Glaucoma classification using a polynomial-driven deep learning approach

Krishna Santosh Naidana, Soubhagya Sankar Barpanda
School of Computer Science and Engineering, VIT-AP University, Amaravati, India

Article Info

ABSTRACT

In this paper, a deep learning-based multi-stage polynomial driven glaucoma classification-net (PDGC-Net) has been proposed for glaucoma identification through retinal images. The proposed approach begins with retinal image purification by noise estimation and reduction. Noise has been estimated using a polynomial coefficient-based approach. Images are classified using PDGC-Net, whose polynomial indeterminate representative blocks are designed using new convolutional neural networks (CNN) architectures. The performance of PDGC-Net has been observed on the ACRIMA, ORIGA, and retinal image database for optic nerve evaluation (RIM-ONE) datasets. The experimentation is carried out on noisy and denoised images separately, and PDGC-Net has achieved 96% to 98% and 98% to 100% accuracy ranges, respectively. The model's elasticity is tested with various stages of PDGC-Net. The quantitative PDGC-Net performance analysis is done with state-of-the-art CNN models. The proposed model's performance has been proven and could be an effective aid to ophthalmologists for glaucoma screening (GS).

Corresponding Author:

Krishna Santosh Naidana
School of Computer Science and Engineering, VIT-AP University
Inavolu, Beside AP Secretariat, Amaravati-522237, AP, India
Email: krishna.18phd7018@vitap.ac.in

1. INTRODUCTION

Glaucoma is the leading cause of optic disorders in the majority of people worldwide [1]. Glaucoma progression is asymptomatic, and a delay in its detection leads to interminable vision deterioration. Currently, there is no medical practice to restore the outright vision of the glaucomatous eye. The only way to save a patient's vision from glaucoma is to diagnose it early on [2]. Ophthalmologists usually diagnose glaucoma in its early stages by clinically examining the retinal region-of-interest (ROI) such as the optic disc (OD), optic cup (OC), and retinal nerves through fundus (retinal) images in terms of cup/disc (C/D) ratio and the OD rim alterations in terms of inferior superior nasal temporal (ISNT)-order [3]. Physical examination of retinal structures is a time-consuming and error-prone operation. Computer-aided diagnosis (CAD) can assist ophthalmologist in the quick and accurate identification of retinal structures as well as measuring their proportions to find aberrant symptoms [4]. The CAD-based approach performs retinal image analysis in various phases: pre-processing to improve image texture, feature extraction to capture the patterns, and classification to assign a class label to the given instance. In this paper, a unique deep learning (DL)-based CAD approach has been proposed for the automatic screening of glaucomatous retinal images. This approach has included the major CAD phases, such as images being pre-processed with the proposed denoising method, and features being extracted and classified

using the proposed DL model: polynomial driven glaucoma classification-Net (PDGC-Net).

− Related works

DL-based glaucoma screening (GS) state-of-the-art practices usually follow a three-phase pattern. Most of the pre-processing approaches deals with image denoising that manipulate individual pixels w.r.t. neighboring pixels using various kernels. He *et al.* [5] utilized bilateral filters for denoising. Elseid *et al.* [6] applied Gaussian, mean, guided, and adaptive mean filters for noise suppression. Subsequently, an nth-order non-linear Wiener filter was applied in [7] for image purification. Erwin *et al.* [8] applied various filters on the green channel of noisy images. Afterwards, soft thresholding is applied to retinal images using shearlet transformation by [9]. Juneja *et al.* [10] employed a window-based approach using an adaptive median kernel on grayscale pixel intensities. Hu *et al.* [11] recently applied shearlet filters to retinal images after noise redistribution. Following purification, retinal images can be used for auto-screening in subsequent CAD phases. The majority of DL techniques in the literature often perform the feature extraction and classification phases simultaneously. A six-layered convolutional neural network (CNN) was introduced by [12], but its performance is quite low due to its simple architecture. Later, the OD area was extracted and processed using a two-way CNN by [13] for GS. However, both CNN branches were similar in architecture. Thereafter, a multi-parallel branched (MB) CNN was introduced by [14] to capture various retinal image features. In this approach, one of the MB CNN's branch is employed for non-ROI's features extraction, which are less significant for GS. Pinto *et al.* [15] utilized transfer learning (TL) on ROI of public datasets with VGG, Xception, Inception V3, and ResNet50. Afterwards, Juneja *et al.* [10] introduced a 76-layered glaucoma classification-Net (GC-Net) CNN for the classification of filter-based denoised retinal input. However, only Conv2D layers were employed, and its variants were not tested. Subsequently, Juneja *et al.* [16] utilized filter-based retinal image denoising and ROI was segmented by the proposed G-Net for GS. Though the proposed CNN is deep, residual connections were not utilized to avoid gradient vanishing problem. Next an ensemble AlexNet model is employed by [17] for GS using threshold-based denoising images. In this study, a single CNN is employed for entire ensembling. Shinde [18] captured the ROI of the input with the bright-spot approach and segmented by the U-Net for structural features. Recently, a two-branched DL network (TWEEC) is proposed by [19] for the classification of denoised retinal images in wavelet domain. In this study, both CNN branches were designed with similar architecture.

The survey made on state-of-the-art GS practices revealed that there are two major choices for employing DL-based diagnosis. The primary one is to utilize existing CNN models (visual geometry group (VGG), AlexNet, and ResNet) by using transfer learning (fine-tuning). The other one is designing and training new CNN architectures (GC-Net, TWEEC-CNN, and MB-CNN) for optimal GS. This study has identified existing approaches' limitations such as; i) most of the approaches reduce the noise without estimating it properly, which can eradicate significant textural patterns, ii) although using CNN models via TL minimizes experimental overhead, it does not address the input-model complexity trade-off, iii) only Conv2D layers are utilized by the current approaches to develop CNN architectures; further variations like depth-wise and separable convolutions are not taken into account.

− Research contribution

To address these identified research gaps, we have proposed a new polynomial-driven CAD-based approach for optimal glaucoma diagnosis. This approach includes the following contributions: i) the retinal image purification has been done with a denoising approach that applies legitimate noise estimation using a polynomial coefficient-based technique, ii) a multi-staged PDGC-Net DL approach has been designed based on indeterminates (variables) of a polynomial expression to capture the image patterns using depth-wise and separable convolutions, and iii) the proposed PDGC-Net has a provision for size and shape elasticity w.r.t. the input complexity. The rest of the paper is organized as follows: section 2 describes the phase-wise proposed CAD approach, section 3 performs a deep investigation of the obtained results, and section 4 concludes the article with future scope.

2. METHOD

In this study, ACRIMA (normal: 309 and glaucoma: 396), ORIGA (normal: 482 and glaucoma: 168), and retinal image database for optic nerve evaluation (RIM-ONE) (normal: 458 and glaucoma: 325) datasets' ROIs are considered individually as well as jointly (ACRIMA+ORIGA+RIM-ONE (AOR)) for glaucoma classification. The overall architectural view of the proposed method is given in Figure 1, and each phase

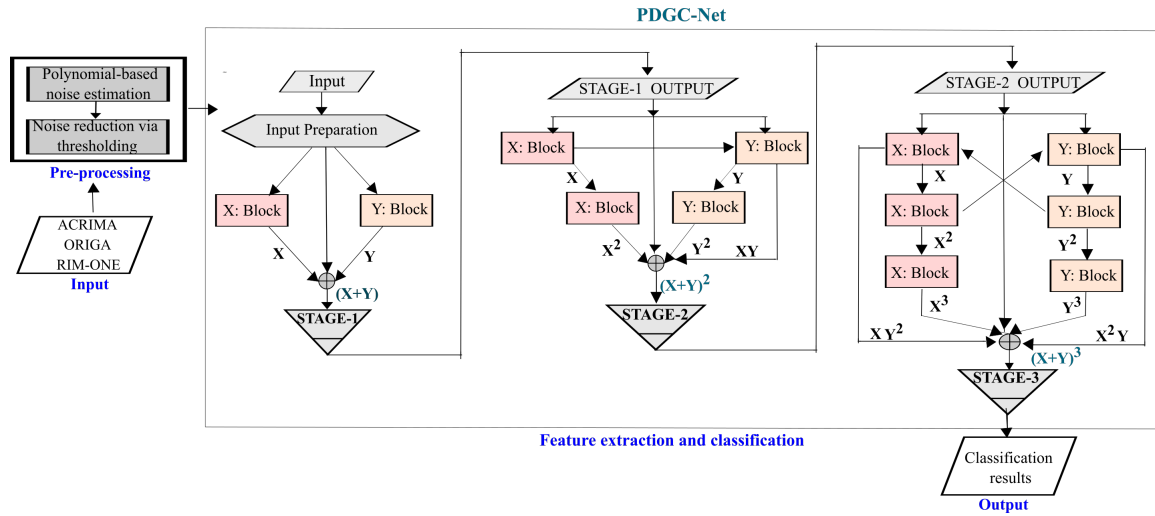is described in detail in the following sub-sections.

Figure 1. The overall architectural view of the proposed CAD approach

2.1. Pre-processing

Retinal imaging is a live procedure that requires adequate lighting, accurate positioning, and quick acquisition. These are highly sensitive to external factors (e.g., sensor quality and external light), which results in the capture of noisy images. The existence of noise seriously affects the automatic processing of images. Fundus image structures, in particular, such as OD and OC, are not in a fixed shape and do not have strong edges. Thus, the presence of noise becomes a severe hurdle for convolutions to discriminate edge, texture, and noise components [20]. This results in unwanted pattern learning. Thus, in this approach, the fundus image's denoising is considered for pre-processing. This phase consists of two sub-phases: noise estimation (Algorithm 1) and reduction (Algorithm 2), which are discussed w.r.t. the corresponding algorithm steps (S: no).

