Indian Journal of Radio & Space Physics
Vol. 50, March 2021, pp. 33-39

Detection of glaucoma from fundus image using pre-trained Densenet201 model

Poonguzhali Elangovan*, Vijayalakshmi D, & Malaya Kumar Nath

Department of Electronics and Communication Engineering, National Institute of Technology Puducherry, Karaikal 609 609, India

Received:12 February 2021; Accepted:5 March 2021

In recent years, the performance of deep learning algorithms for image recognition has improved tremendously. The inherent ability of a convolutional neural network has made the task of classifying glaucoma and normal fundus images more appropriately. Transferring the weights from the pre-trained model resulted in faster and easier training than training the network from scratch. In this paper, a dense convolutional neural network (Densenet201) has been utilized to extract the relevant features for classification. Training with 80% of the images and testing with 20% of the images has been performed. The performance metrics obtained by various classifiers such as softmax, support vector machine (SVM), k-nearest neighbor (KNN), and Naive Bayes (NB) have been compared. Experimental results have shown that the softmax classifier outperformed the other classifiers with 96.48% accuracy, 98.88% sensitivity, 92.1% specificity, 95.82% precision, and 97.28% F1-score, with DRISHTI-GS1 database. An increase in the classification accuracy of about 1% has been achieved with enhanced fundus images.

Keywords: Deep learning, Fine-tuning, Ocular disease, Transfer learning

1 Introduction

Glaucoma indicates a cluster of eye disorders that impair the vision of humans. Peripheral vision loss has been caused by the weakening of optic nerves because of increased eye pressure[1]. Early screening and detection are essential, as this disease hasno symptoms. Glaucoma has been characterized by increased cup size[2]. Intraocular pressure measurement, visual field test, and optic nerve head examination are the standard clinical procedures followed by the clinicians for identifying the presence of glaucoma. Such examinations, however, suffer from intra/inter observer errors. Computer-aided diagnosis (CAD) approaches[3-5] may resolve the limitations of conventional clinical examinations. While CAD-based algorithms yield better results, the key downside is that these techniques depend primarily on handcrafted features. Deep learning algorithms, on the other hand, perform the classification by extracting the features directly from the images.

Transfer learning is a typical deep learning approach in which the weights and layers from pre-trained models can be transferred to a new classification problem. This results in faster and easier training. Different pre-trained models with varying depth have been developed to enhance the performance of image classification and image recognition tasks. Densenet201, one among the standard pre-trained model which has been effectively used for image classification due to its deep architecture. By introducing the dense connection between each layer, the performance of the model has been improved.

Figure 1 represent glaucoma infected and normal fundus images from DRISHTI-GS1 database[6]. The distinct features may not be noticeable in some fundus images. By improving the contrast of those images the performance of the network may get improved. From the literature, it has been observed that better results are obtained by contrast enhancement based on two-dimensional histogram[7]. While enhancing the images using two-dimensional histogram methods, the contextual information around each pixel has been considered. This overcomes the drawbacks of using one-dimensional histograms significantly[8].

The main objective is to enhance the performance of classification system by incorporating the suitable pre-processing technique and also to compare the performance of Densenet201 model with various classifiers towards the task of classifying glaucoma and normal fundus images.

The major contributions of this work are: (i) improving the quality of fundus images using Residual spatial entropy-based contrast enhancement (RESE algorithm) (ii) accurate selection of different hyperparameters to boost the performance of the

* Corresponding author (E-mail: epoonguzhali3@gmail.com)

