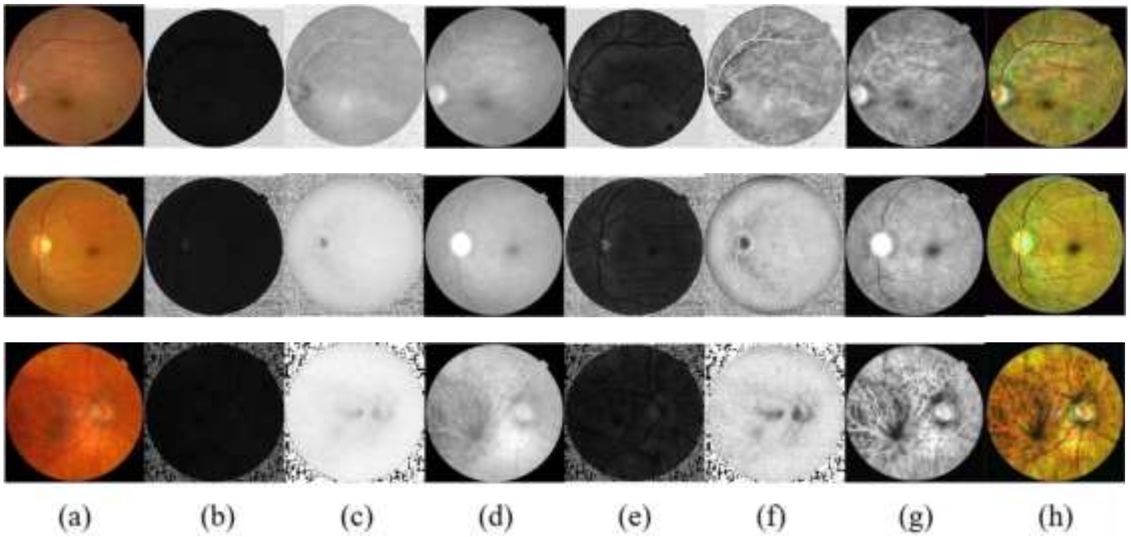
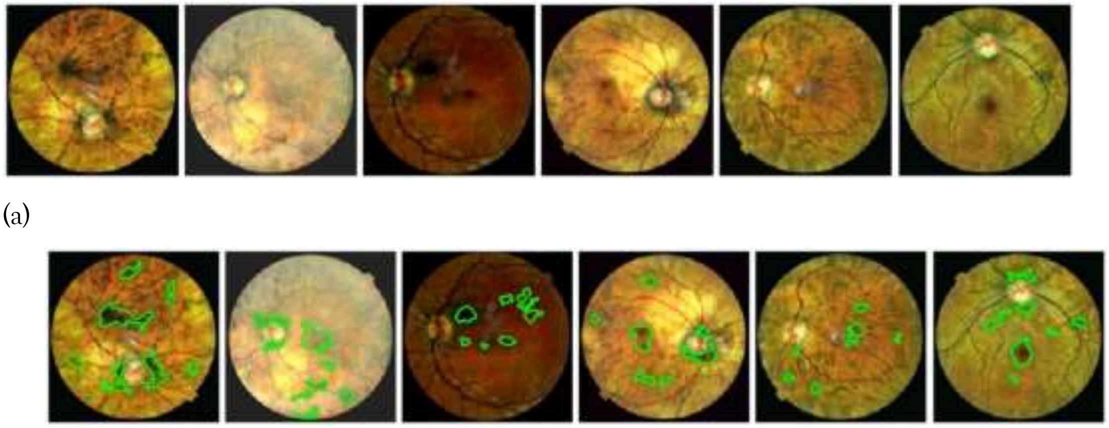


Fig. 6: Preprocessing result obtained in the suggested scheme (a) Input fundus picture (b)-components (c) $\boldsymbol{S}$-components (d) $\boldsymbol{V}$-components (e) CLAHE enhanced $\boldsymbol{H}$-component (f) CLAHE enhanced$\boldsymbol{S}$-component (g) CLAHE enhanced$\boldsymbol{V}$-component (h) Pre-processed fundus picture

The preprocessingresultobtainedby thesuggestedPy-CNN-basedapproachisdepictedinFig.

