Diabetic Retinopathy Grading with Deep Visual Attention Network

https://doi.org/10.3991/ijoe.v18i09.30075

S. Geetha[1], Mansi Parashar[1], JS Abhishek[1], Raj Vishal Turaga[1],
Isah A. Lawal[2], Seifedine Kadry[2](✉)
[1]School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India
[2]Department of Applied Data Science, Noroff University College, Kristiansand, Norway
skadry@gmail.com

Abstract—Diabetic Retinopathy is a serious complication arising in diabetes afflicted patients. Its effective treatment depends on early detection, and the course of action varies decisively with the intensity of the affliction. Computer-aided diagnosis helps to detect not only the presence or absence of the disease, but also the severity, making it easier for ophthalmologists to construct a treatment plan. Diabetic retinopathy grading is the task of classifying images of the eye's fundus of diabetic patients into 5 different grades ranging from 0–4 based on the severity of the disease. In this work, we propose a deep neural network architecture to address the grading problem. The method utilizes additional attention layer in the neural network model to capture the spatial relationship between the region of interests in the images during the training process to better discriminate between the different severity stage of the disease. Also, we analyze the impact of different image processing techniques on the classification results. We assessed the performance of our proposed method using a dataset of eye fundus images and obtained classification accuracy of 89.20% on average. This performance surpass that reported for other state-of-the-art methods on the same dataset. The effectiveness of the proposed method will facilitate the procedural workflow of identifying severe cases of diabetic retinopathy.

Keywords—diabetic retinopathy, diabetic retinopathy grading, deep learning, attention net, CLAHE, Gaussian Blur

1 Introduction

Diabetes is a disease that causes the accumulation of glucose in the blood. In a healthy person, glucose in the blood is transported by insulin to various cells to be stored or used for energy. A diabetic person is unable to produce sufficient insulin which adversely affects the body. When a person with diabetes is left untreated for an extended period, multiple organs such as the eye, nerves, kidneys get affected. In 2019, the number of adults aged between 18 and 79 years diagnosed with diabetes globally was 463 million [ref]. The estimated value of global diabetic patients in 2030 is 578 million (10.2% increase) and in 2045 is 700 million (10.9% increase) [1]. Diabetic

Retinopathy (DR) is one of the significant diabetic complications that affect the patient eyes over a long time if left unchecked [3]. The high sugar level in the human body blocks the blood vessels entering the retina, resulting in the swelling of the blood vessels due to accumulation of blood and the eventual busting of the vessel around the retina. Moreover, the eyes accumulate fluid during long periods of high blood sugar. This accumulation of fluid changes the shape and curve of the lens, causing vision changes and if the patient does not go for regular check-ups, DR's complications can lead to permanent vision loss. According to the International Diabetes Federation, the global prevalence of Diabetic Retinopathy from 2015 to 2019 was 27% [ref]. Almost 80% of diabetic patients are reported to develop symptoms of DR eventually [ref]. According to a survey conducted by the Economic Times in 2017, 63% of Indians were not aware that diabetes severely impacts the eye. A study by Stela et al. [2], reported that an increase in the prevalence of type-two diabetes, increases the DR-affected visual impairment cases. Also, one out of every three diabetic patients is reported to have some degree of DR, and one out of every ten diabetic patients is said to have a high chance of developing vision-threatening DR [4].

DR can develop in anyone who has diabetes for a long time. In severe cases of DR, the eye may grow new blood vessels to recover the damage done. These new blood vessels are weaker and will bleed and leak easily. This type of DR is called Proliferative Diabetic Retinopathy (PDR) [4]. The early stage of DR on the other hand, is known as Non-Proliferative Diabetic Retinopathy (NPDR) [21]. There is no cure for DR, but treatments are available to slow the process of vision loss. Thus, the best way to prevent it is to make sure that a person gets DR to check up regularly and maintain a fitness regime. The traditional method for checking the development of DR is cumbersome and time-consuming. Firstly, the doctor dilates the eye of the patient, followed by either Fluorescein Angiography (FA) or Optical Coherence Test (OCT). The FA involves injecting the patient with a liquid dye called fluorescein via the vein. The injected dye will then flow through the blood vessels in the eye and highlight them. If the dye leaks into the retina or stains the blood vessel, it indicates some abnormality in the blood vessel. The OCT on the other hand, is a non-invasive imaging scan that provides high-resolution cross-sectional images of the retina. It gives critical information like the retina's thickness and allows the doctors to look for swellings or leakage of fluids in the blood vessels around the retina. The fundus images are checked manually by an ophthalmologist. In the diagnosis for NPDR, the presence of lesions like microaneurysms, dot hemorrhages (HM), abnormalities in the blood vessels are checked in the fundus image. However, in the diagnosis of PDR, the presence of new blood vessels often referred to as microvascular abnormalities are checked in the fundus image. The manual checking of fundus images is prone to human errors especially in remote areas with limited number of trained personnel. Therefore, it is important to assist the doctors with an automate way of diagnosing the DR that is faster and efficient. The work in this paper addresses the problem of automating the diagnostic process for DR. We propose a deep neural network-based method that perform analysis of the fundus images and automatically predict the stage of the DR on the patient. This can help the doctors customize treatment for the patient depending on the predicted stage of the DR.

2 Related works

The research to find efficient and accurate ways to detect DR has been going on for decades now, and with the increasing number diabetic patients, the topic has become even more important for the research community. In this section, we review the state-of-the-art methods for the detection and classification of DR. The discussion is divided into two parts. In the first part, we review studies that focus on classical machine learning methods for DR detection, while in the second part, we review those that employed deep learning-based methods for DR detection and classification.