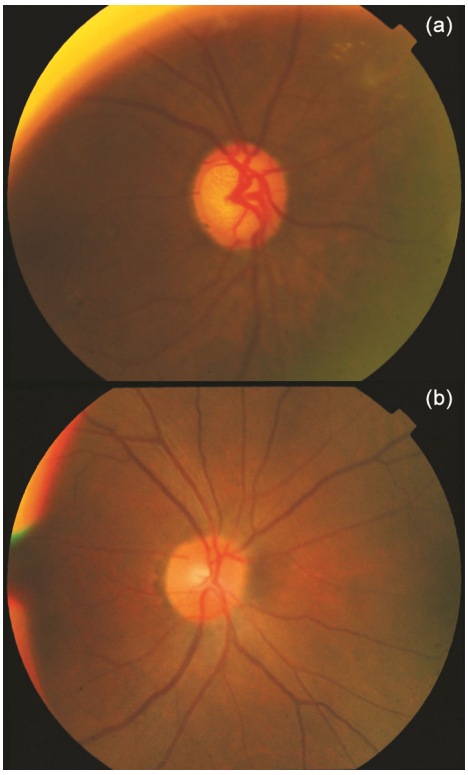

Fig. 1 — Fundus images[6] (a) glaucoma, and (b) normal.

classification system (iii) fine-tuning the Densenet 201 model for better classification (iv) comparing the performance of various classifiers.

With the invention of different pre-trained models, the performance of image classification problem has been greatly improved. As pre-trained networks are trained with a large number of images, transferring the weights and layers to new classification problem will make the training easier and faster. Several works for the automatic detection of glaucoma from fundus images using CNN have been published in the literature. In some current works[9-13], classification is done using standard CNN model. Transfer learning based glaucoma diagnosis is employed in[14-15].

Chen *et al.*[9] have classified glaucoma and normal images using a sixteen-layer CNN model. The network weights are learned using the SGDM optimizer. Authors have reported AUC values of 0.831 and 0.887 for ORIGA and Singapore Chinese Eye Study (SCES) databases, respectively.

A twenty-layer CNN model for two class classification problem has been developed by Raghavendra *et al.*[10] The performance of CNN for various learning rates has been investigated. Authors have reached 98.13%, 98%, and 98.3%, respectively, for accuracy, sensitivity, and specificity with 1426 images.

Li *et al.*[12] have classified glaucoma and normal images using attention-based CNN (AG-CNN). For evaluation, a large-scale attention-based glaucoma (LAG) database comprising 5824 fundus images has been developed. In the LAG and RIMONE2 databases, the authors achieved an overall accuracy of 95.3% and 85.2%, respectively.

Juan *et al.*[16] have investigated the performance of various pre-trained models for glaucoma detection. Results have been compared with a standard sixteen layer CNN. Among the networks, VGG19 yields the highest AUC of 0.94, sensitivity of 87.01%, and specificity of 89.01% with 2313 images. Andres *et al.*[14] examined the performance of different pre-trained models. The authors have concluded that the Xception model has outperformed other models using 1707 images with overall accuracy, sensitivity, and specificity of 96.05%, 93.46%, and 85.80%, respectively.

Zhen *et al.*[15] have compared around eight pre-trained models in their work. Authorshave reported that the densenet model performs better with the highest classification accuracy of 75.50% using 5978 images. Orlando *et al.*[17] have exploited the transfer learning approach using overfeat and VGG-S for glaucoma classification. The contrast of the fundus images is enhanced using Contrast Limited Adaptive Histogram Equalization (CLAHE). With DRISHTI-GS1 database, an overall AUC of 0.7626 and 0.7180 are obtained using overfeat and VGG-S, respectively.

Private databases (images taken from hospitals) are used in most of the existing works. So it is difficult to determine the quality of those images. The nature and quality of fundus images will greatly improve the performance of the deep neural network. However, literature studies reveal that at the pre-processing stage in the deep learning-based image classification, minimal work has been performed in the field of image enhancement. This inspired us to create a framework that could improve the performance metrics by considering the enhanced images.

2 Materials and Methods

In recent years, the classification of glaucomatous and normal fundus images using a deep neural network had increased in popularity. Training the neural network from scratch was time-consuming. Also, it requires an effective hyperparameter selection technique. Instead, transferring the weights from the standard pre-trained network was easy and had better performance metrics for classification problems. The

proposed method for glaucoma detection using the Densenet201 network was shown in Fig. 2.