− Noise estimation: the process begins by dividing colour channels (Ch) of input (IMG_{IN}) into N_B blocks (Blk_C) of size n (S: 2). Discrete wavelet transformation (DWT) sub-bands (SB) W_e of each Blk_C are utilized to construct a coefficient matrix $DwtMat_{coeff}$ of size $W_e \times N_B$ (S: 3). Row-wise energy (E_{row}) of $DwtMat_{coeff}$ is measured (S: 5) to generate G_{mat} (S: 6). Initial noise deviation (N_{SD}) is obtained from i_L^{th} row of G_{mat} corresponding to the minimum E_{row} (S: 8). Column-wise variation (C_{var}) of G_{mat} is measured (S: 9) to discard non-homogeneous columns ($G_{mat-rec}$) using a threshold (S: 10-13). This estimates the initial noise (N_{IE}) (S: 14). Then, the final noise (N_{FE}) is estimated using polynomial coefficients (P_{coeff}) obtained from X_{vec}, Y_{vec}, and, W_{vec} (S: 19). X_{vec} represents known noise levels (S: 15). Non-edge areas (C_{ne}) of retinal datasets are calculated using [21] and enhanced by the procedure introduced in our previous contribution [22]. These are considered a standard for noise estimation. Y_{vec} (S: 17) is formed using these standard retinal images' non-edge area average (AVG) values. The weight vector (W_{vec}) is diagonal ($Diag$) values of the X_{vec} function (S: 18). Using P_{coeff}, the final noise ($N_{FE_{ch}}$) is estimated (S: 20).

− Noise reduction: an image-specific threshold (ThL) is determined (S: 4) using a scaling factor (F_S) (S: 2) and estimated $N_{FE_{ch}}$. Using ThL, DWT coefficients of Ch are depreciated to remove the noise (S: 7). Finally, denoised images (REC_{IMG}) are reconstructed using fine-tuned SBs (dwt_{Ch}^{new}) (S: 9).

2.2. Feature extraction and classification

Glaucoma classification is carried out using the polynomial driven glaucoma classification-Net (PDGC-Net). Its architecture is composed of 85 layers (both sequential and parallel) based on the two indeterminates (i.e., X and Y) of an nth order (degree) polynomial expression (constant coefficients are omitted). The idea of PDGC-Net originated from the PolyNet [23]. The proposed PDGC-Net's X: block and Y: block are constructed using new approaches in three different stages, as follows:

− Stage-1: this stage is configured based on a linear polynomial that accepts preprocessed input (IN_{Pre}) and processed using X and Y blocks. Algebraically, the stage-1 operation can be reflected as in (1).

$$Stage: 1_{OUT} = (I + X + Y) \cdot IN_{Pre} = IN_{Pre} + X(IN_{Pre}) + Y(IN_{Pre}) \tag{1}$$

- Stage-2: this stage is configured based on a quadratic polynomial that accepts the $Stage: 1_{OUT}$. Algebraically, the stage-2 operation can be expressed as in (2).

$$Stage: 2_{OUT} = \left(I + (X + Y)^2\right) \cdot Stage: 1_{OUT} = Stage: 1_{OUT} + X(X(Stage: 1_{OUT})) \\ + Y(Y(Stage: 1_{OUT})) + Y(X(Stage: 1_{OUT})) \tag{2}$$

- Stage-3: this stage is configured based on cubic polynomial that accepts the $Stage: 2_{OUT}$. Algebraically, the stage-3 operation can be characterized as in (3).

$$Stage: 3_{OUT} = \left(I + (X + Y)^3\right) \cdot Stage: 2_{OUT} = Stage: 2_{OUT} + X(X(X(Stage: 2_{OUT}))) \\ + Y(Y(Y(Stage: 2_{OUT}))) + Y(X(X(Stage: 2_{OUT}))) + X(Y(Y(Stage: 2_{OUT}))) \tag{3}$$

The entire functioning of the PDGC-Net is characterised by X and Y blocks, whose structural layout is given in Figure 2 and described in the following sub-sections.

Algorithm 1 Retinal images' noise estimation

Require: ROIs of retinal datasets
Ensure: RGB channels' final estimated noise $N_{FE_{Ch}}$
1: **for** $IMG_{Ch} \in IMG_{IN}$ **do**
2:     $Blk_C \leftarrow Split(IMG_{Ch}); \{SB\} \leftarrow DWT(Blk_C),\ where\ C = 1\ to\ N_B$
3:     $[DwtMat_{coeff}]_{W_e \times N_B} \leftarrow append_{row-wise}(SB(:))$
4:     **for** $q = 1\ to\ W_e$ **do**
5:         $E_row(q) \leftarrow \sum_{p=1}^{N_B} DwtMat_{coeff}(p, q)$ ▷ Row-wise energy calculation
6:         $G_{mat} \leftarrow Extract_{1\ :\ L}(Sort(E_{row}(q))),\ L < q$ ▷ G_{mat} calculation
7:     **end for**
8:     $N_{SD} \leftarrow K * V_{med}(G_{mat}(i_L, c) - R_m)$ c=1,2,...,N_B,$0 < K < 9$ ▷ Initial noise deviation
9:     $C_{var}(n) \leftarrow \frac{1}{L} \sum_{l=1}^{L} G_{mat}(l, n)^2,\ n \in N_B$
                                            ▷ G_{mat} column-wise variation
10:    $T_{N_{SD}} \leftarrow N_{SD}^2 \left(1 + \sqrt{\frac{K-1}{L}}\right)$ ▷ Threshold for homogeneous columns
11:    **if** $C_{var}(n) > T_{N_{SD}}$ **then**
12:        $G_{mat-rec} \leftarrow Discard_{n^{th}column}(G_{mat})$
13:    **end if**
14:    $N_{IE} \leftarrow (K * V_{med}(G_{mat-rec}(i_L, r) - R_m))^2$ ▷ Initial noise estimation
15:    $X_{vec} \leftarrow [10, 20, 30, 40, 50]$ ▷ X_{vec}, Y_{vec} construction
16:    $SqMean = \sqrt{Mean\left[\left(\sum_{u,v \in SB} C_{ne}^2(u, v)\right)\right]}, SB \in HH, LH, HL$
17:    $Y_{vec} \leftarrow AVG_{X_{vec}}\left(\frac{2}{C_S} * SqMean\right)$
      C_s : Constant
18:    $W_{vec} \leftarrow Diag\left[\exp\left(X_{vec} - \frac{X_{vec}^2}{2*SD(X_{vec})}\right)\right]$ ▷ Weights generation
19:    $P_{coeff} \leftarrow polyfitweighted(X_{vec}, Y_{vec}, W_{vec}, D_n)$ ▷ Polynomial coefficients generation
20:    $N_{FE_{Ch}} \leftarrow \sum_{d=D_n}^{0} P_{coeff_d} * (N_{IE})^d$ ▷ Final noise estimation
21: **end for**

Algorithm 2 Retinal images' noise reduction

Require: Estimated retinal RGB channel-wise noise ($N_{FE_{Ch}}$)
Ensure: Denoised retinal images (REC_{IMG})

1: **for** $IMG_{Ch} \in IMG_{IN}$ **do**

2:      $F_{S_{Ch}} \leftarrow \sqrt{\frac{\log\left(Med_{HL(r,c)}\right)}{2^{Skw_{HL(r,c)}}}/SD_{HL(r,c)}}$ ▷ Threshold calculation

3:      $N_{FE_{Ch}} \in IMG_{Ch},\ HL \in dwt_{2D}\left(IMG_{Ch}\right)$

4:      $ThL_{Ch}(r,c) \leftarrow F_{S_{Ch}} * N_{FE_{Ch}}$

5: `//Median (Med), Skewness(Skw), and SD are generated on local`
   `neighbourhood(`n_L`) of corresponding DWT's sub-bands.`

6:   ▷ DWT sub-band tuning

7:      $dwt^{new}_{Ch(r,c)} \leftarrow dwt_{2D}\left(IMG_{Ch(r,c)}\right) * ThL_{Ch}(r,c)$

8: **end for**   ▷ Denoised image reconstruction

9: $REC_{IMG} \leftarrow DWT_{Inverse}\left(dwt^{new}_{Ch(r,c)}\right)$

2.2.1. Input preparation

The objective is to prioritize input channels using a content-aware approach. Initially, each channel is processed using a convolutional (Conv2D) layer of sixteen 3×3 filters and squeezed with a 2D global-average (GA) pooling. Two channel complexity reduction (CCR) ratios $R1$ and $R2$ have been chosen based on the 50% and 25% of existing filters, i.e., $R1$ is 8 and $R2$ is 4. The channels are then smoothed using parallel, fully connected (FC) layers for a soft channel gating function. The average of these results is used to prioritize (weight) each feature map before being passed to the X: block and Y: block.

2.2.2. X:block

The X: block has been constructed using Conv2D, depth-wise (DwConv2D), and separable convolutions (SpConv2D). The operation begins with a sequence of Conv2D and DwConv2D layers of eight, 3×3 filters with relu activation. It is connected to an inverted residual (IR) block that has a sequence ($L_{seq} : X_1$) of three sub-blocks: expansion, depth-wise, and projection. The expansion block contains a Conv2D layer, batch-normalization (BN), and relu activation. Filters are defined using input shape and a scalar value s_v. In depth-wise block, DwConv2D operations are carried out using 6×6 filters with Selu activation. The projection block carried out Conv2D operations with 24 filters and BN. The outcome is added to the IR block's input using residual connections (dashed lines in Figure 2). The purpose of $L_{seq} : X_1$ is to expanded initial dimension features and then compress output features to get back to their initial dimensions. The tensor output is then processed by a sequence of Conv2D layer, DwConv2D, and SpConv2D layer with selu and sigmoid activations. Now the result is again processed with the sequence $L_{seq} : X_1$ with 32 channels. The flow is then passed through a flipped sequence (F_{seq}) formed by DwConv2D-SpConv2D-DwConv2D (D-S-D) layers with 32 and 64 filters. The result is then added to the input using ResNet-of-ResNet (RoR) [24] connections. The final output is produced by a Conv2D layer with thirty two, 3×3 filters with relu activation. In stage-2, the $L_{seq} : X_1$ is attached to $L_{seq} : X_2$, which includes all blocks with 32 channels, and the F_{seq} is stretched with the D-S-D-S-D sequence. In stage-3, both $L_{seq} : X_1$ and $L_{seq} : X_2$ are attached to $L_{seq} : X_3$ sequence with 64 channels, and the F_{seq} is improved with the D-S-D-S-D-S-D sequence. All the residual and RoR connections are maintained accordingly.