6.The CLAHE algorithm enhances all three components($\boldsymbol{H}$,$\boldsymbol{S}$,$\boldsymbol{V}$)to obtain the enhanced image. The optical disc, blood vessels, and microaneurysms are clearly visible in the pre-processedpicture.TheCLAHEschemeutilizesa grid sizeof 8 × 8 andacliplimitof40.

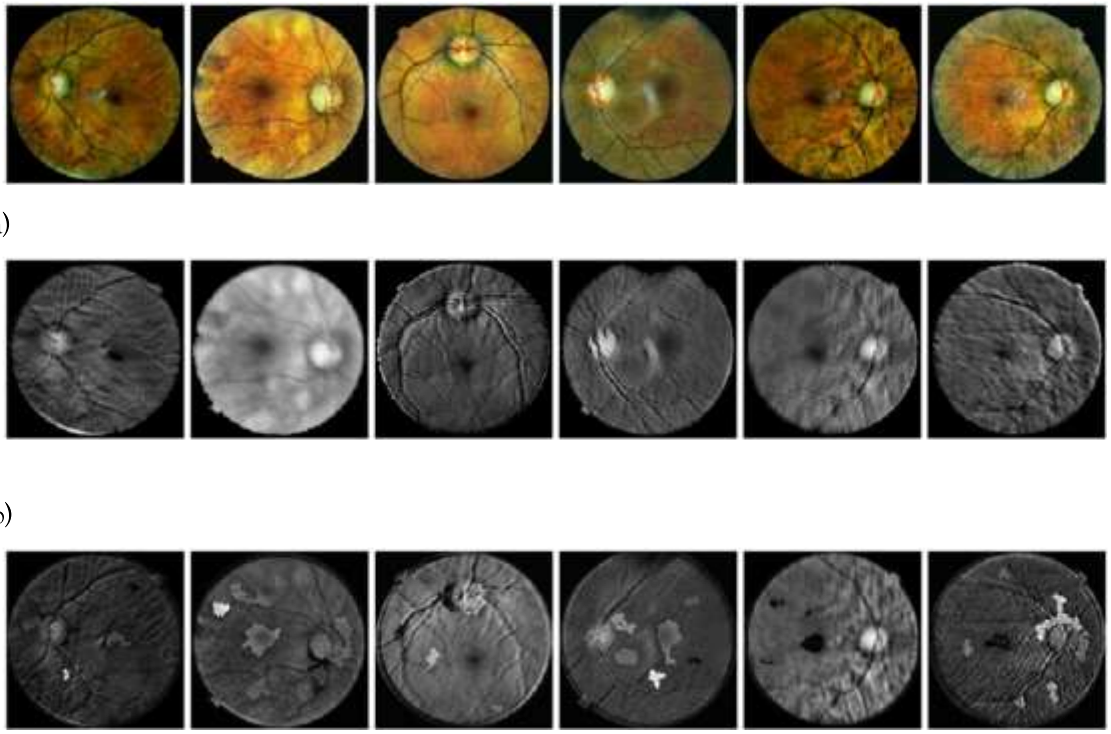


(a)

(b)

Fig. 7: Representation of top-7 ranked lesion candidates estimated in the Py-CNN-based approach (a) Input image (b) Top-7 lesion candidates

The top-7 ranked lesion candidates selected by the energy and edge parameter-based rankingare illustrated in Fig.7. These top-7 ranked lesion candidates are utilized to extract deep lesion candidate descriptors if the Kvalue is chosen as 7.



(a)

(b)

(c)

Fig. 8: Feature maps obtained by the trained convolutional filters used in Network-A and Network-B (a) Global descriptor extraction process by Network-A(b) Diffused descriptor extraction process by Network-B

Sample feature map obtained by Network-Aand Network-B is illustrated in Fig. 8(b) and Fig. 8(c) respectively. Fig. 8(b) illustrates the global feature map, while Fig. 8(c) illustrates the diffused feature maps. When comparing the global feature map and the diffused feature map,thediffusedfeaturemapcontainsmoreinformationonthe selectedlesioncandidate region. FortrainingthesuggestedPy-CNNapproach,theAdamoptimizerwithalearningrateof 0.01isused. Thetrainingwasperformedwithabatchsizeof16for60epochs.