The block diagram comprises of three stages: pre-processing, feature extraction, and classification. Improving the quality of images, enlarging the number of images, and resizing were performed in pre-processing stage. In the feature extraction stage, the relevant features were extracted from the Densenet201 model using transfer learning approach. Finally, classification of glaucomatous and normal images were performed in classification stage.

2.1 Contrast enhancement

The contrast of the image was improved by the Residual entropy based contrast enhancement technique[8]. In this method, histograms were calculated based on the distributions concerning the spatial locations called two-dimensional histograms and used for enhancement.

The two dimensional histogram and the spatial entropy were obtained by using the following equations:

$$h_j = h_j(x,y), 1 \le x \le M, 1 \le y \le N \quad \ldots(1)$$

where, $h_j(x,y)$ denotes the two dimensional histogram of the j^{th} intensity value and x,y represents the spatial grid values on the image. M and N represents the total number of grids in the input image. Spatial entropy (S_j) was calculated from the h_j.The mapping fuction was obtained from the S_j

$$S_j = -\sum_{x=1}^{M}\sum_{y=1}^{N} h_j(x,y)\log_2\left(h_j(x,y)\right) \quad \ldots(2)$$

Discrete function f_j was derived from S_j, and it was represented as

$$f_j = \frac{S_j}{\sum_{l=1,l\neq j}^{k} S_l} \quad \ldots(3)$$

The joint spatial entropy ($S_{j,m}$) was obtained from joint histogram ($h_{j,m}$) by the following equation

$$h_{j,m}(x,y) = \max\left(h_j(x,y)h_m(x,y)\right) \quad \ldots(4)$$

$$S_{j,m} = -\sum_{x=1}^{M}\sum_{y=1}^{N} h_{j,m}(x,y)\log_2 h_{j,m}(x,y) \quad \ldots(5)$$

Residual entropy (R_j) was obtained from by using $S_{j,m}$

$$R_j = \frac{w_j s_j - w_{j,q}\sum_{m=1,l\neq k}^{K} S_{j,m}}{w_j + w_{j,q}} \quad \ldots(6)$$

w_i and $w_{j,q}$ are two dimensional histograms weighting functions and joint histograms respectively, defined as

$$w_j = \sum_{x=1}^{M}\sum_{x=1}^{N} h_j(x,y) \quad \ldots(7)$$

$$w_{j,q} = \sum_{l=1,l\neq j}^{K}\sum_{x=1}^{M}\sum_{y=1}^{N} h_{j,m}(x,y) \quad \ldots(8)$$

The discrete function f_j and a CDF F_j was obtained by

$$f_j = \frac{R_j}{\sum_{l=1,l\neq i}^{k} R_l} \quad \ldots(9)$$

The mapping function $y_j$was generated by F_j

$$y_j = \left\lfloor F_j \times (y_u - y_d) + y_d \right\rfloor \quad \ldots(10)$$

y_u represents the minimum intensity and y_d denotes the maximum intensity of the dynamic scale.

2.2 Data augmentation

Data augmentation was essential, as a deep neural network with a large number of images works better. The fundus images were enlarged using rotation augmentation technique. As DRISHTI-GS1 database had unbalanced (89 glaucoma images and 12 normal images), the rotation steps may get varied. Each glaucoma image was randomly rotated between zero and nineteen degrees in a clockwise direction. Rotating each normal image between zero and ninety-nine degrees in the clockwise direction would generate 1200 images. The fundus images were

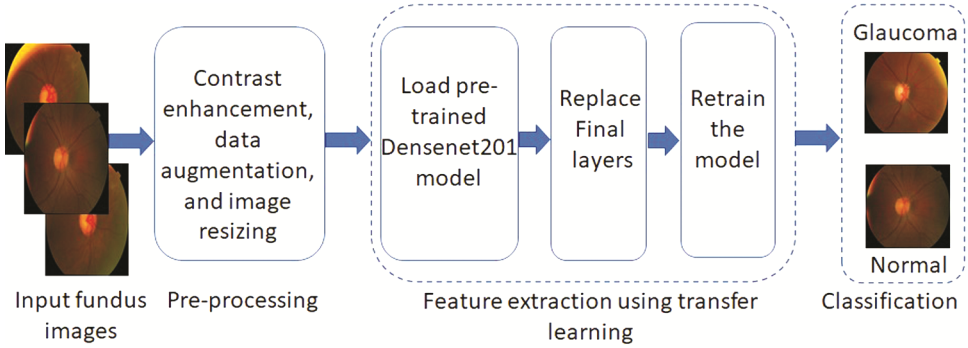

Fig. 2 — Glaucoma detection from fundus images using pre-trained Densenet201 model[18].

resized to 224×224, as this standard input image size was required by the Densenet201 model.