2.2.3. Y: block

In addition to convolutions, the output channels weightage is utilized in Y: block that has been partitioned into 8 sub-blocks: $Y_{11}, Y_{12}, Y_{21}, Y_{22}, Y_{31}, Y_{32}, Y_{41},$ and Y_{42}. Each Y_{ij} performs Conv2D, DwConv2D, or SpConv2D on the input with varying channels (ch) and kernel sizes. Then GA pooling is applied to the output and passed it to two branches for their complexity scale down using complexity-reduction (CR) ratios: *ratio1* and *ratio2*. It is followed by the application of non-linearity using Act_f : *relu, selu,* or *sigmoid* activations, which are then passed to the next FC layers. The products of C_{opr} feature maps and outputs are combined to form a weighted sub-branch output (Y_{ij-Out}). Though these sub-blocks (Y_{ij}) are structurally similar, their operation differs depending on the parameters $C_{opr}, ch,$ and Act_f, which are defined as:

- $Y_{11} : Conv2D, 8, relu$
- $Y_{21} : DwConv2D, 16, selu$
- $Y_{31} : SpConv2D, 32, sigmoid$
- $Y_{41} : Conv2D, 64, relu$

- $Y_{12} : Conv2D, 64, relu$
- $Y_{22} : SpConv2D, 32, sigmoid$
- $Y_{32} : DwConv2D, 16, selu$
- $Y_{42} : Conv2D, 8, relu$

The Y_{i1} sub-blocks' parameter order is reversed for Y_{i2} to make the Y: Block robust. Parallel connections are also implemented for sub-block inter-dependency by averaging $Y_{i,1}$ and $Y_{i,2}$ and adding them to the $Y_{i+1,1}$ and $Y_{i+1,2}$ resultants. The output feature maps are added with the input using residual connections to generate the final Y: block's output. The performance of PDGC-Net is deeply investigated in the next section.
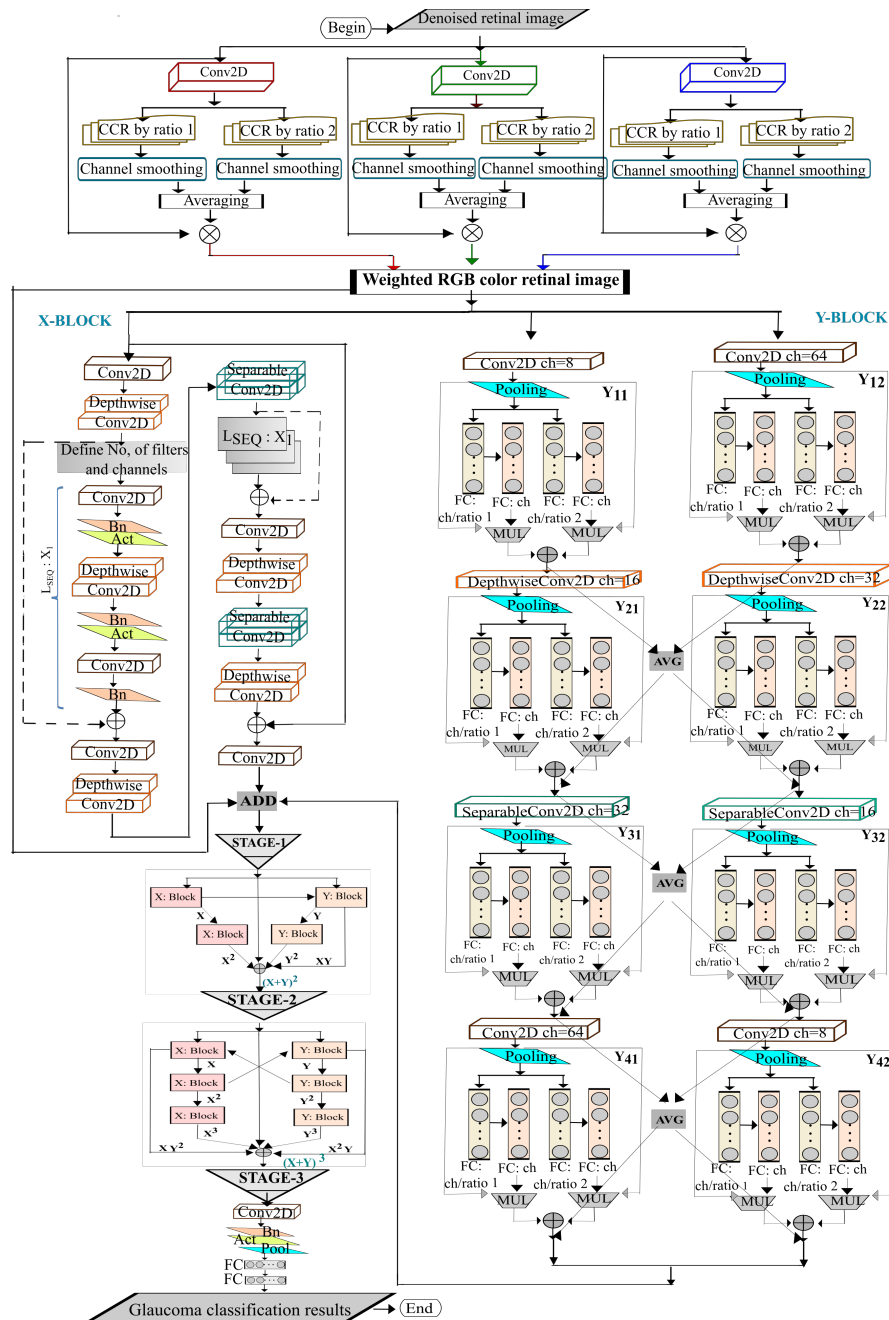

Figure 2. In-depth architectural view of the proposed PDGC-Net's X and Y blocks

2.3. The novelty of the proposed approach

The proposed DL model for glaucoma classification is constructed by considering polynomials of various degrees which has the following unique characteristics:

− Each retinal image's channel-wise noise has been estimated using a weighted polynomial coefficients-based approach before being reduced. This avoids the significant image texture patterns' erosion
− New independent architectures are used for the construction of PDGC-Net blocks corresponding to polynomial indeterminates. The X: Block mainly concentrated on pattern extraction using various convolutional operations along with residual connections. The Y: Block extracts the image texture information through channel prioritization and complexity reduction
− The RoR connections in X: block provide a vanishing gradient-free model. The feed-forward connections among sub-blocks of Y: block cause the model to receive mixed output channels
− The PDGC-Net is built with a flexible (i.e., elastic) architecture. Thus, it can be extended to the next stages (i.e., stages 4 and 5) by considering the higher degree (4th, and 5th) polynomials based on the input complexity

3. RESULT AND DISCUSSION

The performance of the PDGC-Net has been evaluated before and after image denoising. The noise estimation begins by using 5th (D_n) order polynomial coefficients and some sample images' (15 images) channel-wise (R_c, G_c, and B_c) estimated noises are shown in Table 1. While doing this, each ROI's colour channel is resized into 256×256 dimensions and divided into 8×8 (i.e., n = 8) blocks. This results in 1024 (N_B) square blocks. The corresponding identified channel-wise noise levels (for 50 retinal ROIs) from the ACRIMA, ORIGA, and RIM-ONE datasets are shown in Figure 3. In the ACRIMA dataset, higher noise levels are estimated in the images' red channels, as shown in Figure 3(a). In a few cases, more noise is detected in the blue and green channels. However, blue and green color channels are almost noise-free. For the ORIGA retinal dataset, higher noise levels are identified in both the red and green colour channels, as shown in Figure 3(b). In some cases, the images' blue colour channel is associated with higher noise levels than the other channels. All channels of ORIGA retinal images are not completely noise-free. In the case of RIM-ONE, most of the noise exists in the red channels of images, as shown in Figure 3(c). Both the green and blue channels are completely noise-free for most of the images. In very limited cases, all three colour channels are noise-free.

Table 1. Retinal ROI's noise estimation by the proposed method

Img. No.	ACRIMA			ORIGA			RIM-ONE		
	Rc	Gc	Bc	Rc	Gc	Bc	Rc	Gc	Bc
Normal retinal images (NI)									
1	0.0229	1.1768	1.5812	0.0102	3.5039	4.1788	0.9790	0.0110	0.0111
2	0.1742	0.0029	0.0042	15.949	1.9611	11.565	0.0051	0.0000	0.0011
3	15.357	2.3128	3.1631	1.8014	1.7866	10.171	4.2092	0.1576	0.1025
4	4.4754	1.4268	1.9272	13.037	5.4454	4.8626	6.1212	0.0107	0.0107
5	9.0567	1.8134	2.4649	2.3836	0.8161	3.5031	16.639	0.0587	0.0409
6	9.0562	0.3716	0.4825	16.976	0.8161	3.5031	5.3857	0.0323	0.0230
7	9.0400	6.1832	8.6450	16.976	0.1356	0.3389	1.6805	0.0114	0.0114
Glaucomatous retinal images (GI)									
8	7.6451	0.4647	0.6076	3.8208	9.9385	6.0135	2.0816	0.1617	0.1050
9	6.5094	3.8440	5.3202	3.1049	0.1928	0.5279	1.3271	0.0017	0.0025
10	3.8186	0.0814	0.1011	5.9865	1.5509	8.3757	2.5142	0.0141	0.0098
11	0.7682	0.0113	0.0113	17.387	4.3002	4.4818	0.9714	0.2766	0.1711
12	0.9865	0.9051	1.2071	8.1655	0.2764	0.8384	1.0384	0.1034	0.0695
13	5.5943	9.8602	13.910	3.0045	0.4260	1.4749	15.914	0.4023	0.2399
14	12.480	1.7425	2.3661	0.8117	1.8896	10.988	11.306	0.3710	0.2230
15	0.5267	12.393	17.555	1.0290	0.0708	0.1531	5.1172	1.9995	1.0123