Table III: Class-wise performance obtained by the Py-CNN-based approach for the Messidor-2 dataset

Classes	F1-score (%)	Recall (%)	Specificity (%)	MCC (%)	Precision (%)	Accuracy (%)
SeNPDR	94.51	95.56	99.39	94.00	93.48	95.56
MoNPDR	97.01	96.01	98.79	95.16	98.04	96.02
MiNPDR	96.81	96.71	98.67	95.43	96.91	96.71
PDR	82.09	87.30	98.97	81.47	77.46	87.30
Healthy	97.55	98.03	99.31	96.98	97.08	98.03

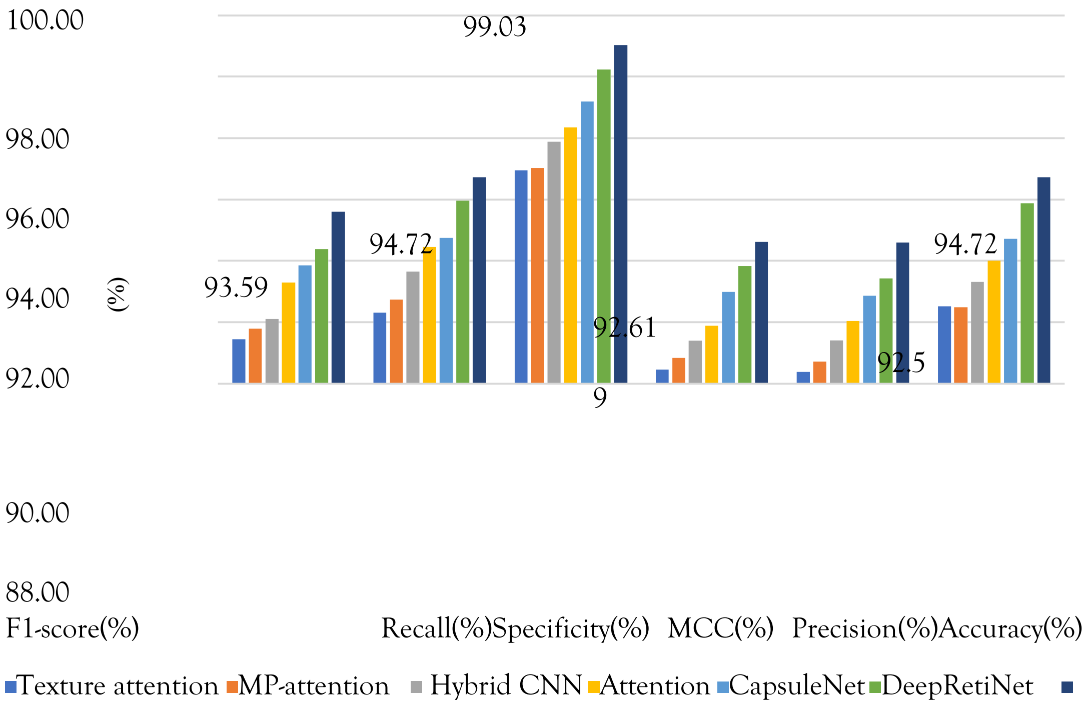
Table III provides the class-wise comparison of performance using the suggested Py-CNN approach when analyzed utilizing the Messidor-2 dataset. For the classes SeNPDR, MoNPDR, MiNPDR, PDR, and Healthy cases, the Py-CNN provides an accuracy of 95.56%, 96.02%, 96.71%, 87.30%, and

98.03% respectively when evaluated using the Messidor-2 dataset. The classification accuracy is maximum for the class healthy, while it is lower for the PDR cases. The performance of the proposed Py-CNN-based DR severity grading approach was compared with other DR severity grading approaches namely texture attention [29], MP-attention [30], Hybrid CNN [31], attention [16], CapsuleNet [17], and DeepRetiNet [14] approaches as provided in Table IV.