2.3 Pre-trained model

For image classification, Densenet201[18] was most commonly used, as it had many advantages such as alleviating the issue of the vanishing gradient, enhancing the propagation of features, facilitating reuse of features, and significantly reducing the number of parameters. Densenet201 includes four dense blocks and three transition blocks. Each dense block had a varying number of dense layers (6, 12, 48, 32 in blocks 1, 2, 3, 4, respectively) which were densely connected. Through feature reuse, Densenet201 realizes the objective of easy training and parameter performance, which incorporates the concatenated feature maps created by all previous layers into the subsequent layer. The kth layer receives the feature maps of all proceeding layers. The output feature map at the kth layer was given as

$$x_k = H_k([x_0, x_1, \ldots x_k - 1]), \quad \ldots(11)$$

where, $[x_0, x_1, \ldots x_k - 1]$ refers to the concatenation of feature maps produced in the appropriate layers. The composite function H_k includes batch normalization[19], rectified linear unit (ReLU)[20], and convolution. By normalizing the activations using Equation 12, batch normalization layer achieves stable distribution of activations throughout the network.

$$X_i = \frac{x_i - \mu_b}{\sqrt{\sigma_b^2 + \epsilon}}, \quad \ldots(12)$$

where, mini batch mean (μ_b) and variance (σ_b^2) are defined as

$$\mu_b = \frac{1}{m}\sum_{i=1}^{m} x_i \quad \ldots(13)$$

$$\sigma_b^2 = \frac{1}{m}\sum_{i=1}^{m} (x - \mu_b)^2 \quad \ldots(14)$$

Further mean and standard deviation were parameterized as trainable parameters using

$$y_i = \gamma X_i + \beta. \quad \ldots(15)$$

To improve the training speed, ReLU activation function defined as $f(x) = max(0, x)$ was employed. The convolution layer extracts the features by convolving the input representation with kernel given by

$$C(i,j) = \sum_{u=-m}^{m}\sum_{v=-m}^{m} I(i-u, j-v).K(u,v) + b, \quad \ldots(16)$$

where, b represents the bias term. Bottleneck layers were introduced by utilizing a 1×1 convolution before each 3×3 convolution. This reduces the computational complexity in dense layers. Hyperparameters such as kernel size, stride, and padding are chosen so that feature map dimensions remains constant within the block. Dimensionality reduction was achieved by a transition layer, which includes batch normalization, 1×1 convolution, and 2×2 average pooling layers.

Fine tuning was performed so that the network weights and initial layers were transferred. Final layers were modified according to the number of classes. The learnable parameters were optimized using Adam optimizer[21] to reduce the cross entropy loss defined by

$$loss = -\sum_{m=1}^{X}\sum_{n=1}^{Y} A_{mn} ln(P_{mn}) \quad \ldots(17)$$

where, X and Y denotes number of samples and number of classes, respectively. A_{mn} indicates the actual output and P_{mn} gives the predicted output. The parameters were updated using

$$W_{i+1} = W_i - \frac{\alpha}{\sqrt{V_i} + \zeta} M_i, \quad \ldots(18)$$

The bias-corrected mean M_i and bias-corrected variance V_i were described by

$$M_i = \frac{m_i}{1 - \beta_1^i}, \quad \ldots(19)$$

$$V_i = \frac{v_i}{1 - \beta_2^i}, \quad \ldots(20)$$

3 Results and Discussion

This work takes into account the publicly accessible DRISHTI-GS1 database to evaluate the performance of the network. Different metrics used in this work were described in Table 1.