2.1 Classical machine learning methods

Shiva et al. [5] performed a comparative analysis of six different classical machine learning algorithms including Support Vector Machine (SVM), K Nearest Neighbors (kNN), Logistic Regression, Probabilistic Neural Network (PNN), Naive Bayes Classifiers (NBC), and Random Forests for detection of DR. The analysis indicated that SVM was the most efficient algorithm. Javeria et al. [6] discussed the impact of exudates on DR and methods to detect them. The study used Gabor filtering to process the images and extract representative features, and then applied classifiers such SVM, NBC and KNN to classify the feature set. Similarly, Ullah et al. [11] considered exudates to be a parameter to classify an image as DR or non-DR case. After the pre-processing of the images, the Otsu segmentation method was used for the segmentation of exudates. Wen et al. [7] presented a study that detect the presence of microaneurysms by classifying the fundus images into microaneurysms and non-microaneurysms. The study also reported SVM to be the effective classifier for this kind of application among the three classifiers used in the experiments. Revathi et al. [8] proposed an ensemble of classifiers to improve the accuracy of detecting the presence of DR. The study extract lesion features from the fundus images for developing the classifiers. Similarly, Rituparna et al. [12] studied the detection of many types of lesions that cause DR. The study employed Fuzzy CMeans clustering for the lesion extraction and classify the extracted lesions as dark or bright, respectively, using a combination of NB and SVM classifiers. Welikala et al. [9] presented a study for the detection and classification of neovascularization in PDR. The used genetic algorithms for feature extraction and SVM classifiers for the classification of the fundus images as containing a new vessels or non-new vessels. Somasundaram and Ali [10] proposed a two-stage detection method whereby features of the disease are extracted using t-distributed Stochastic Neighbor Embedding and then classified using ensemble of classifiers. The proposed method could detect if an image has PDR or NPDR. Emran et al. [14] compared two ensemble classifier models namely Fuzzy Random Forests (FRF) and Dominance-Based Rough Set Balanced Rule Ensemble (DBR-SBRE) for DR detection and reported that FRF gave better results for the datasets used.

One major limitation of the methods reviewed thus far is that they relied on the fact that the fundus images used in training process are labeled by an expert in advance [13, 19, 20]. To overcome this limitation, Yuego et al. [15] proposed a Self-supervised Fuzzy Clustering Network (SFCN) method that take as input unlabeled retinal images and classify them as having DR or not. Although the models trained are effective in detecting presence of the DR in the fundus images, however, they still do not solve the fundamental problem of specifying the severity of the disease. It is important to explore the means of also grading the severity of the disease to understand its stage in the patients' body. Recently researchers are looking at analyzing the different sets of lesions to extract information about the severity of the DR. Sohini et al. [16] proposed a three-stage fundus image screening method that combine segmentation of the lesion region, classification of the segmented region and analysis of the severity of the DR as seen in the image all in a single framework. Gule et al. [17] presented a model that performs a grade classification of fundus images into healthy, mild NPDR, and moderate NPDR. In the study, the authors detect the presence of exudates and microaneurysms using the closing morphological operation and canny edge detection, respectively, while the classification of the DR was done using the SVM classifier. Charu et al. [18] proposed a Hierarchical Severity Grading (HSG) model for detection and classification of DR. HSG is a three-step hierarchical method namely Retinal landmark segmentation, DR lesion segmentation, and DR severity grade classification. Abbas et al. [33] introduced a new semi-supervised hierarchical approach for the multi-grade classification of DR. Feature extraction was done using Dense Colour Scale-Invariant Transform and Gradient Location-Orientation Histogram. The classification was done using Deep Learning Neural Network architecture along with discriminative fine-tuning phase. Similarly, Jadhav et al. [35] proposed a new type of Rider Optimization Algorithm (ROA) called the Modified Gear and Steering-based ROA (MGS-ROA) that improved the feature extraction and selection processes in DR classification. The best features selected are then classified using a Deep Belief Network model.

2.2 Deep learning methods

The methods reviewed in the Section 2.1, required the generation of handcraft feature from the fundus images for classification. The extraction of representative features requires the skills of an expert in the domain area. Besides the accuracy of the developed model can degrade with poor feature sets. Deep learning (DL) methods allows the classifications of images without explicit need for feature extraction [23]. It is also capable of solving complex image classification problems. Hence, many DL models have been implemented for DR classification. Some studies combine both DL and classical machine learning for improved DR classification. Shanker et al. [21] applied the Synergic Deep Learning model to classify the fundus images according to their severity. In the study, the author performed histogram-based segmentation on the green channel of the fundus images before classifying them into four different severity levels. Eman et al. [22] proposed a method that combined Convolutional Neural Network (CNN) with SVM for DR classification. The fundus images were segmented using CNN and the segmented regions are presented as input to the SVM for classification. Similarly, Mansour et al. [24] proposed a system that uses two set of deep neural networks for feature extraction in DR classification. The extracted features were then classified using a polynomial kernel based SVM classifier.

While some researchers use a combination of classical machine learning and deep learning algorithm, others focus on designing the entire DP detection and classification pipeline using just deep neural network [25, 27, 28, 32, 37]. Zhan et al. [26] proposed a Coarse-to-Fine DR Network (CF-DRNet) which comprises of two sets of deep neural networks namely Coarse Network and Fine Network. The Coarse Network is used to detect the presence of DR while the Fine Network is used to predict the severity of the DR. Juan et al. [29] proposed a hierarchical multi-task deep learning framework for simultaneous diagnosis of DR severity and DR-related features in fundus images. A hierarchical structure is designed to incorporate the causal relationship between DR-related features and DR severity levels into the multi-task deep neural network.

Another approach for detection and classification of DR is using transfer learning whereby pre-trained neural network parameters are used to build a custom model. This technique has the potential of achieving good classification accuracy even with small sets of training fundus images. Wan et al. [30] used transfer learning and hyperparameter tuning to get the optimum weights set for identifying DR in fundus images. These weights were then used to train new sets of CNN models for identification and classification of DR. Similarly, Varun et al. [31] used an adapted pre-trained deep neural network to detect DR in fundus images. The trained model achieved a sensitivity of 97.5% and a specificity of 93.4% on the *EyePACS* dataset.

Shankar et al. [34], proposed a Deep Neural Network with Moth Search Optimization (DNN-MSO) for improving the multi-grade classification of the fundus images. Feature extraction was done using another custom made deep neural network. Along et al. [36] on the other hand, presented a Category Attention Block Network (CABNet) which allows learning class-specific features of the DR, thus helping the differentiation of the different severity levels. Kalyani et al. [38] presented a study that shows the improvement in classification accuracy for DR when the pooling layers in CNN were replaced with class capsule layers. A capsule network was developed in the study which consists of convolutional layers, primary capsule layer, and class capsule layer. When compared with a classical CNN model, the performance of the proposed model surpasses that of the CNN in classifying the severity level of the DR.