The qualitative sample original and denoised fundus images from ACRIMA, ORIGA, and RIM-ONE datasets are shown in Figures 4(a)–(c). In this approach, no external noise has been added to fundus images prior to the noise estimation. Fundus images are denoised without affecting the texture, based on the existing noise estimation. By careful observation, one can identify the clear appearance of retinal structures such as

the OD, OC, and blood vessels in denoised fundus images. Especially, the boundaries of OD and OC and the retinal area in between blood vessels are now clearly discernible.
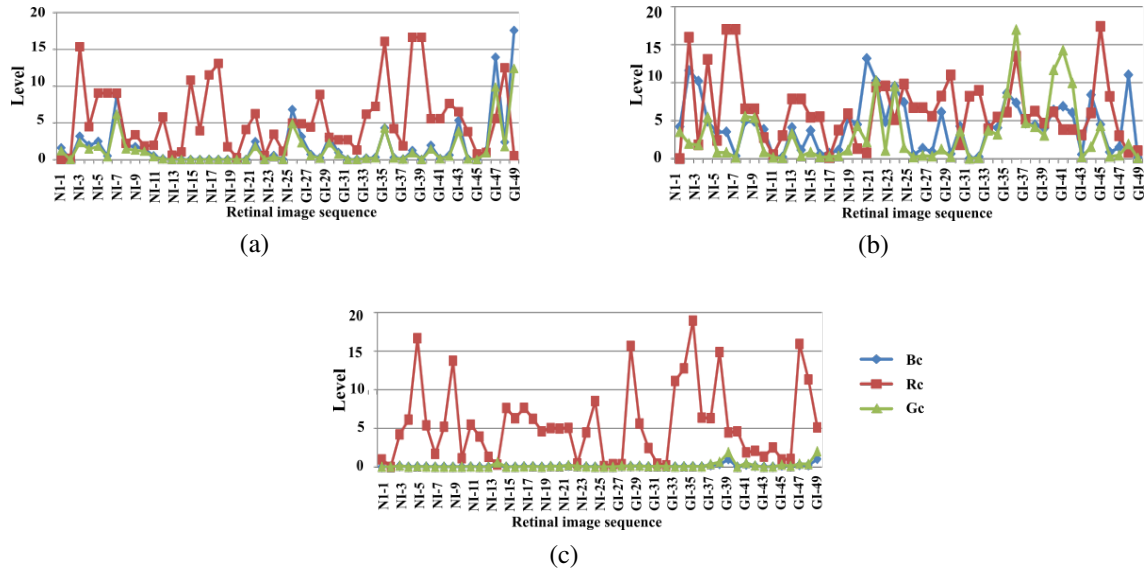

(a)

(b)

(c)

Figure 3. Sample noise estimation plots (a) ACRIMA, (b) ORIGA, and (c) RIM-ONE

Original   Denoised   (a)   Original   Denoised

Original   Denoised   (b)   Original   Denoised

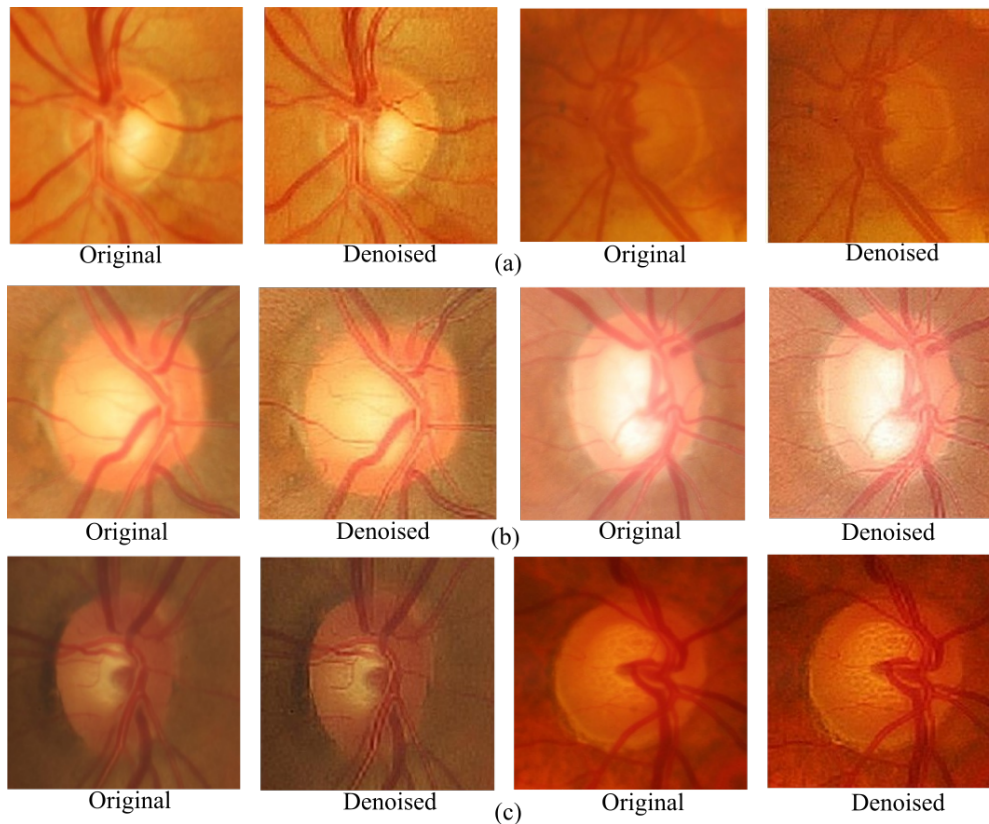Original   Denoised   (c)   Original   Denoised

Figure 4. Sample qualitative denoised retinal images from (a) ACRIMA, (b) ORIGA, and (c) RIM-ONE datasets

        The performance of the proposed image purification is measured quantitatively using mean squared error (MSE) and peak signal-to-noise ratio (PSNR) metrics. Sample (for the same 50 retinal ROIs from Figure 3) MSE value plots for ACRIMA, ORIGA, and RIM-ONE datasets are shown in Figures 5(a)–(c). The corresponding PSNR plots are shown in Figure 5(d)–(f). The lower MSEs indicate that the reconstructed images are very close to the original images. That means the image purification approach preserves the retinal structures. A higher PSNR value implies a higher-quality reconstructed image. In Figure 5, both MSE and PSNR plots have been included with and without applying noise estimation to fundus images. The proposed polynomial coefficient-based noise estimation improved both plots by minimising MSE values and maximising PSNR values. In very limited cases, denoised images' noise estimation-based MSE and PSNR values are close to the values obtained without noise estimation. Overall, the quality of denoised images has been improved with the proposed polynomial coefficient-based noise estimation and reduction.
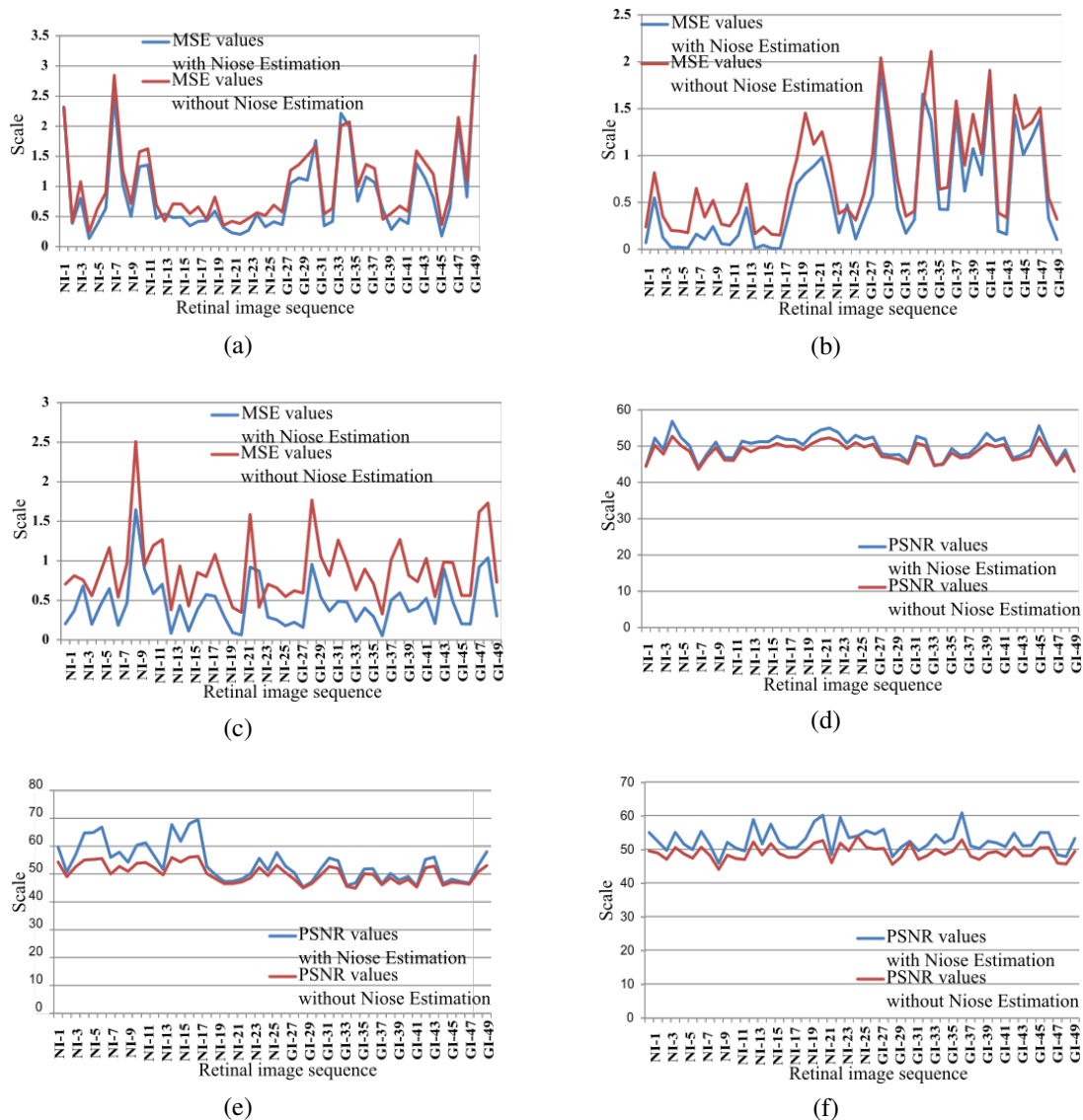

Figure 5. Sample denoised images MSE and PSNR value plots by the proposed approach; (a) ACRIMA data, (b) ORIGA dataset, (c) RIM-ONE dataset, (d) ACRIMA data, (e) ORIGA dataset, and (f) RIM-ONE dataset

        The proposed image denoising strategy is quantitatively compared with the well known denoising techniques such as: approaches by [25], [26], DnCNN [27], WNNM [28], and PGPD [29]. Table 2 presents the

corresponding average MSE and PSNR values obtained from the entire fundus datasets. The values obtained by the polynomial coefficient-based denoising approach are more optimal than those obtained by the DnCNN and PGPD approaches. The WNNM approach reconstructed the denoised images better than the [25], [26] approaches. However, in the proposed denoising approach, there is a considerable improvement in both the MSE and PSNR value ranges with the presence of noise estimation. The ACRIMA, ORIGA, and RIM-ONE retinal images have been qualitatively purified and restored with average MSEs of 0.5524, 0.5406, and 0.4333 and PSNRs of 53.4178, 53.5709, and 53.4320, respectively. Instead of adding external noise, retinal images' pixel intensities are carefully analysed in terms of DWT SB's energies with weighted polynomial coefficients for proper noise estimation. As a result, the original intrinsic noises (if they exist) are correctly identified and eliminated without affecting the significant texture patterns. This enables the reconstruction of denoised images that are closer to the original retinal images, which is a highly significant characteristic for medical images. Thus, optimal MSE and PSNR values are obtained on all considered retinal datasets. This emphasises the significance of accurate noise estimation in retinal (medical) image denoising.