Table IV: Performance comparison between the suggested Py-CNN-based DR grading approach and other approaches when evaluated using the Messidor-2 dataset

Schemes	1-score (%)	Recall (%)	ecificity (%)	MCC (%)	ecision (%)	ccuracy (%)
Textureattention[29]	89.45	90.31	94.96	88.46	88.38	90.52
MP-attention[30]	89.79	90.74	95.02	88.84	88.72	90.49
HybridCNN[31]	90.10	91.65	95.88	89.39	89.41	91.31
Attention[16]	91.30	92.45	96.35	89.88	90.04	91.99
CapsuleNet[17]	91.86	92.75	97.19	90.98	90.87	92.72
DeepRetiNet[14]	92.39	93.96	98.24	91.83	91.43	93.88
Proposed	93.59	94.72	99.03	92.61	92.59	94.72

For the suggested Py-CNN approach the F1-Score, recall, specificity, MCC, precision, and accuracy were estimated as 93.59%, 94.72%, 99.03%, 92.61%, 92.59%, and 94.72% respectively. When comparing the suggested Py-CNN with the DeepRetiNet approach, the F1- score, recall, specificity, MCC, precision, and accuracy improve by 1.2%, 0.77%, 0.79%, 0.78%, 1.17%, and 0.84% respectively when evaluated using the Messidor-2 dataset. The graphical comparison is provided in Fig. 9.



Proposed

Fig. 9: Graphical comparison of performance between suggested Py-CNN and similar DR grading schemes when evaluated using Messidor-2 data

The class-wise performance comparison between different DR classes when evaluated usingtheAPTOS 2019 dataset is illustrated in Table V. In the case of theAPTOS 2019 dataset, the proposed Py-CNN yields an accuracy of 97.98%, 97.78%, 97.60%, 96.05%, and 97.97% for the DR grades SeNPDR, MoNPDR, MiNPDR, PDR, and healthy respectively.

Table V: Class-wise performance obtained by the Py-CNN-based approach for theAPTOS 2019 dataset

Classes	F1-score (%)	Recall (%)	pecificity (%)	MCC (%)	recision (%)	ccuracy (%)
SeNPDR	96.87	97.98	99.43	96.46	95.77	97.98
MoNPDR	97.67	97.78	98.94	96.66	97.56	97.78
MiNPDR	97.67	97.60	99.35	97.00	97.74	97.60
PDR	97.14	96.05	99.63	96.54	98.27	96.05
Healthy	97.70	97.97	99.43	97.18	97.43	97.97

The average F1-score, recall, specificity, MCC, precision, and accuracy estimated by the suggested Py-CNN is 97.41%, 97.47%, 99.36%, 96.77%, 97.35%, and 97.47% respectively when evaluated usingtheAPTOS 2019 dataset.The average F1-score, recall, specificity, MCC, precision, and accuracy estimated by the suggested Py-CNN approach are 0.72%, 0.58%, 0.57%, 0.79%, 0.45%, and 0.52%respectively when evaluated using the APTOS 2019dataset.

Schemes	1-score (%)	ecall (%)	pecificity (%)	MCC (%)	recision (%)	ccuracy (%)
Textureattention[29]	93.67	93.63	95.67	92.66	93.49	93.48
MP-attention[30]	93.88	94.06	95.85	93.15	93.94	94.09
HybridCNN[31]	94.82	94.60	96.69	93.99	94.31	94.51
Attention[16]	95.23	95.54	97.38	94.81	95.17	95.42
CapsuleNet[17]	95.96	96.22	98.00	95.69	95.89	96.06
DeepRetiNet[14]	96.68	96.90	98.79	95.98	96.90	96.95
Proposed	97.41	97.47	99.36	96.77	97.35	97.47

Table VI: Performance comparison between the suggested Py-CNN-based DR grading approach and other approaches when evaluated using the APTOS 2019 dataset

The graphical comparison of performance between the suggested Py-CNN approach and other similar schemes when evaluated using the APTOS 2019 dataset is presented in Fig. 10. The comparison shows that the performance of the suggested scheme is higher than otherapproaches used for comparison.