Table 1 — Performance metrics[13]

Metrics	Mathematical Expression
Accuracy	$ACC = \frac{TP + TN}{TP + TN + FP + FN}$
Sensitivity	$SN = \frac{TP}{TP + FN}$
Specificity	$SP = \frac{TN}{TN + FP}$
Precision	$PR = \frac{TP}{TP + FP}$
F1-score	$F1 - score = \frac{2*Recall*Precision}{Recall+Precision}$

Table 2 — Metrics obtained using softmax classifier[18]

Performance metrics / Data folds	ACC	SN	SP	PR	F1-score
Fold 1	97.3	95.8	100	100	97.85
Fold 2	97.9	100	94	96.8	98.37
Fold 3	94.3	98.6	86.5	92.9	95.66
Fold 4	96.8	100	91	95.2	97.54
Fold 5	96.1	100	89	94.2	97.01
Average	96.48	98.88	92.1	95.82	97.28

Table 3 — Metrics obtained using SVM classifier[18]

Performance metrics / Data folds	ACC	SN	SP	PR	F1-score
Fold 1	97.1	95.6	99.5	99.7	97.60
Fold 2	98.4	100	95.5	97.6	98.78
Fold 3	93.6	96.7	88	93.5	95.07
Fold 4	95.5	100	87.5	93.5	96.64
Fold 5	95.5	100	87.5	93.5	96.64
Average	96.34	98.34	92.7	96.1	97.17

Hyperparameters such as batch size, learning rate, optimizer, and maximum epochs can decide the overall performance of the model. Using Adam optimizer, the network weights were optimized as it had many benefits compared to other optimizers[22]. The initial learning rate and batch size had a major influence on network performance. These two hyperparameters were selected using two-stage tuning approach[13]. Based on the results, a batch size of 32 and an initial learning rate of 0.00001 yield better glaucomatous and normal image classification. A maximum epoch of 20 was used in this work.

The Densenet201 model was trained with 1420 glaucoma and 1000 normal images. With 360 glaucoma and 200 normal images, the model testing was completed. The simulation experiment with various folds of enhanced images was replicated. 80% of training images and 20% of testing images were included in each fold. Five different folds of images were used in this work.

Using softmax classifier, the better classification accuracy of 96.48%, sensitivity of 98.88%, specificity of 92.1%, precision of 95.82%, and F1-score of 97.28% were obtained. The results were indicated in Table 2. For the classification of glaucomatous and normal images using SVM, KNN, and NB classifiers, distinct features were extracted from the global average pooling layer. It was evident from Table 3 that the metrics obtained by the SVM classifier were approximately the same as the Softmax classifier. Tables 4 and 5 indicates the results obtained using KNN and NB classifiers. Average accuracy of 96.34%, 93.88%, and 95.56% were obtained using SVM, KNN, and NB, respectively.

Table 4 — Metrics obtained using KNN classifier[18]

Performance metrics / Data folds	ACC	SN	SP	PR	F1-score
Fold 1	95.4	93.1	99.5	99.7	96.28
Fold 2	96.1	100	89	94.2	97.01
Fold 3	90.7	94.7	83.5	91.2	92.91
Fold 4	94.5	98.3	87.5	93.4	95.78
Fold 5	92.7	98.9	81.5	90.6	94.56
Average	93.88	97	88.2	93.82	95.30

Table 5 — Metrics obtained using NB classifier[18]

Performance metrics / data folds	ACC	SN	SP	PR	F1-score
Fold 1	97.9	100	94	96.8	98.37
Fold 2	99.6	100	99	99.4	99.69
Fold 3	93.2	95.3	89.5	94.2	94.74
Fold 4	92.1	92.5	91.5	95.1	93.78
Fold 5	95	95.4	92.5	95.9	95.64
Average	95.56	96.64	93.3	96.28	96.44

Table 6 — Metrics obtained using networks[9,11,18]