2.3 Research gaps and motivation

The existing research on fundus images focuses heavily on binary classification of input images into DR and non-DR diagnosis. DR severity grading, that is, classifying fundus images into 5 grades depending on the nature and stage of the disease increases the applicability of the system for computer aided diagnosis. This is because ophthalmologists are given greater insights into the problem for making an informed decision about the course of action to undertake.

DR severity grading techniques involved using multiple models for feature extraction and classification. These methods are cumbersome and sometimes incur high computational cost. Moreover, the methods that explore the use of deep learning in DR severity grading do not explore the role of image processing in improving the DR analysis. The spatial relationship between the regions of interest in fundus images have not been studied extensively, even when they can help in gaining new insights about fundus image grading. Additionally, deep learning-based methods with higher parameters yield a computational burden in the classification task and may lead to overfitting.

To address these research gaps, we propose a novel deep learning architecture leveraging attention and choosing a pooling layer which offers high discriminability and faster training. We train our deep learning model in an end-to-end approach. Therefore, it does not require an additional classifier for the classification purpose. Furthermore, with the help of the attention module as a deep learning layer, we capture the spatial relationship between the region of interests during the training process to better discriminate fundus images. Moreover, our model requires lower number of parameters as it leverages the appropriate layer of the deep learning model. We further give depth to our study by creating 2 separate image processing pipelines, contrast limited adaptive histogram equalization and Gaussian blurring, which are embedded into our proposed architecture, thus, providing important insights of the role of image enhancement as preprocessing step for deep learning based Diabetic Retinopathy Grading. In this work,

- We propose a novel deep learning model for DR severity grading that combines both CNN (VGG-16 architecture [41]) model and the attention module. This combination allows our model to able to capture more accurately the deteriorated regions in both local and global levels of fundus images.
- The proposed deep learning model can be trained in an end-to-end fashion without the need for a separate model for feature extraction and classifications.
- Our proposed method requires a lower number of tunable parameters as we use the 4th pooling layer.
- We evaluate our model on 2 different pre-processing pipelines, demonstrating the better fit for DR severity grading.

3 Methodology

3.1 Data generation and augmentation

In this study we use of the Kaggle EyePacs dataset [ref] for our modelling and evaluation. The dataset contains 35,126 labeled fundus image samples. There are five different levels (i.e., 0–4) of the DR severity depicted in each of fundus images in the dataset. Because there is uneven number of samples for each of the five different class of DR in the dataset, we addressed the class imbalance by generating additional samples for each class using data augmentation techniques as follows:

1. Cropping and resizing the images—The images must not be randomly cropped as the fundus has to appear wholly on the image and cropping randomly may lead to partial or no fundus to appear on the resultant picture. By dynamically obtaining the colour of the leftmost border pixel, we trim the whitespace around the fundus image by maintaining the aspect ratio and thumbnail of the size 224 × 224.
2. Generating samples—It is worth nothing that each eye image can depict a specific DR grade, regardless of whether the images are taken from the same patient. There is no guarantee that the left eye image sample taken from a patient will have the same DR grade as the right eye of the same patient. This fact can be exploited to generate fundus image samples from the existing ones in a particular grade by reflection of images. This helps to address not only the inter class imbalance, but also the intra

class imbalance of the number of samples for left eye images or right images. Thus, enhancing the robustness of our methodology. Figure 1 shows the samples from each class after cropping and resizing.

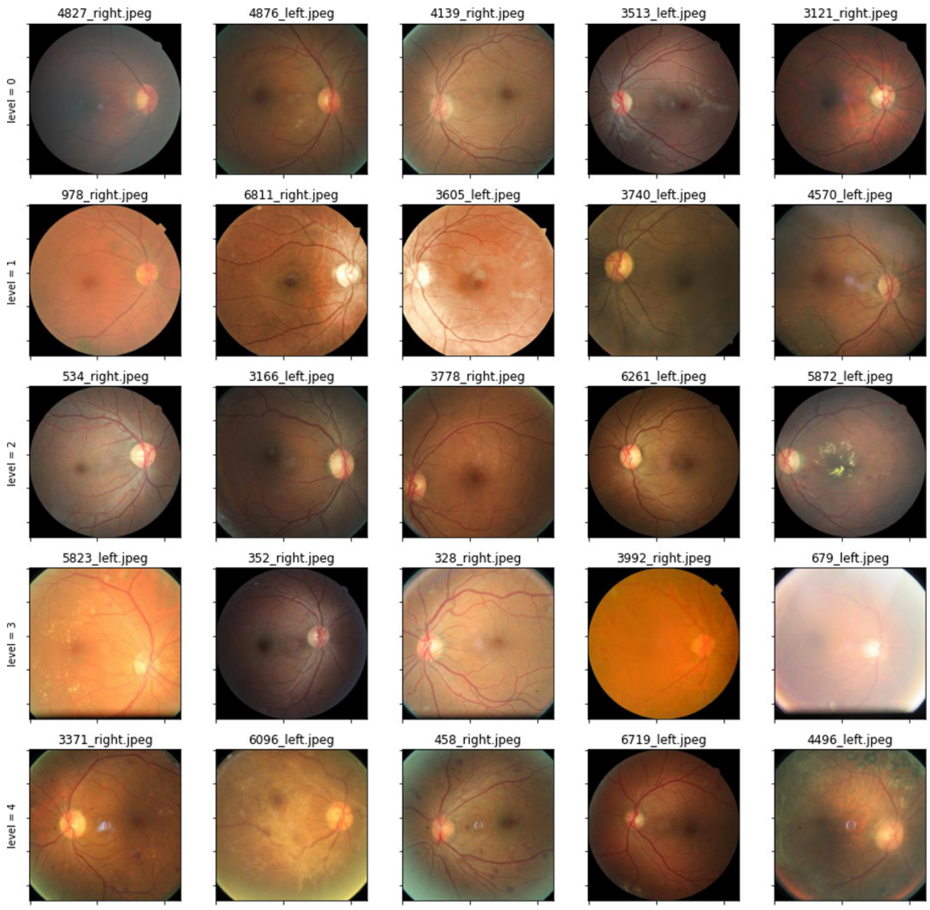


Fig. 1. Sample fundus images. level refers to the DR severity grade. 5 samples for each DR grade are displayed

3.2 Data preprocessing

It can be seen from Figure 1 that there exists intra-class variability in the sample images due to difference in brightness, contrast, and acquisition conditions. To overcome this, we apply two different preprocessing techniques; Contrast Limited Adaptive Histogram Equalization (CLAHE) [39] and Gaussian Blur [40], which have been shown to improve image appearances in many applications.