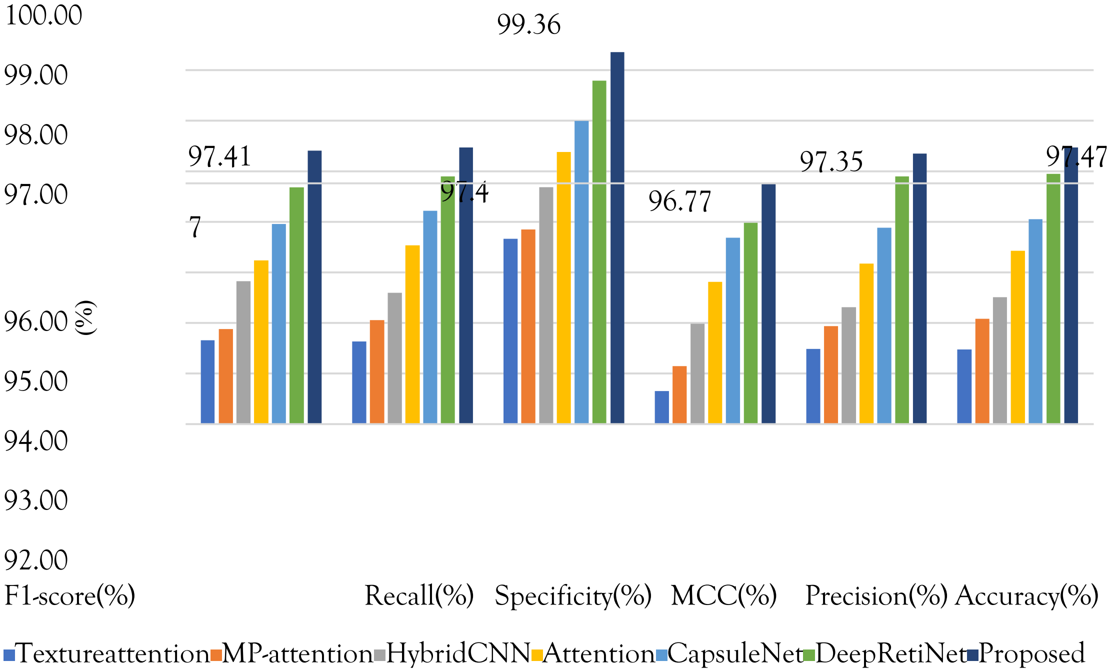
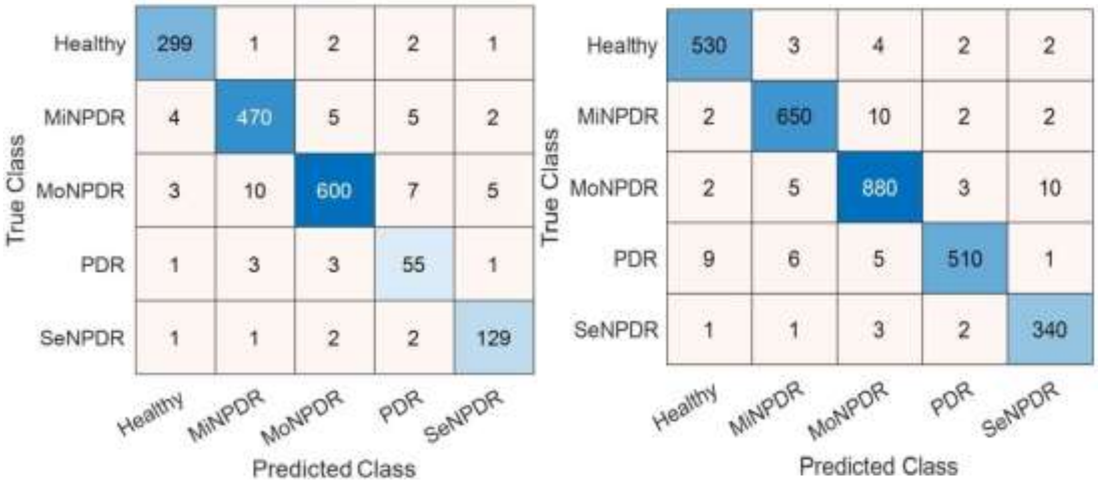


Fig. 10: Graphical comparison of performance between suggested Py-CNN and similar DR grading schemes when evaluated using APTOS 2019 data