The proposed PDGC-Net's training is performed by randomly selecting 70% of the retinal dataset images (ACRIMA, ORIGA, RIM-ONE, and AOR), and remaining 30% of images were utilized for testing and validation purposes. The PDGC-Net has been evaluated with several hyper-parameter combinations before settling on a learning rate (LR) of 0.01 and batch size (BS) of 32 with 30 epochs. Table 3 shows the PDGC-Net's stage-wise classification performance (in terms of percentages) before and after image preprocessing. In the case of the ACRIMA dataset, the PDGC-Net has attained 100% accuracy in stage-3 (i.e., w.r.t. cubic polynomials) with the denoised images and 96% accuracy, 97% sensitivity, and 93% specificity with the original (noisy) images. When the model is shrunk vertically w.r.t. quadratic polynomials (i.e., stage-2), it attains 98% accuracy, 100% sensitivity, and 92% specificity with purified images, and 96% accuracy, 95% sensitivity, and 100% specificity with original images.

Furthermore, the PDGC-Net has classified the purified and original retinal images with 96%, 94% accuracy, 95%, 92% sensitivity, and 100% specificity, respectively. While classifying the ACRIMA dataset, the PDGC-Net has reached its highest accuracy only in stage-3 with purified images. Its performance is almost similar in both stage-3 and stage-2 in the case of the original ACRIMA images. In stage-1, the PDGC-Net classification performance is consistent with purified images and slightly lower with noisy images. When it comes to true and false predictions, the PDGC-Net has attained an extremely low false-positive prediction rate, i.e., 0% to 2% for all experimentation instances. Thus, PDGC-Net has optimally identified glaucomatous retinal images from the dataset.

The false-negative prediction rates are gradually decreased, i.e., healthy retinal image identification is optimal when PDGC-Net is stretched vertically w.r.t. higher-order polynomials. Using the structure corresponding to cubic polynomials, the PDGC-Net has classified the ORIGA dataset with a maximum of 98% accuracy, 89% sensitivity, and 100% specificity. With purified ORIGA retinal images, the performance of the PDGC-Net with the structure corresponding to a second-order polynomial is similar to that of the third-order polynomial equivalent structure. In these cases, the ORIGA dataset has been classified with 96% accuracy, 78%-80% sensitivity, and 100% specificity. In the case of lower-order polynomial-based PDGC-Net, the ORIGA dataset's retinal images were classified with a range of 92%-94% accuracy, 78%-88% sensitivity, and 95% specificity for both the original and denoised cases, respectively. In all test cases, the PDGC-Net's predictions attained a very low percentage of false-negatives, i.e., 2%-4%, which indicates that the proposed model recognizes the non-glaucomatous ORIGA dataset's images perfectly.

However, with the PDGC-Net corresponding to higher-order polynomials, the false-positive prediction rates are dramatically reduced. In the case of the RIM-ONE dataset, the PDGC-Net has exhibited optimal classification performances corresponding to first-, second-, and third-order polynomial-based structures. The proposed DL model performed an optimal screening of RIM-ONE retinal images with ranges of 98%-100%, 92%-100%, and 97%-100% of accuracy, sensitivity, and specificity, respectively. On the RIM-ONE dataset, the false-positive prediction rate is less than 2% and the false-negative prediction rate is less than 5%. This indicates that both glaucomatous and non-glaucomatous retinal images from RIM-ONE are almost accurately identified. In the case of the combined dataset, i.e., the AOR dataset, the PDGC-Net's performance is highly satisfactory in all its stages corresponding to various polynomial orders. Only for the one input case, both sensitivity and specificity measures reached 94% and 92% by the PDGC-Net's classification. In all other cases, the classification measures reached their maximum possible range. With this combined dataset, both false-positive and negative predictions are almost nullified for all stages corresponding to various polynomial orders.

Table 2. Quantitative analysis of the proposed image denoising approach

Input		ACRIMA	ORIGA	RIM-ONE
The proposed denoising with noise estimation	Avg. MSE	0.8130	0.7901	0.6934
	Avg. PSNR	50.0227	50.1418	49.039
The proposed denoising without noise estimation	Avg. MSE	0.5524	0.5406	0.4333
	Avg. PSNR	53.4178	53.5709	53.4320
Approach by [25]	Avg. MSE	2.8266	2.8094	2.8711
	Avg. PSNR	43.6326	43.6575	43.5580
Noise estimation is done by [26] and reduction is done by the proposed approach	Avg. MSE	6.2329	6.1713	5.7733
	Avg. PSNR	40.8991	40.9399	42.6075
DnCNN	Avg. MSE	29.4771	30.1076	19.6199
	Avg. PSNR	33.6318	33.0098	35.2947
WNNM	Avg. MSE	1.9712	2.1575	1.9696
	Avg. PSNR	45.1882	44.7953	45.2045
PGPD	Avg. MSE	10.2399	15.8817	8.5605
	Avg. PSNR	36.0789	35.7797	38.8176

Table 3. The PDGC-Net's glaucoma classification performance

Input		Noise-stage:3	DeNoise-stage:3	Noise-stage:2	DeNoise-stage:2	Noise-stage:1	DeNoise-stage:1
	Acc	96	100	96	98	94	96
	Sen	97	100	95	100	92	95
	Spe	93	100	100	92	100	100
ACRIMA	Tp	0.97	1.0	1.0	0.97	1.0	1.0
	Fn	0.07	0.0	0.13	0.0	0.17	0.13
	Fp	0.02	0.0	0.13	0.0	0.17	0.13
	Tn	0.92	1.0	0.86	1.0	0.82	0. 86
	Acc	96	98	94	96	92	94
	Sen	80	89	83	78	78	88
	Spe	100	100	98	100	95	95
ORIGA	Tp	1.0	1.0	0.90	1.0	0.77	0.77
	Fn	0.04	0.02	0.04	0.04	0.04	0.02
	Fp	0.0	0.0	0.9	0.0	0.22	0.22
	Tn	0.95	0.97	0.95	0.95	0.95	0.97
	Acc	98	100	98	100	98	98
	Sen	92	100	100	100	100	94
	Spe	100	100	97	100	97	100
RIM-ONE	Tp	1.0	1.0	0.94	1.0	0.94	1.0
	Fn	0.02	0.0	0.0	0.0	0.0	0.02
	Fp	0.0	0.0	0.05	0.0	0.05	0.0
	Tn	0.97	1.0	1.0	1.0	1.0	0.97
	Acc	100	100	98	100	98	100
	Sen	100	100	100	100	94	100
	Spe	100	100	92	100	100	100
ACRIMA+ORIGA+RIM-ONE (AOR)	Tp	1.0	1.0	0.97	1.0	1.0	1.0
	Fn	0.0	0.0	0.0	0.0	0.02	0.0
	Fp	0.0	0.0	0.02	0.0	0.0	0.0
	Tn	1.0	1.0	1.0	1.0	0.97	1.0

Note: accuracy (Acc), sensitivity (Sen), specificity (Spe), true positive (Tp), false positive (Fp), true negative (Tn), and false negative (Fn)

Through these four input datasets, the significance of PDGC-Net's flexibility (i.e., its elastic nature) has been clearly verified. Regardless of dataset size, the PDGC-Net, with its structure corresponding to higher-order polynomials, performed optimally in retinal image classification. The PDGC-Net's performance on individual datasets is suffering from a few false predictions with the architecture corresponding to lower-order polynomials (i.e., linear and quadratic). This problem is gradually mitigated and has almost vanished with the PDGC-Net architecture's corresponding cubic polynomials. However, in the case of a large dataset (i.e., the AOR dataset), its performance is optimal even with stage-1. Thus, the proposed PDGC-Net's stages (i.e., polynomial orders) can be extended or reduced vertically based on the retinal dataset size for optimal classification. This exclusive property makes the PDGC-Net more robust in retinal (medical) image classification.

The performance of PDGC-Net is quantitatively compared to that of existing CNN models: VGG16, ResNet50, InceptionV3, EfficientnetB0, MobilenetV2, and SqueezeNet. Various hyper-parameter combina-

tions have been applied to individual DL models as shown in Table 4. These well-known CNNs are designed with various characteristics, such as deep (depth) architectures (VGG16), preservation of gradients (ResNet), optimization (InceptionV3), adaptability (MobilenetV2), advanced convolutions (EfficientnetB0), and compression (SqueezeNet). The PDGC-Net has been designed by considering all these significant properties. Thus, PDGC-Net's performance is compared with the existing CNNs to prove its ability as shown in Table 5.