CLAHE. CLAHE was developed for medical imaging and is widely used in radiology image processing workflow. Histogram equalization is employed to increase the global contrast of low-contrast images, transforming the samples by redistributing the intensities across the dynamic range of the image. This also results in an increase of the foreground-background difference. To overcome the challenge of overamplification of noise in homogenous regions due to adaptive histogram equalization, CLAHE uses clipped amplification. Hence, a clip limit is provided as a parameter to the CLAHE method as a value for which the histogram is clipped. However, applying CLAHE to our dataset, which consists of RGB images, directly would affect the relative distribution of RGB Channels and dramatically impact the overall colour balance. To avoid this, the images are converted to LAB colour space—Luminance and A & B chromatic components. The luminance channel (L) is isolated and subjected to the CLAHE algorithm. While applying CLAHE, we experiment with the high contrast limit (clip limit) and size of the local neighbourhood for histogram equalization (tile size) features. Figure 2 showcases the results with parameter variations.

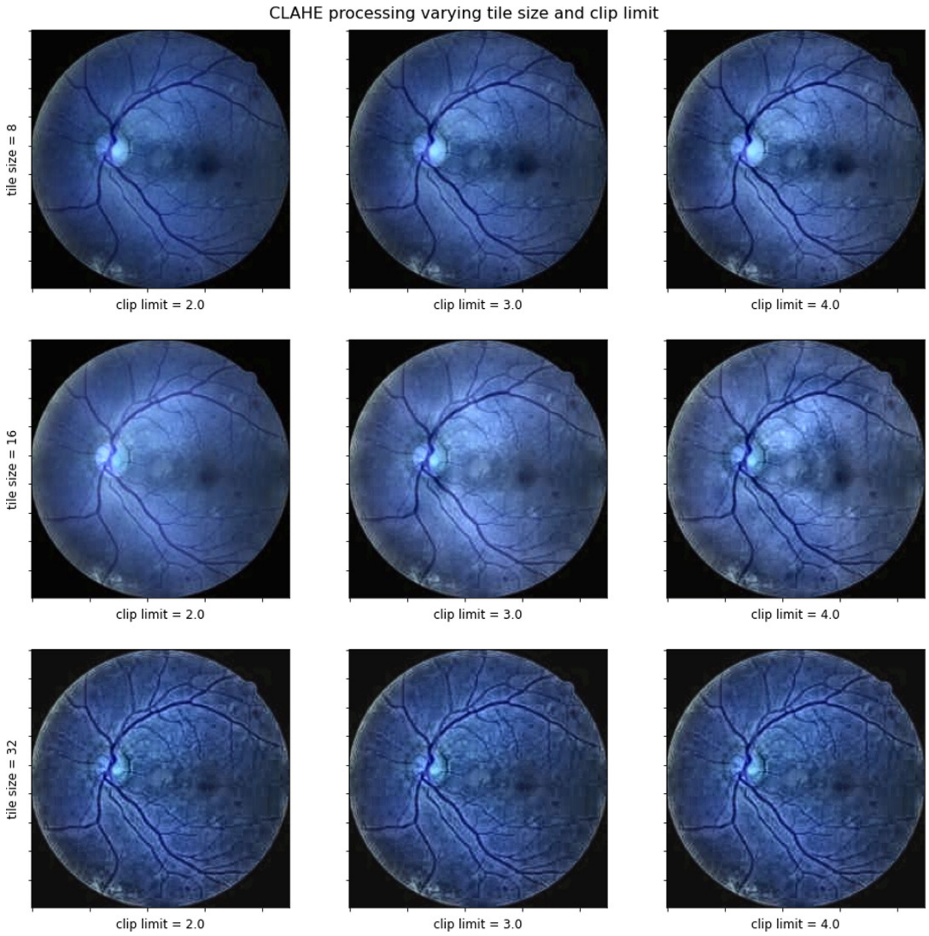


Fig. 2. CLAHE operation output. The output obtained by varying parameters like tile size and clip limit are visualized

CLAHE parameters are chosen to enhance the sample images to make markers prominent to facilitate feature extraction and learning. Figure 3 shows the final output of the pre-processing.

Fig. 3. Original image (left) and CLAHE enhanced image (right)

Gaussian blur. Gaussian blur [40] is a method used to bring out the distinctive features of the fundus images. The original image is convolved with a Gaussian kernel to get the blurred image. During the operation, the pixels near the center of the kernel are given more weight than those far away from the center. This is done on a channel-by-channel basis and the average channel values become the new value for the filtered pixel. Larger kernels have more values factored into the average, and this implies that a larger kernel will blur the image more than a smaller kernel.

Once the images are blurred by convolving the image with gaussian filter, then a weighted addition of the blurred image and the original image is done to get the final preprocessed image. Here the image is multiplied with their respective weight associated with them and then finally added to get the final image as shown in Equation 1.

$$dst = \alpha \cdot img1 + \beta \cdot img2 + \gamma \tag{1}$$

Figure 4 shows the final output of the Gaussian blurring with blood vessels and the optic disk highlighted.