	Healthy	MiNPDR	MoNPDR	PDR	SeNPDR
Healthy	299	1	2	2	1
MiNPDR	4	470	5	5	2
MoNPDR	3	10	600	7	5
PDR	1	3	3	55	1
SeNPDR	1	1	2	2	129

(a)

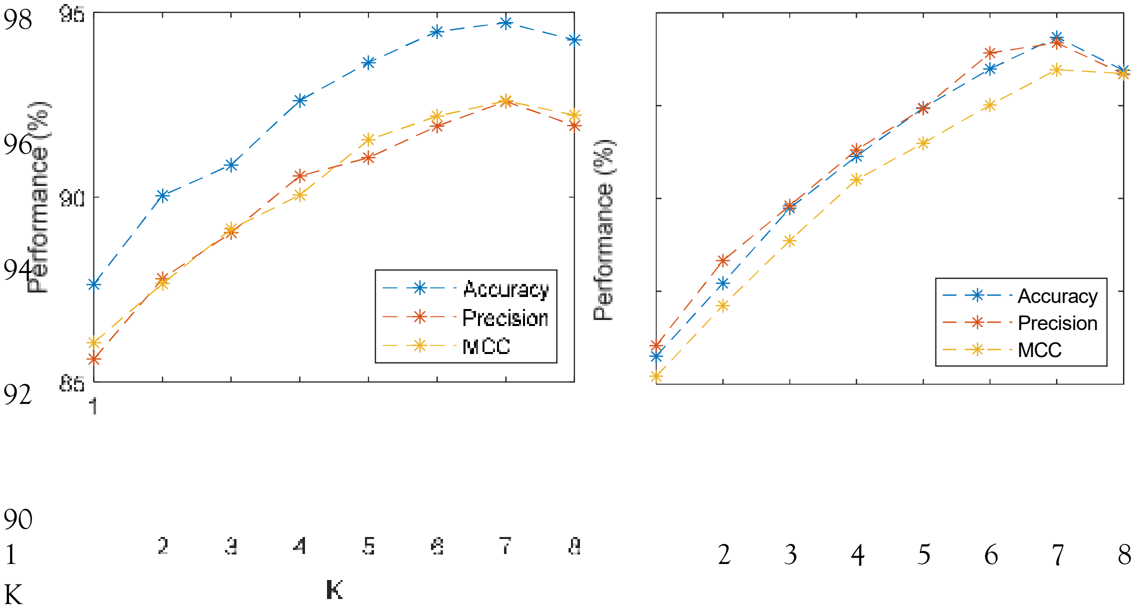
	Healthy	MiNPDR	MoNPDR	PDR	SeNPDR
Healthy	530	3	4	2	2
MiNPDR	2	650	10	2	2
MoNPDR	2	5	880	3	10
PDR	9	6	5	510	1
SeNPDR	1	1	3	2	340

(b)

Fig.11:ConfusionmatrixestimatedduringthetestingphaseofPy-CNNmodel(a)Mesidor-2

(b)APTOS2019dataset

The confusion matrix obtained by the suggested Py-CNN during the testing process utilizingthe Messidor-2 andAPTOS 2019 is illustrated in Fig. 11(a) and Fig. 11(b) respectively. When comparing the two datasets, the APTOS 2019 dataset provides a higher proportion of true positives than the Messidor-2 dataset. This shows that the suggested approach provides higher performance for the APTOS 2019 dataset than the Messidor-2 dataset.



(a) (b)

Fig. 12: Variation of performance with respect to the parameter (a) Messidor-2 (b)APTOS 2019 dataset

The variation of measures such as accuracy, precision, and MCC with respect to$K$is provided in Fig. 12.Asthe K valueincreases from 1, the performanceincreases.The performanceattains maximum with $K = 7$. This resembles that the usage of more lesion candidates extracts more deep features. For further increases in K value, the performance is slightly reduced. This reduction is due to the increase in redundant features or due to the involvement of normal regions in the extraction of deep features.