Table 4. Hyper parameters used by the various DL models for glaucoma classification

DL model	Layers	LR	Optimizer	Inner layer activation	Outer layer activation
VGG16	23	0.01	Adadelta	Relu	Sigmoid
ResNet50	178	0.05	SGD	Relu	Sigmoid
InceptionV3	314	0.1	Adadelta	Relu	Sigmoid
EfficientnetB0	237	0.01	Adam	Relu	Swish
MobilenetV2	53	0.01	Adam	Relu	Sigmoid
SqueezeNet	18	0.01	Adam	Relu	Sigmoid
PDGC-Net	**85**	**0.001**	**Adam**	**Relu, Selu**	**Sigmoid**

Table 5. Existing CNN models' glaucoma classification performances

Input		ACRIMA-noise	ACRIMA-DeNoise	ORIGA-noise	ORIGA-DeNoise	RIM-ONE-noise	RIM-ONE-DeNoise	AOR-noise	AOR-DeNoise
	Acc	94	94	88	94	84	84	94	96
	Sen	93	93	67	78	90	76	94	100
	Spe	100	100	93	98	80	88	93	85
VGG16	Tp	1.0	1.0	0.66	0.87	0.75	0.76	0.97	0.94
	Fn	0.23	0.23	0.07	0.04	0.07	0.12	0.12	0.0
	Fp	0.0	0.0	0.33	0.12	0.25	0.23	0.02	0.05
	Tn	0.76	0.76	0.92	0.95	0.92	0.87	0.87	1.0
	Acc	92	96	92	94	90	94	94	98
	Sen	92	95	57	73	90	93	94	97
	Spe	93	100	98	100	90	94	93	100
ResNet50	Tp	0.97	1.0	0.80	1.0	0.69	0.86	0.97	1.0
	Fn	0.18	0.16	0.06	0.07	0.02	0.02	0.12	0.07
	Fp	0.02	0.0	0.02	0.0	0.30	0.13	0.02	0.0
	Tn	0.81	0.83	0.93	0.92	0.97	0.97	0.87	0.92
	Acc	96	98	78	78	92	94	96	98
	Sen	95	97	71	73	100	82	94	100
	Spe	100	100	83	83	89	100	100	92
InceptionV3	Tp	1.0	1.0	0.75	0.82	0.78	1.0	1.0	0.97
	Fn	0.13	0.07	0.20	0.25	0.0	0.08	0.11	0.0
	Fp	0.0	0.0	0.25	0.17	0.21	0.0	0.0	0.02
	Tn	0.86	0.92	0.80	0.74	1.0	0.91	0.88	1.0
	Acc	92	96	84	88	90	94	94	96
	Sen	100	100	80	88	100	100	92	95
	Spe	64	85	86	98	87	91	100	100
Efficientnet B0	Tp	0.90	0.94	0.70	0.75	0.70	0.83	1.0	1.0
	Fn	0.0	0.0	0.09	0.10	0.0	0.0	0.17	0.13
	Fp	0.09	0.05	0.29	0.25	0.29	0.16	0.0	0.0
	Tn	1.0	1.0	0.90	0.89	1.0	1.0	0.82	0.86
	Acc	94	96	86	88	84	84	94	94
	Sen	94	94	83	85	80	75	92	93
	Spe	93	100	100	100	86	88	100	100
MobilenetV2	Tp	0.97	1.0	1.0	1.0	0.70	0.75	1.0	1.0
	Fn	0.12	0.11	0.14	0.12	0.10	0.11	0.23	0.17
	Fp	0.02	0.0	0.0	0.0	0.29	0.25	0.0	0.0
	Tn	0.87	0.88	0.89	0.87	0.90	0.88	0.76	0.82
	Acc	90	92	84	86	90	90	92	94
	Sen	86	89	89	83	93	75	92	82
	Spe	100	100	83	100	89	97	93	100
SqueezeNet	Tp	1.0	1.0	0.53	1.0	0.77	0.92	0.97	1.0
	Fn	0.26	0.26	0.02	0.14	0.03	0.10	0.18	0.08
	Fp	0.0	0.0	0.46	0.0	0.23	0.07	0.02	0.0
	Tn	0.73	0.73	0.97	0.85	0.96	0.89	0.81	0.89

The existing CNNs performance has been tested both on original and denoised retinal images. In the case of the AOR dataset, VGG16 classification performance is optimal, with an accuracy of 96%, a sensitivity of 100%, and 85% specificity. Its performance is unsatisfactory with the ORIGA dataset without denoising. In this case, classification was done with 88% accuracy, 67% sensitivity, and 93% specificity. In most cases, VGG16 suffers from false-negative and positive prediction rates of 12% to 23% and 12% to 33%, respectively. The proposed polynomial coefficient-based denoising approach boosts the performance of ResNet50 on the AOR dataset with an accuracy of 98%, a sensitivity of 97%, and a specificity of 100%. In the absence of denoising, ResNet50 underperformed, with 90% accuracy, sensitivity, and specificity. However, for the majority of the experimental cases, ResNet50 suffers from erroneous predictions, with a maximum of 18% false-positives and 30% false-negatives. In the case of smaller-sized input, i.e., the ORIGA dataset, InceptionV3's classification performance is dramatically reduced to 78% accuracy, even with the denoised images. The false-prediction rates of InceptionV3 decreased only with the larger datasets. The GS performance of EfficientnetB0 is improved by 3%–4% in accuracy, 4% in sensitivity, and 12%–13% in specificity with the proposed denoising technique. EfficientnetB0 performed better with larger datasets than with smaller datasets. Throughout the experiment, MobilenetV2 suffers from more false negative predictions than false positive predictions. However, the presence of the proposed denoising approach reduces false-prediction rates by 2%–3% on average. The SqueezeNet's performance is reached at its maximum level only with the larger datasets (i.e., the AOR dataset). However, SqueezeNet's glaucoma classification measures are weaker than those of other existing CNN models. Image denoising has improved its auto-screening ability on individual datasets, with an average of 2% accuracy, 3% sensitivity, and 9% specificity.

The majority of existing CNN models' classification performance is not optimal on individual retinal datasets. Their performance is best with the larger dataset, i.e., the combined retinal dataset. In the case of individual datasets, these models either suffer from false-positive predictions or false-negative predictions. This demonstrates that their architectures are inflexible in terms of dataset size and complexity. However, in the majority of cases, the CNN models GS ability is considerably improved with the proposed denoising approach. When compared to existing models' results, the PDGC-Net's classification measures are more stable and ideal, as shown in Figure 6. The PDGC-Net's glaucoma classification accuracy is consistent, ranging from 95% to 100%, whereas the accuracy of other CNN models is partly consistent and has fallen below 80% for some cases. The PDGC-Net's performance is more consistent in the identification of positive cases, i.e., glaucomatous retinal images, than the other existing CNN models. However, in terms of classification sensitivity, there is a huge performance gap between PDGC-Net and other CNN models. Except for a very few cases, the PDGC-Net's sensitivity is consistently above 92%. Existing CNN models' classification performance suffers from heavy fluctuations and attains optimal values in very few instances. It shows strong evidence for the PDGC-Net's outperformance over the other CNN models in the identification of glaucomatous retinal images.
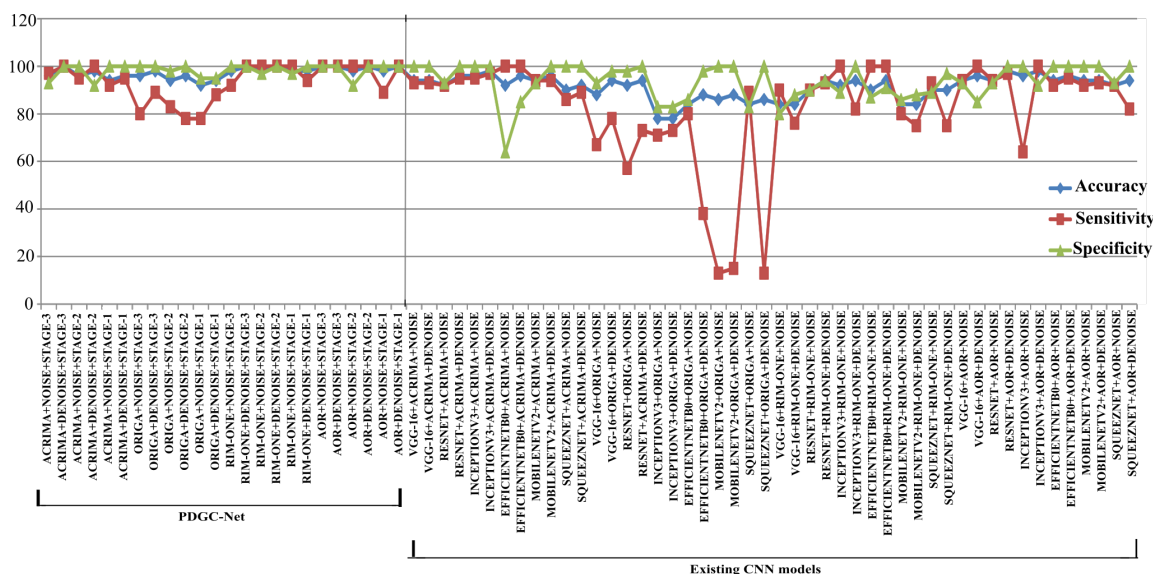

Figure 6. Performance comparison of the PDGC-Net with the other CNN models

The proposed PDGC-Net architecture is carefully designed to avoid overfitting situations using precautionary techniques such as; i) lowering the model's capacity: the model becomes more general with a lower number of trainable parameters. However, too much reduction in the model's complexity results in underfitting. Thus, the model's capacity has been optimized by choosing the number of parameters, i.e., 0.15 million approximately, which is less than the parameters of other DL models: ResNet (23 million), VGG16 (138 million), InceptionV3 (24 million), EfficientneB0 (5.29 million), and MobilenetV2 (3.4 million); ii) applying regularization and dropout layers: L2 regularizer is applied to dense layers, and the L1 regularizer is applied to convolutional layers along with the dropout layers at various stages at a rate of 0.2 (i.e., 20% of inputs are nullified); and iii) data augmentation: it is applied to improve the input size with various transformations. Images are randomly rotated within a range of 30. ROIs are shifted by 0.1 percentage vertically and horizontally. Shearing is applied on input with an intensity of 0.15. Horizontal flipping, along with zooming, is also applied to the input.