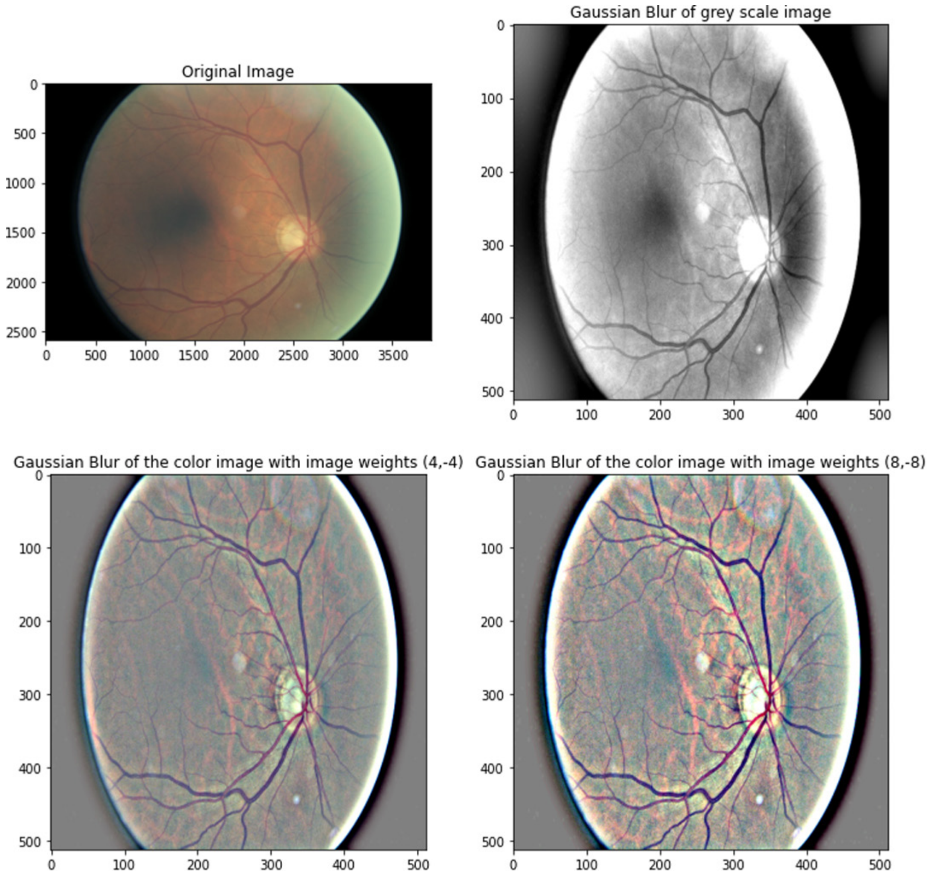


Fig. 4. Original image (top left) with Gaussian preprocessed images of different weights

3.3 Classifier modelling

We propose an attention-convolution model with VGG-16 [41]-like architecture as the base model. VGG-16 is a form of CNN that is used to extracts low level features from images. Its smaller kernel size makes it suitable for extracting features from DR RGB images [30]. To improve the performance of the model and prevent overfitting, we used pre trained weights from ImageNet [42] to initiate the training of the model and fine-tuned it accordingly to translate its performance to our target domain of fundus images. The proposed attention-convolution model relies on Attention, Convolution, Full-Connected layers along with a SoftMax classifier for multi-class classification. Figure 5 shows the block diagram of the proposed framework.

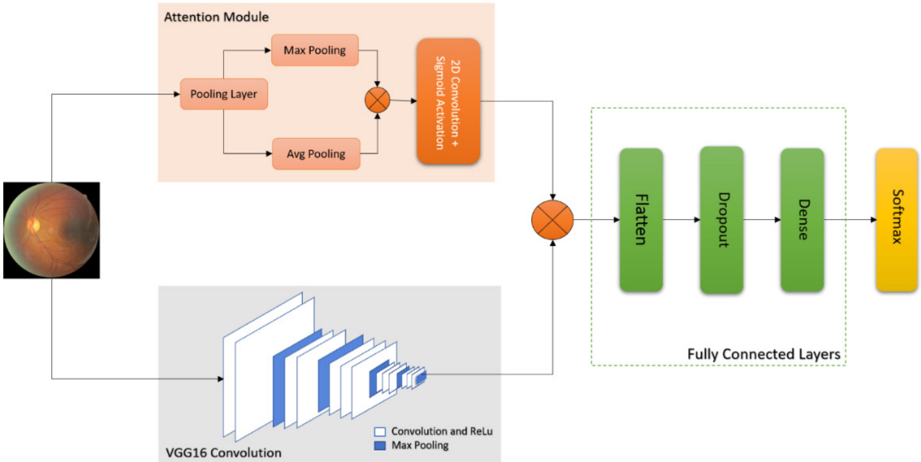


Fig. 5. Proposed system block diagram. The attention module shows the operations introduced for inducing attention and the VGG-16 module showcases the convolution operation. The results from both modules are concatenated and processed by the fully connected layers to be classified by the SoftMax layer

The visual cues in fundus images and their relationship in spatial domain are captured with the attention block. Woo et al. [18] first introduce the concept of spatial attention map. Unlike [18], our method uses the 4th pooling layer of VGG-16 as the input tensor, after which the output of the pooling layer is fed to another set of max and average pooling, respectively. The pooled tensors are concatenated using convolution of size 7 × 7 and stride 1 with sigmoid activation. If F^s_{avg} and F^s_{max} are 2D tensors obtained after average pooling and max pooling respectively, then the output of the attention module would be a spatial attention map $M_s(F)$ that encodes the information of the regions to accentuate and subdue. This map is computed as follows:

$$M_s(F) = \sigma\,(f^{7\times 7}([F^s_{avg}; F^s_{max}])),$$

where, σ is the sigmoid activation function and $f^{7\times 7}$ is the convolution operation of filter size 7 × 7.

To capture ROIs in input fundus image, scale invariant convolution module is introduced. The results of the convolution and attention modules are concatenated. The concatenated features are then flattened, a dropout of 0.5 is used, and a dense layer of 256 nodes is outputted. The function of the fully connected module is to transform the extracted features into a 1-Dimensional tensor which is to be used for Diabetic Retinopathy grading. The DR severity grading is produced using SoftMax classifier. The SoftMax layer forms the final dense layer which determines the multinomial distribution of probability scores. We fix number of nodes in the output layer to 5, according to the DR severity levels. Table 1 summaries the proposed model architecture.