Performance metrics / network	ACC	SN	SP	PR	F1-score
Chen *et al*.	93.8	98.74	84.9	92.18	94.64
Bajwa *et al*.	93.22	97.26	86	92.68	95.34
Densenet201	96.48	98.88	92.1	95.82	97.28

In addition, Densenet201 model was also trained with different folds of original images for comparison. Figure 3 illustrates the obtained results for original and enhanced images. With original images, the highest average sensitivity was achieved with softmax, SVM, KNN, and NB classifiers of 99.5%, 99.56%, 97.62%, and 99.44% , respectively. By taking into account the enhanced images, the highest average classification accuracy was obtained. It was also evident from the results that, except KNN classifier, the accuracy of classification increased by about 1% when considering enhanced images.

Further the results obtained by Chen *et al*.[9] and Bajwa *et al*.[11] were compared with the proposed Densenet201 network. Table 6 describes the metrics obtained by various models. Densenet201 model classifies better compared to other two models.

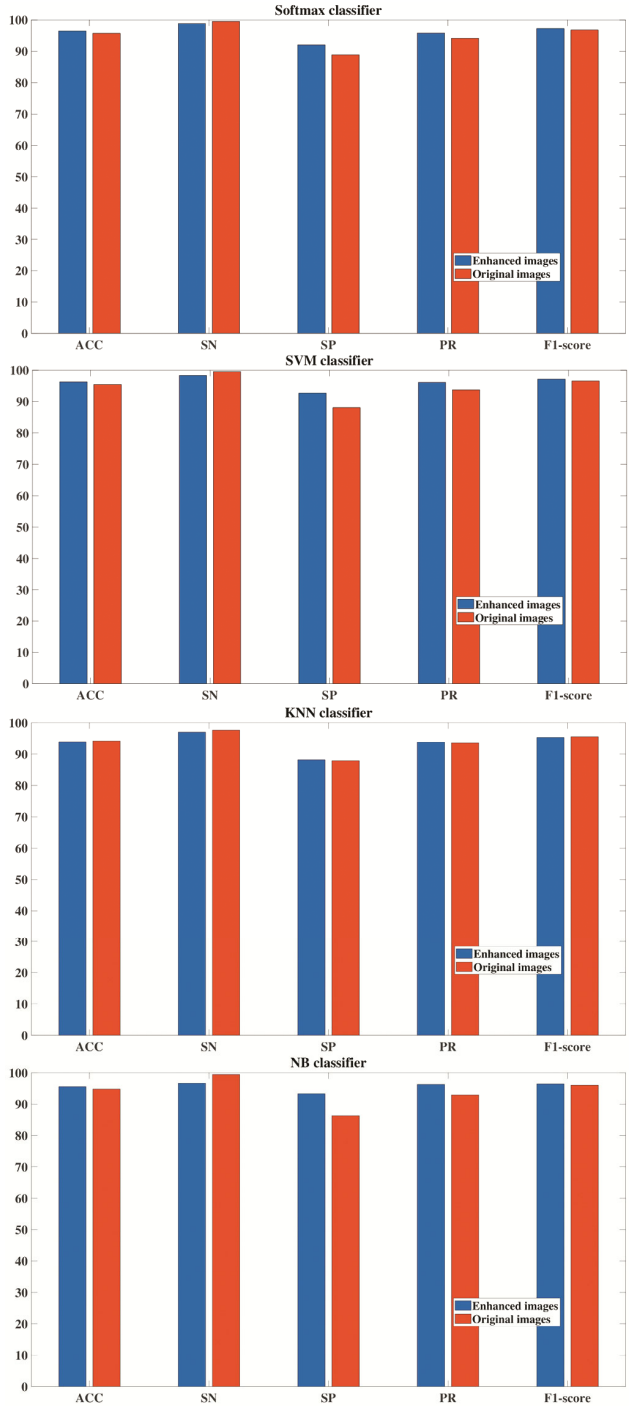

Fig. 3 — Comparison of metrics with enhanced[8] and original fundus images[6] (a) softmax classifier, (b) SVM classifier, (c) KNN classifier, and (d) NB classifier.