The impact of these precautionary techniques is clearly demonstrated in Figure 7 using the accuracy graphs (for some optimal cases). It is observed from the VGG16 plot (Figure 7(a)) that there is no consistency in the validation accuracy. In most of the epochs, the gap between training and validation accuracy is greater, and peaks are inconsistent. In the case of the ResNet plot (Figure 7(b)), the gap between the two plots has gradually increased over the epochs. This could cause model overfitting. Further, plots of InceptonV3 (Figure 7(c)) are initially correlated, but there is a large gap between them for most of the epochs, which is undesirable. While in the plot (Figure 7(d)) of EfficientNetB0, the validation accuracy is constant for the maximum training epochs and suddenly reaches higher values. These rapid changes could not lead to a generic model. In the case of MobileNetV2 (Figure 7(e)), the validation accuracy was steady initially but dropped to its lowest value in the middle of the training. Even after several epochs, there is a large gap between two plots, which could lead to overfitting. SqueezeNet's validation accuracy (Figure 7(f)) is always lower than its training accuracy, which is not suitable for optimal classification. A systematic and gradual increase in accuracy could be identified in the PDGC-Net plots. In the case of ACRIMA (Figure 7(g)), there are minor fluctuations in validation accuracy, but throughout training, a lower gap is maintained between training and validation plots. In the case of the ORIGA DS (Figure 7(h)), initially there is a large gap between two plots since the DS size is moderately low. However, this gap has gradually minimized and both plots are nearly merged after the 20th epoch. The PDGC-Net accuracy plot with the RIM-ONE DS (Figure 7(i)) shows an optimal performance. Both plots exhibit steady progression with minimal gaps between them. Thus, it has been proved that the proposed PDGC-Net is well designed and trained for GS.

The performance comparison of the proposed PDGC-Net with state-of-the-art GS approaches is shown in Table 6. In this comparison, various recent CAD-based GS approaches have been considered that are applied to the ACRIMA, ORIGA, and RIM-ONE retinal datasets. All these approaches split the retinal dataset into a 70:30 ratio for training and testing/validation purposes. Among the considered state-of-the-art approaches, most of them introduced new DL architectures.

The remaining techniques used pre-trained CNN models for glaucoma identification. Several approaches' [30]-[32] classifications on the ACRIMA dataset suffer from either false positives or negative predictions. The PDGC-Net's accuracy on the ORIGA dataset is 98%, which is significantly better than other existing approaches [31], [33]. The performance of TL-based approaches [32], [34] on the ORIGA dataset is affected by false predictions. Though the approaches' [35]-[37] accuracy is above 93% to 98% on the RIM-ONE dataset, they are suffering from either false positives or negative predictions. On the other side, the PDGC-Net outperforms the currently existing methods. The significant observation was that, rather than pre-trained CNNs, newly built, customised CNN architectures performed better in glaucoma classification. Thus, the proposed PDGC-Net has been designed with a flexible architecture for various retinal datasets. The above investigation yields valid inferences such as: i) image channels noise exposure is non-uniform across datasets, and hence, channel-wise noise estimation makes the method more resilient, ii) though the PDGC-Net's architecture is less complex, its performance is on par with the current CNN models, iii) the PDGC-Net's positive and negative predictions are more consistent and robust than the existing CNN predictions, iv) the PDGC-Net's architecture is elastic in nature, i.e., it can be extended to the next stages (i.e., stage-4, and stage-5) by considering the higher degree polynomials, and this has been proved with the experimental results, and v) the feed-forward and RoR connections provide a vanishing gradient-free model.
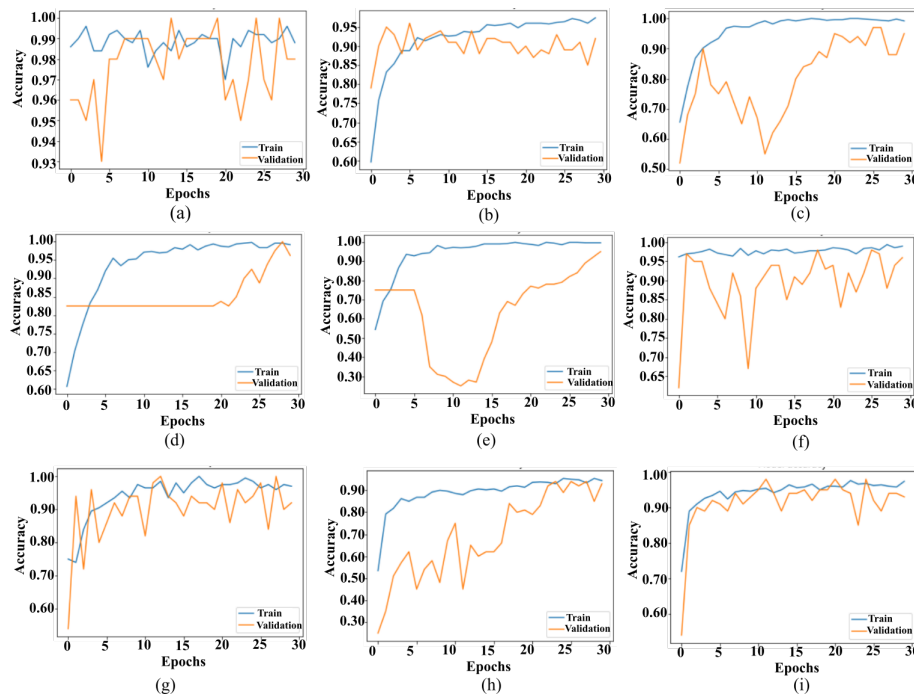
Figure 7. Classification accuracy graphs of (a) VGG16(on ACRIMA), (b) ResNet50 (on ACRIMA), (c) InceptionV3 (on RIM-ONE), (d) EfficientnetB0 (on RIM-ONE), (e) MobileNetV2 (on ACRIMA), (f) SqueezeNet (on ACRIMA), (g) PDGC-Net (on ACRIMA), (h) PDGC-Net (on ORIGA), and (i) PDGC-Net (on RIM-ONE)

Table 6. Performance analysis of the PDGC-Net with the state-of-the-art approaches

Reference	Year	Approach	Train-test data split	Performance		
				Acc	Sen	Spe
		ACRIMA				
[30]	2019	TL by ResNet and GoogLeNet	70:30	70	-	87
[31]	2021	An 18-layered CNN is introduced	70:30	96.64	96.07	97.39
[32]	2022	TL by VGG and ImageNet with the optimization	70:30	98.86	100	98.2
PDGC-Net		A three-stage CNN has been introduced	70:30	100	100	100
		ORIGA				
[31]	2021	An 18-layered CNN is introduced	70:30	78.32	58.06	92.44
[32]	2022	TL by VGG and ImageNet with the optimization	70:30	96.55	97.5	100
[33]	2018	Synthetic minority over-sampling technique (SMOTE)	70:30	76.90	79.90	73.8
[34]	2020	TL by AlexNet	70:30	91.21	92.1	90.56
[38]	2019	ROIs were identified using RCNN and classified by a new CNN model	70:30	95.4	71.17	85
PDGC-Net		A three-stage CNN has been introduced	70:30	98	89	100
		RIM-ONE				
[31]	2021	An 18-layered CNN is introduced	70:30	85.97	81.75	89.3
[35]	2020	Pre-trained CNN models were utilized	70:30	93.15	100	85
[36]	2018	DenseNet with fully convolutions	70:30	98.63	65.85	99.66
[37]	2020	AlexNet with local binary patterns	70:30	98.90	100	97.50
PDGC-Net		A three-stage CNN has been introduced	70:30	100	100	100
		ACRIMA+ORIGA+RIM-ONE (AOR)				
[39]	2022	DCGANs with the hybrid VGG-CapsNet	70:30	90.12	89	86
PDGC-Net		A three-stage CNN has been introduced	70:30	100	100	100

4. CONCLUSION

The proposed CAD-based glaucoma detection technique has been designed by considering polynomials as its basis. Three public retinal datasets were tested independently and in combination in this study.

Retinal images are purified using proper noise estimation during the pre-processing phase. The obtained average MSE and PSNR values were compared to the results obtained using popular image denoising techniques to demonstrate the importance of noise estimation prior to its elimination. Feature extraction and classification are carried out using the proposed PDGC-Net, whose architecture is based on polynomials of various degrees. The architecture of the PDGC-Net is adaptable to the input specifications for optimal classification. The PDGC-Net has been built stage-wise, corresponding to higher-degree polynomials. The PDGC-Net attained 100% classification accuracy, sensitivity, and specificity on the ACRIMA, RIM-ONE, and AOR datasets. On the ORIGA dataset, the proposed model secured 98% accuracy, 89% sensitivity, and 100% specificity. The flexibility (elasticity) of PDGC-Net has been verified stage-wise on individual datasets using various combinations. The PDGC-Net's complexity can be minimised or maximised depending on the size of the dataset, corresponding to lower or higher degree polynomials, without resulting in under- or over-fitting, as demonstrated experimentally. In terms of true and false prediction rates, the role of noise estimation and reduction is validated. The PDGC-Net's performance is compared analytically with the well-known CNN models. The PDGC-Net architecture corresponding to higher degree polynomials mitigate existing CNNs' performance limitations. In future work, we plan to use a variety of CNN model topologies to extend the PDGC-Net's model architecture laterally (more indeterminates) and vertically (more stages) in the context of glaucoma diagnosis.