Table 1. Proposed model architecture with output shape at each layer

Layer (Type)	Output Shape
VGG-16 (Model)	(None, 16, 16, 512)
Lambda (Average Pooling)	(None, 16, 16, 1)
Lambda (Max Pooling)	(None, 16, 16, 1)
Concatenate	(None, 16, 16, 2)
Conv2D	(None, 16, 16, 1)
Concatenate	(None, 16, 16, 513)
Flatten	(None, 131328)
Dropout	(None, 131328)
Dense	(None, 256)
Dense (SoftMax)	(None, 5)

4 Results and discussions

This section presents the experimental setup, model results and evaluation. We draw comparison of our model's performance with state-of-the-art models introduced in the field of Diabetic Retinopathy Grading which have experimented with attention. The dataset described in Section 4.1 was used for the experimentation.

4.1 Experimental setup

The proposed models were implemented in TensorFlow deep learning framework [43] using Python [46] and Keras [45]. We performed the model training and validation on a Google Colab VM running NVIDIA K80 GPU with 12GB RAM. The training dataset is localized and cropped as discussed in Section 4.2 and then fed into 2 preprocessing pipelines; CLAHE processing and Gaussian Blur processing separately, and the resulting images were then analyzed by our proposed classification model. The pipeline for training and validation is shown in Figure 6. Adaptive Moment Estimation (Adam) optimizer [44], with an initial learning rate of 0.0001 is used as an optimizer and a batch size of 32 is used for both training and validation.

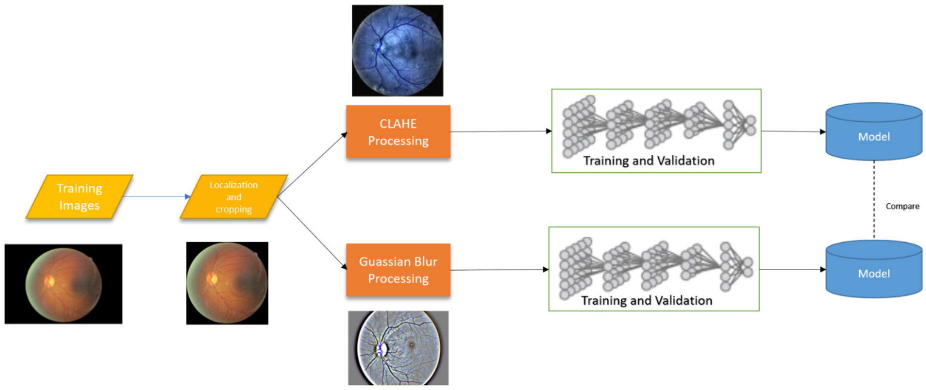
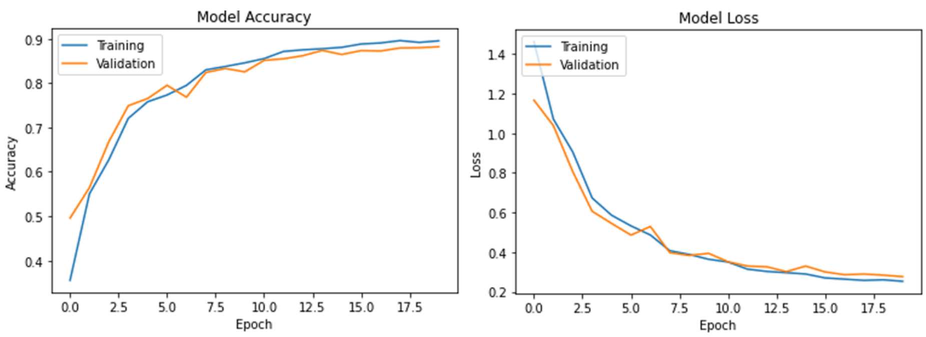


Fig. 6. Overview of training process. Note that the model is purely representative

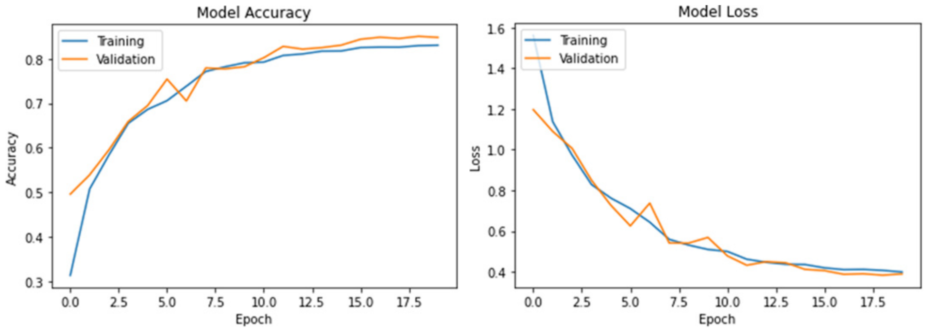
4.2 Performance analysis

The performance of the two image enhancement pipelines discussed in Section 4.2.1 and 4.2.2 are analyzed and compared. The performance metric used is a weighted average of the accuracy. Figures 7 and 8, show the model performances of the proposed model on the training and validation sets. Table 2 shows the training and validation accuracy. The model with the CLAHE image enhancement mechanism produces the best results with 92.51% training accuracy and 89.20% validation accuracy, compared with the model with the Gaussian Blur image enhancement technique that recorded accuracy of 83.79% and 82.64% for training and validation, respectively.



(a) (b)

Fig. 7. Qualitative analysis of (a) model accuracy and (b) model loss for CLAHE+proposed model pipeline



(a) (b)

Fig. 8. Qualitative analysis of (a) model accuracy and (b) model loss for Gaussian blur+proposed model pipeline

Table 2. Comparison of results for CLAHE and Gaussian blur image processing with our proposed architecture

Model	Training Accuracy	Validation Accuracy
CLAHE + Proposed Model	**0.9251**	**0.8920**
GaussianBlur + Proposed Model	0.8379	0.8264

We also compare the performance of our best model with that obtained by using other state-of-the-art method that employed attention augmentation in their work on the same datasets. Tables 3 and 4 show the comparison of our best performing model that uses CLAHE image enhancement. We notice that our method outperforms the state-of-the-art methods with an improvement margin of at most 64.24%.