4. CONCLUSION

The work proposed a Pyramidal structured convolution-based (Py-CNN) DR grading approach that uses a rank-based lesion candidate estimation. The scheme initially preprocesses the fundus image to enhance the DR lesion candidate's region. The DR lesion candidates are then segmented and a rank is computed for each candidate. The Py-CNN structure has the pyramidal structure of convolutional filters, where features are extracted from the top-ranked candidates using a convolutional filter having a larger length. As the rank reduces, the filter length also reduces. Features such as diffused features, lesion candidate features, and global features are extracted using the Py-CNN. Two dense networks are used in the Py-CNN classifier, where the first dense network uses the combined features namely the global feature and thecandidate feature, while the second densenetwork uses the diffused features.The actual predicted probability is estimated from the average probability of the two dense networks to classify the DR grades such as PDR, healthy, and NPDR types (Severe, mild, and moderate). Publicly available DR grading datasets such as Messidor-2 and APTOS 2019 are utilized for evaluating the Py-CNN classifier performance. The proposed Py-CNN classifier provides an accuracy of 94.72% and 97.47% when evaluated using the Messidor-2 and APTOS 2019 datasets.

REFERENCES

[1]Fong,D.S.,Aiello,L.P.,FerrisIII,F.L.,&Klein,R.(2004).Diabeticretinopathy.Diabetes care, 27(10).

[2].Stitt,A.W.,Curtis,T.M.,Chen,M.,Medina,R.J.,McKay,G.J.,Jenkins,A.,...&Lois, N.(2016).Theprogressinunderstandingandtreatmentofdiabeticretinopathy. Progressin retinal and eye research, 51, 156-186.

[3]. Vujosevic, S.,Aldington, S. J., Silva, P., Hernández, C., Scanlon, P., Peto, T., & Simó, R. (2020). Screening for diabetic retinopathy: new perspectives and challenges. The Lancet Diabetes & Endocrinology, 8(4), 337-347.

[4]. Kollias,A. N., & Ulbig, M. W. (2010). Diabetic retinopathy: early diagnosis and effective treatment. Deutsches Arzteblatt International, 107(5), 75.

[5]. Suchetha, M., Ganesh, N. S., Raman, R., & Dhas, D. E. (2021). Region of interest-based predictive algorithm for subretinal hemorrhage detection using faster R-CNN. Soft Computing, 25(24), 15255-15268.

[6]. Wan, S., Liang, Y., & Zhang, Y. (2018). Deep convolutional neural networks for diabetic retinopathy detection by image classification. Computers & Electrical Engineering, 72, 274- 282.

[7].Krishnan,S.H.,Vishwa,C.,Suchetha,M.,Raman,A.,Raman,R.,Sehastrajit,S.,&Dhas, D. E. (2023). Comparative performance of deep learning architectures in classification of diabeticretinopathy.InternationalJournalofAdHocandUbiquitousComputing,44(1),23-35.

[8]. Qummar, S., Khan, F. G., Shah, S., Khan, A., Shamshirband, S., Rehman, Z. U., ... & Jadoon, W. (2019).Adeep learning ensemble approach for diabetic retinopathy detection. Ieee Access, 7, 150530-150539.

[9]. Kalyani, G., Janakiramaiah, B., Karuna, A., & Prasad, L. N. (2023). Diabetic retinopathy detection and classification using capsule networks. Complex & Intelligent Systems, 9(3),2651-2664.

[10]. Kar, S. S., & Maity, S. P. (2017).Automatic detection of retinal lesions for screening of diabetic retinopathy. IEEE Transactions on Biomedical Engineering, 65(3), 608-618.

[11].Wang,Z.,Yin,Y.,Shi,J.,Fang,W.,Li,H.,&Wang,X.(2017).Zoom-in-net:Deepmining lesions for diabetic retinopathy detection. In Medical Image Computing and ComputerAssisted Intervention— MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part III 20 (pp. 267-275). Springer International Publishing.

[12]. Das, S., Kharbanda, K., Suchetha, M., Raman, R., & Dhas, E. (2021). Deep learning architecture based on segmented fundus image features for classification of diabeticretinopathy. Biomedical Signal Processing and Control, 68, 102600.

[13]. Qiao, L., Zhu, Y., & Zhou, H. (2020). Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms. IEEE Access, 8, 104292-104302.

[14]. Chellaswamy, P., &Kamalam,C.J.R. N. R. (2025). Attention-enhancedDeepRetiNet for robust hard exudates detection in diabetic retinopathy. Biomedical Signal Processing and Control, 100, 106903.