REFERENCES

[1]   R. R. A. Bourne *et al.*, "Causes of vision loss worldwide, 1990–2010: a systematic analysis," *The Lancet Global Health*, vol. 1, no. 6, pp. 339–349, 2013, doi: 10.1016/S2214-109X(13)70113-X.
[2]   S. Dua, U. R. Acharya, P. Chowriappa, and S. V. Sree, "Wavelet-based energy features for glaucomatous image classification," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 1, pp. 80–87, 2012, doi: 10.1109/TITB.2011.2176540.
[3]   D. Y. Park, E. J. Lee, J. C. Han, and C. Kee, "Applicability of ISNT rule using BMO-MRW to differentiate between healthy and glaucomatous eyes," *Journal of Glaucoma*, vol. 27, no. 7, pp. 610–616, 2018, doi: 10.1097/IJG.0000000000000970.
[4]   U. Raghavendra, A. Gudigar, S. V. Bhandary, T. N. Rao, E. J. Ciaccio, and U. R. Acharya, "A two layer sparse autoencoder for glaucoma identification with fundus images," *Journal of Medical Systems*, vol. 43, no. 9, pp. 1–9, 2019, doi: 10.1007/s10916-019-1427-x.
[5]   Y. He, Y. Zheng, Y. Zhao, Y. Ren, J. Lian, and J. Gee, "Retinal image denoising via bilateral filter with a spatial kernel of optimally oriented line spread function," *Computational and Mathematical Methods in Medicine*, vol. 2017, pp. 1–13, 2017, doi: 10.1155/2017/1769834.
[6]   A. A. G. Elseid, M. E. Elmanna, and A. O. Hamza, "Evaluation of spatial filtering techniques in retinal fundus images," *American J. of Artificial Intell.*, vol. 2, no. 2, pp. 16–21, 2018, doi: 10.11648/j.ajai.20180202.11.
[7]   O. M. A. Hazaimeh, K. M. O. Nahar, B. A. Naami, and N. Gharaibeh, "An effective image processing method for detection of diabetic retinopathy diseases from retinal fundus images," *International Journal of Signal and Imaging Systems Engineering*, vol. 11, no. 4, pp. 206–216, 2018, doi: 10.1504/IJSISE.2018.10015063.
[8]   Erwin, R. Zulfahmi, D. S. Noviyanti, G. R. Utami, A. N. Harison, and P. S. Agung, "Improved image quality retinal fundus with contrast limited adaptive histogram equalization and filter variation," in *2019 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, 2019, pp. 49–54, doi: 10.1109/ICIMCIS48181.2019.8985198.
[9]   G. Palanisamy, N. B. Shankar, P. Ponnusamy, and V. P. Gopi, "A hybrid feature preservation technique based on luminosity and edge based contrast enhancement in color fundus images," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 2, pp. 752–763, 2020, doi: 10.1016/j.bbe.2020.02.006.
[10]   M. Juneja, N. Thakur, S. Thakur, A. Uniyal, A. Wani, and P. Jindal, "GC-NET for classification of glaucoma in the retinal fundus image," *Machine Vision and Applications*, vol. 31, no. 5, pp. 1–18, 2020, doi: 10.1007/s00138-020-01091-4.
[11]   Y. Hu, J. Ren, J. Yang, R. Bai, and J. Liu, "Noise reduction by adaptive-SIN filtering for retinal OCT images," *Scientific Reports*, vol. 11, no. 1, pp. 1–14, 2021, doi: 10.1038/s41598-021-98832-w.
[12]   X. Chen, Y. Xu, D. W. Kee Wong, T. Y. Wong, and J. Liu, "Glaucoma detection based on deep convolutional neural network," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 715–718, doi: 10.1109/EMBC.2015.7318462.
[13]   Y. Chai, L. He, Q. Mei, H. Liu, and L. Xu, "Deep learning through two-branch convolutional neuron network for glaucoma diagnosis," in *Smart Health, Cham: Springer*, 2017, pp. 191–201, doi: 10.1007/978-3-319-67964-8_19.
[14]   Y. Chai, H. Liu, and J. Xu, "Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models," *Knowledge-Based Systems*, vol. 161, pp. 147–156, 2018, doi: 10.1016/j.knosys.2018.07.043.
[15]   A. D. -Pinto, S. Morales, V. Naranjo, T. Köhler, J. M. Mossi, and A. Navea, "CNNs for automatic glaucoma assessment using fundus images: an extensive validation," *BioMedical Engineering OnLine*, vol. 18, no. 1, pp. 1–19, 2019, doi: 10.1186/s12938-019-0649-y.
[16]   M. Juneja *et al.*, "Automated detection of glaucoma using deep learning convolution network (G-net)," *Multimedia Tools and Applications*, vol. 79, no. 21–22, pp. 15531–15553, 2020, doi: 10.1007/s11042-019-7460-4.
[17]   D. S. David, "Enhanced glaucoma detection using ensemble based CNN and spatially based ellipse fitting curve model," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–12, 2021, doi: 10.1007/s12652-021-03467-4.
[18]   R. Shinde, "Glaucoma detection in retinal fundus images using U-Net and supervised machine learning algorithms," *Intelligence-Based Medicine*, vol. 5, pp. 1–15, 2021, doi: 10.1016/j.ibmed.2021.100038.
[19]   L. Abdel-Hamid, "TWEEC: computer-aided glaucoma diagnosis from retinal images using deep learning techniques," *International Journal of Imaging Systems and Technology*, vol. 32, no. 1, pp. 387–401, 2022, doi: 10.1002/ima.22621.
[20]   A. E. Ilesanmi and T. O. Ilesanmi, "Methods for image denoising using convolutional neural network: a review," *Complex &*

Intelligent Systems, vol. 7, no. 5, pp. 2179–2198, 2021, doi: 10.1007/s40747-021-00428-4.

[21] V. A. Pimpalkhute, R. Page, A. Kothari, K. M. Bhurchandi, and V. M. Kamble, "Digital image noise estimation using DWT coefficients," *IEEE Transactions on Image Processing*, vol. 30, pp. 1962–1972, 2021, doi: 10.1109/TIP.2021.3049961.

[22] N. K. Santosh and S. S. Barpanda, "Wavelet and PCA-based glaucoma classification through novel methodological enhanced retinal images," *Machine Vision and Applications*, vol. 33, no. 1, pp. 1–42, 2022, doi: 10.1007/s00138-021-01263-w.

[23] X. Zhang, Z. Li, C. C. Loy, and D. Lin, "PolyNet: a pursuit of structural diversity in very deep networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 3900–3908, doi: 10.1109/CVPR.2017.415.

[24] S. Targ, D. Almeida, and K. Lyman, "Resnet in resnet: generalizing residual architectures," *Arxiv-Computer Science*, pp. 1–7, 2016, doi: 10.48550/arXiv.1603.08029.

[25] M. Biswas and H. Om, "A new adaptive image denoising method based on neighboring coefficients," *Journal of The Institution of Engineers (India): Series B*, vol. 97, no. 1, pp. 11–19, 2016, doi: 10.1007/s40031-014-0166-0.

[26] M. Ponomarenko, N. Gapon, V. Voronin, and K. Egiazarian, "Blind estimation of white Gaussian noise variance in highly textured images," *International Symposium on Electronic Imaging Science and Technology*, vol. 13, pp. 1–6, 2018, doi: 10.2352/ISSN.2470-1173.2018.13.IPAS-382.

[27] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: residual learning of deep CNN for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017, doi: 10.1109/TIP.2017.2662206.

[28] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2862–2869, doi: 10.1109/CVPR.2014.366.

[29] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng, "Patch group based nonlocal self-similarity prior learning for image denoising," in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 244–252, doi: 10.1109/ICCV.2015.36.

[30] S. Serte and A. Serener, "A generalized deep learning model for glaucoma detection," in *2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 2019, pp. 1–5, doi: 10.1109/ISMSIT.2019.8932753.

[31] P. Elangovan and M. K. Nath, "Glaucoma assessment from color fundus images using convolutional neural network," *International Journal of Imaging Systems and Technology*, vol. 31, no. 2, pp. 955–971, 2021, doi: 10.1002/ima.22494.

[32] K. Balasubramanian, K. Ramya, and K. G. Devi, "Improved swarm optimization of deep features for glaucoma classification using SEGSO and VGGNet," *Biomedical Signal Processing and Control*, vol. 77, p. 103845, 2022, doi: 10.1016/j.bspc.2022.103845.

[33] F. Guo *et al.*, "Yanbao: a mobile app using the measurement of clinical parameters for glaucoma screening," *IEEE Access*, vol. 6, pp. 77414–77428, 2018, doi: 10.1109/ACCESS.2018.2882946.

[34] S. Ajitha, J. Akkara, and M. Judy, "Identification of glaucoma from fundus images using deep learning techniques," *Indian Journal of Ophthalmology*, vol. 69, no. 10, pp. 2702–2709, 2021, doi: 10.4103/ijo.IJO_92_21.

[35] F. J. F. Batista *et al.*, "Rim-one DL: a unified retinal image database for assessing glaucoma using deep learning," *Image Analysis & Stereology*, vol. 39, no. 3, pp. 161–167, 2020, doi: 10.5566/ias.2346.

[36] B. Al-Bander, B. Williams, W. Al-Nuaimy, M. Al-Taee, H. Pratt, and Y. Zheng, "Dense fully convolutional segmentation of the optic disc and cup in colour fundus for glaucoma diagnosis," *Symmetry*, vol. 10, no. 4, pp. 1–16, 2018, doi: 10.3390/sym10040087.

[37] S. Maheshwari, V. Kanhangad, and R. B. Pachori, "CNN-based approach for glaucoma diagnosis using transfer learning and LBP-based data augmentation," *Arxiv-Electrical Engineering and Systems Science*, pp. 1–7, 2020, doi: 10.48550/arXiv.2002.08013.

[38] M. N. Bajwa *et al.*, "Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, pp. 1–16, 2019, doi: 10.1186/s12911-019-0842-8.

[39] L. K. Singh, M. Khanna, and Pooja, "A novel multimodality based dual fusion integrated approach for efficient and early prediction of glaucoma," *Biomedical Signal Processing and Control*, vol. 73, p. 103468, 2022, doi: 10.1016/j.bspc.2021.103468.

BIOGRAPHIES OF AUTHORS

Krishna Santosh Naidana ⬤ 🔗 SC 🔾 received his B.Tech degree in computer science and engineering from the Jawaharlal Nehru Technological University, Hyderabad, and M.Tech degree in software engineering from the GITAM University. He is currently doing Ph.D. in VIT-AP University, Amaravati, Andhra Pradesh, India. His research area is medical image processing. He can be contacted at email: krishna.18phd7018@vitap.ac.in.

Dr. Soubhagya Sankar Barpanda ⬤ 🔗 SC 🔾 received his B.Tech degree in computer science and engineering from the Biju Patnaik University of Technology, Rourkela, M.Tech degree in software engineering from the NIT, Rourkela, and Ph.D. degree from the NIT, Rourkela. He is currently working as associate professor in school of computer science and engineering, VIT-AP University, Amaravati, Andhra Pradesh, India. His research area is image processing, software engineering, and biometrics. He can be contacted at email: soubhagya.barpanda@vitap.ac.in.