Table 3. Comparison of our model performance with existing Attention nets

Model	Accuracy
BIRA-NET [35]	0.5431
SeaNET [47]	0.5994
CABNet [36]	0.8618
Proposed Model	**0.8920**

Table 4. Comparison of our model performance with existing Transfer learning methods

Model	Accuracy
VGG-16 [48]	0.6554
ResNet-50 [48]	0.6900
Inceptionv3 [48]	0.7029
Proposed Model	**0.8920**

From our analysis it is apparent that leveraging image enhancement improves the feature extraction capabilities of convolutional neural networks. While both the studied image enhancement methods show significant increase in performance, CLAHE helps overcome the variability in training dataset classes, which is especially imperative in the case of DR grading. The VGG-16 helps in extracting ROIs in medical fundus images, while the attention captures the interesting regions in the input image and leveraging the 4th pooling layer of VGG-16 helps construct a model pipeline which is not only performs well but is efficient as well.

5 Conclusion

In this paper we propose a spatial attention augmented VGG-16 model for Diabetic retinopathy grading of RGB fundus eye images. We extend the proposed architecture by incorporating an image enhancement pipeline. To choose the image enhancement most suitable for Diabetic retinopathy severity grading, we study the CLAHE technique and Gaussian Blurring in tandem with our proposed deep neural network architecture. The evaluation results show that the CLAHE image enhancement technique improve the results more than Gaussian Blurring method. Also, the experimental results show the effectiveness of the proposed method in comparison to other state-of-the-art methods on the same dataset. Future work will explore using other forms of deep neural network architecture other than VGC-16 as base classifier. Additionally, the use of advanced sample generation techniques such as Convolutional Autoencoder before classification would be explored for the task of Diabetic retinopathy severity grading.

Conflicts of Interest: "The authors declare no conflict of interest."

6 References

[1] Gaurav Saxena, et al., Improved and robust deep learning agent for preliminary detection of diabetic retinopathy using public datasets, Intelligence-Based Medicine 3–4 100022 (2020). https://doi.org/10.1016/j.ibmed.2020.100022

[2] Stela Vujosevic, MD, et al., Screening for diabetic retinopathy: new perspectives and challenges, The Lancet 8 4 (2020) 337–347. https://doi.org/10.1016/S2213-8587(19)30411-5

[3] Chintala Narsaiah, et al., Study on awareness and assessment of diabetic retinopathy in diabetic patients attending ophthalmology clinic at a tertiary care hospital, Telangana state, IJCMR 6 11 (2019). https://doi.org/10.21276/ijcmr.2019.6.11.39

[4] Tey, K. Y., Teo, K., Tan, A. C. S. et al., Optical coherence tomography angiography in diabetic retinopathy: a review of current applications, Eye and Vis 6, 37 (2019). https://doi.org/10.1186/s40662-019-0160-3

[5] Shiva Shankar Reddy, et al., Extensive analysis of machine learning algorithms to early detection of diabetic retinopathy, Materials Today: Proceedings (2020) https://doi.org/10.1016/j.matpr.2020.10.894

[6] Javeria Amin, et al., A method for the detection and classification of diabetic retinopathy using structural predictors of bright lesions, Journal of Computational Science 19 (2017) 153–164. https://doi.org/10.1016/j.jocs.2017.01.002

[7] Wen Cao, et al., Microaneurysm detection using principal component analysis and machine learning methods, IEEE Transactions on NanoBioscience 17 3 (2018) 191–198. https://doi.org/10.1109/TNB.2018.2840084
[8] Revathy R, et al., Diabetic retinopathy detection using machine learning, IJERT 9 6 (2020) https://doi.org/10.17577/IJERTV9IS060170
[9] Welikala, et al., Genetic algorithm based feature selection combined with dual classification for the automated detection of proliferative diabetic retinopathy, Computerized Medical Images and Graphics 43 (2015) 64–77. https://doi.org/10.1016/j.compmedimag.2015.03.003
[10] S K, S., P, A., A machine learning ensemble classifier for early prediction of diabetic retinopathy. J Med Syst 41 201 (2017). https://doi.org/10.1007/s10916-017-0853-x
[11] Ullah, H, Saba, T, Islam, N, et al., An ensemble classification of exudates in color fundus images using an evolutionary algorithm based optimal features selection, Microsc Res Tech. 82 (2019) 361–372. https://doi.org/10.1002/jemt.23178
[12] Saha R., Chowdhury A. R., Banerjee S. (2016) Diabetic Retinopathy Related Lesions Detection and Classification Using Machine Learning Technology. In: Rutkowski L., Korytkowski M., Scherer R., Tadeusiewicz R., Zadeh L., Zurada J. (eds) Artificial Intelligence and Soft Computing. ICAISC 2016. Lecture Notes in Computer Science, vol 9693. Springer, Cham. https://doi.org/10.1007/978-3-319-39384-1_65
[13] Georgios Leontidis, et al., A new unified framework for the early detection of the progression to diabetic retinopathy from fundus images, Computers in Biology and Medicine 90 (2017) 98–115. https://doi.org/10.1016/j.compbiomed.2017.09.008
[14] Emran Saleh et al., Learning ensemble classifiers for diabetic retinopathy assessment, Artificial Intelligence in Medicine 85 (2018) 50–63. https://doi.org/10.1016/j.artmed.2017.09.006
[15] Hoque, M. E., Kipli, K., Deep learning in retinal image segmentation and feature extraction: a review, International Journal of Online and Biomedical Engineering (iJOE) 17 14 (2021) 103–118. https://doi.org/10.3991/ijoe.v17i14.24819
[16] Sohini Roychowdhury, et al., DREAM: Diabetic Retinopathy Analysis Using Machine Learning | IEEE Journals & Magazine, IEEE Journal of Biomedical and Health Informatics 18 5 (2014) 1717–1728. https://doi.org/10.1109/JBHI.2013.2294635
[17] Gule Sama, et al., Automatic detection and severity classification of diabetic retinopathy, Multimed Tools and Application 79 (2020) 31803–31817. https://doi.org/10.1007/s11042-020-09118-8
[18] Charu Bharadwaj, et al., Hierarchical severity grade classification of non-proliferative diabetic retinopathy, Journal of Ambient Intelligence and Human Computing 12 (2020) 2649–2670. https://doi.org/10.1007/s12652-020-02426-9
[19] Vijayan T, et al., Gabor filter and machine learning based diabetic retinopathy analysis and detection, Microprocessors and Microsystems 103353 (2020). https://doi.org/10.1016/j.micpro.2020.103353
[20] Gayathri, S., Gopi, V. P., Palanisamy, P., Automated classification of diabetic retinopathy through reliable feature selection, Phys Eng Sci Med 43 (2020) 927–945. https://doi.org/10.1007/s13246-020-00890-3
[21] K. Shankar, et al., Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model, Pattern Recognition Letters 133 (2020) 210–216. https://doi.org/10.1016/j.patrec.2020.02.026
[22] Eman Abdelmaksoud, et al., Automatic diabetic retinopathy grading system based on detecting multiple retinal lesions, IEEE Access 9 (2021) 15939–15960. https://doi.org/10.1109/ACCESS.2021.3052870
[23] Gayathri S, et al., A lightweight CNN for diabetic retinopathy classification from fundus images, Biomedical Signal Processing and Control 62 102115 (2020). https://doi.org/10.1016/j.bspc.2020.102115