[15]. Navaneethan, R., & Devarajan, H. (2024). Enhancing diabetic retinopathy detection through preprocessing and feature extraction with MGA-CSG algorithm. Expert Systems with Applications, 249, 123418.

[16]. Romero-Oraá, R., Herrero-Tudela, M., López, M. I., Hornero, R., & García, M. (2024). Attention-based deep learning framework for automatic fundus image processing to aid in diabetic retinopathygrading. Computer Methods and Programs inBiomedicine, 249, 108160.

[17]. Oulhadj, M., Riffi, J., Khodriss, C., Mahraz,A. M., Yahyaouy,A., Abdellaoui, M., ... & Tairi, H. (2024). Diabetic retinopathy prediction based on vision transformer and modified capsule network. Computers in Biology and Medicine, 175, 108523.

[18]. Usman, T. M., Saheed, Y. K., Ignace, D., & Nsang, A. (2023). Diabetic retinopathy detection using principal component analysis multi-label feature extraction and classification. International Journal of Cognitive Computing in Engineering, 4, 78-88.

[19]. Bilal,A., Zhu, L., Deng,A., Lu, H., & Wu, N. (2022).AI-based automatic detection and classification of diabetic retinopathy using U-Netand deep learning. Symmetry, 14(7), 1427.

[20]. Vives-Boix, V., & Ruiz-Fernández, D. (2021). Diabetic retinopathy detection through convolutional neural networks with synaptic metaplasticity. Computer Methods and Programsin Biomedicine, 206, 106094.

[21]. Kamble, V. V., & Kokate, R. D. (2020).Automated diabetic retinopathy detection using radial basis function. Procedia Computer Science, 167, 799-808.

[22].Pao,S.I.,Lin,H.Z., Chien,K.H.,Tai,M.C.,Chen,J.T.,&Lin,G. M.(2020).Detection of diabetic retinopathy using bichannel convolutional neural network. Journal of Ophthalmology, 2020(1), 9139713.

[23]. Aziz, T., Charoenlarpnopparut, C., & Mahapakulchai, S. (2023). Deep learning-based hemorrhage detection fordiabetic retinopathy screening. Scientific Reports, 13(1), 1479.

[24]. Zago, G. T.,Andreão, R. V., Dorizzi, B., & Salles, E. O. T. (2020). Diabetic retinopathy detection using red lesion localization and convolutional neural networks. Computers inbiology and medicine, 116, 103537.

[25]. Setiawan,A. W., Mengko, T. R., Santoso, O. S., & Suksmono,A. B. (2013, June). Color retinalimageenhancementusingCLAHE.In InternationalconferenceonICTforsmart society (pp. 1-3). IEEE.

[26]. Ganguly, S., Ganguly, S., Srivastava, K., Dutta, M. K., Parthasarathi, M., Burget, R., & Riha, K. (2014, November).An adaptive threshold based algorithm for detection of red lesions of diabetic retinopathy in a fundus image. In 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom)(pp. 91-94). IEEE.

[27].Abràmoff, M. D., Folk, J. C., Han, D. P.,Walker, J. D.,Williams, D. F., Russell, S. R., ... & Niemeijer, M. (2013). Automated analysis of retinal images for detection of referablediabetic retinopathy. JAMA ophthalmology, 131(3), 351-357.

[28]. Karthik, Maggie, and Sohier Dane. APTOS 2019 Blindness Detection. https://kaggle.com/competitions/aptos2019-blindness-detection, 2019. Kaggle.

[29]. Alahmadi,M.D.(2022).Textureattentionnetworkfordiabeticretinopathy classification. IEEE Access, 10, 55522-55532.

[30]. Zhang, C., Chen, P., & Lei, T. (2023). Multi-point attention-based semi-supervised learning for diabetic retinopathy classification. Biomedical Signal Processing and Control, 80, 104412.

[31]. Ali, G., Dastgir, A., Iqbal, M. W., Anwar, M., & Faheem, M. (2023). A hybrid convolutional neural network model for automatic diabetic retinopathy classification from fundus images. IEEE Journal of Translational Engineering in Health and Medicine, 11, 341- 350.