4 Conclusion

In this paper, the transfer learning approach to glaucoma detection from fundus images has been discussed. Residual Spatial Entropy-based Image Contrast Enhancement (RESE) technique has been employed to enhance the quality of fundus images taken from the DRISHTI-GS1 database. Densenet201 model has been fine-tuned effectively to extract the relevant features for classification. Using different metrics such as accuracy, sensitivity, specificity, precision, and F1-score, the model's performance has been evaluated. With a suitable selection of hyperparameters, better results have been achieved by the softmax classifier. Also, the classification accuracy has been improved by 1% using enhanced images.

References

1 Mardeen J, *Nurs Times*, 110 (2014) 20.
2 Poonguzhali E, Ravichandran G, Nath M K, & Acharya O P, *2018 International Conference on Applied Electromagnetics, Signal Processing and Communication (IEEE, Bhubaneswar*), ISBN: 978-1-5386-8334-7, 2018, p.1.
3 Zilly J, Buhmann J M, & Mahapatra D, *Comput Med Imaging Graph*, 55 (2017) 28.
4 Nath M K, & Dandapat S, *Int J Imaging Syst Technol*, 22 (2017) 161.
5 Poonguzhali E, Nath M K, & Mishra M, *Procedia Comput Sci*, 171 (2020) 2675.
6 Sivaswamy J, Krishnadas S R, Joshi J D, Jain M, & Tabish A U S, 2104 *IEEE 11th International Symposium on Biomedical Imaging (IEEE, Beijing*), ISSN: 1945-7928, 2014, p. 53.
7 Vijayalakshmi D, Nath M K, & Acharya O P, *Sens Imaging*, 21 (2020) 1.
8 Celik T, & Li H C, *J Mod Opt*, 63 (2016) 1600.
9 Chen X, Xu Y, Wong D, Wong T Y, & Liu J, *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE, Milan*), ISBN: 978-1-4244-9271-8, 2015, p. 715.
10 Raghavendra U, Fujita H, Bhandary S V, Gudigar A, Tan J H, & Acharya U R, *Inf Sci*, 144 (2018) 41.
11 Bajwa M N, Malik M I, Siddiqui S A, Dengel A, Shafait F, Neumeier W, & Ahmed S, *BMC Medical Inform Decis Mak*, 19 (2019) 1472.
12 Li L, Xu M, Liu H, Li Y, Wang X, Jiang L, Wang Z, Fan X, & Wang N, *IEEE Trans Med Imaging*, 39 (2020) 413.
13 Poonguzhali E, & Nath M K, *Int J Imaging Syst Technol*, 31 (2021) 955. 14.
14 Pinto A D, Morales S, Naranjo V, Kohler T, Mossi J M, & Navea A, *BioMedEng Online*, 18 (2019) 1.
15 Zhen Y, Wang L, Liu H, Zhang J, & Pu J, *Med Image Anal*, 9 (2018) 297.
16 Valverde J G, Anton A, Fatti G, Liefers B, Herranz A, Santos A, Sanchez C, & Carbayo M L, *Br J Ophthalmol*, 10 (2019) 892.
17 Orlando J I, Prokofyeva E, Fresno M D, & Blaschko M B, 12th International Symposium on Medical Information Processing and Analysis (SPIE, Tandil), ISBN: 978-1-5106-0778-1, 2017, p. 241.
18 Huang G, Liu Z, Maaten L V D, & Weinberger K Q, 2017 IEEE Conference on Computer Vision and Pattern

Recognition (IEEE, Honolulu), ISBN: 978-1-5386-0457-1, 2017, p. 2261.
19 Ioffe S, & Szegedy C, 32nd International Conference on Machine Learning (*Journal of Machine Learning Research, Lily France*), ISSN: 1938-7228, 2015, p. 448.
20 Ide H, & Kurita T, 2017 International Joint Conference on Neural Networks (IEEE, Anchorage), ISBN: 978-1-5090-6182-2, 2017, p. 2684.
21 Kingma D P, & Ba J, *CoRR*, 1 (2015) 1.
22 Ruder S, *CoRR*, 1 (2016) 1.