[24] Mansour, R. F., Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy, Biomed. Eng. Lett. 8 (2018) 41–57. https://doi.org/10.1007/s13534-017-0047-y
[25] T. Shanthi, R. S. Sabeenian, Modified Alexnet architecture for classification of diabetic retinopathy images, Computers & Electrical Engineering 76 (2019) 56–64, ISSN 0045-7906. https://doi.org/10.1016/j.compeleceng.2019.03.004
[26] Zhan Wu, et al., Coarse-to-fine classification for diabetic retinopathy grading using convolutional neural network, Artificial Intelligence in Medicine 108 (2020) 101936. https://doi.org/10.1016/j.artmed.2020.101936
[27] Sambyal, N., Saini, P., Syal, R., Gupta, V., Aggregated residual transformation network for multistage classification in diabetic retinopathy. Int J Imaging Syst Technol (2020) 1–12. https://doi.org/10.1002/ima.22513
[28] Pan, X., Jin, K., Cao, J. et al., Multi-label classification of retinal lesions in diabetic retinopathy for automatic analysis of fundus fluorescein angiography based on deep learning. Graefes Arch Clin Exp Ophthalmol 258 (2020) 779–785. https://doi.org/10.1007/s00417-019-04575-w
[29] J. Wang, Y. Bai and B. Xia, "Simultaneous Diagnosis of Severity and Features of Diabetic Retinopathy in Fundus Photography Using Deep Learning," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 12, pp. 3397–3407, Dec. 2020. https://doi.org/10.1109/JBHI.2020.3012547
[30] Shaohua Wan, Yan Liang, Yin Zhang, Deep convolutional neural networks for diabetic retinopathy detection by image classification, Computers & Electrical Engineering 72 (2018) 274–282, ISSN 0045–7906. https://doi.org/10.1016/j.compeleceng.2018.07.042
[31] Gulshan V., Peng L., Coram M., et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316 22 (2016) 2402–2410. https://doi.org/10.1001/jama.2016.17216
[32] Sehrish Qummar, et al., A deep learning ensemble approach for diabetic retinopathy detection, IEEE Access 7 (2019) 150530–150539. https://doi.org/10.1109/ACCESS.2019.2947484
[33] Abbas, Q., Fondon, I., Sarmiento, A. et al., Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features. Med Biol Eng Comput 55 (2017) 1959–1974. https://doi.org/10.1007/s11517-017-1638-6
[34] Shankar, K., Perumal, E., Vidhyavathi, R. M., Deep neural network with moth search optimization algorithm based detection and classification of diabetic retinopathy images, SN Appl. Sci. 2 (2020) 748. https://doi.org/10.1007/s42452-020-2568-8
[35] Jadhav, A. S., Patil, P. B., Biradar, S., Optimal feature selection-based diabetic retinopathy detection using improved rider optimization algorithm enabled with deep learning, Evol. Intel. (2020). https://doi.org/10.1007/s12065-020-00400-0
[36] Along He, et al., CABNet: Category attention block for imbalanced diabetic retinopathy grading, IEEE Transactions of Medical Imaging 40 1 (2020) 143–153. https://doi.org/10.1109/TMI.2020.3023463
[37] Yanni, R. M. T., El-Ghitany, N. E.-K., Amer, K., Riad, A., El-Bakry, H., A new model for image segmentation based on deep learning. International Journal of Online and Biomedical Engineering (iJOE), 17 7(2021) 28–47. https://doi.org/10.3991/ijoe.v17i07.21241
[38] Kalyani, G., Janakiramaiah, B., Karuna, A. et al. Diabetic retinopathy detection and classification using capsule networks, Complex Intell. Syst. (2021). https://doi.org/10.1007/s40747-021-00318-9
[39] Reza, A. M., Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement, The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology 38, 35–44 (2004). https://doi.org/10.1023/B:VLSI.0000028532.53893.82

[40] M. Basu, "Gaussian-based edge-detection methods-a survey," in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 32, no. 3, pp. 252–260, Aug. 2002. https://doi.org/10.1109/TSMCC.2002.804448
[41] Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv:14091556
[42] B R, N., A, G. K., H S, C., M S, D., S, M., An ensemble deep neural network approach for oral cancer screening. International Journal of Online and Biomedical Engineering (iJOE) 17 2 (2021) 121–134. https://doi.org/10.3991/ijoe.v17i02.19207
[43] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, Xiaoqiang Zheng (2016), TensorFlow: A system for large-scale machine learning, arXiv:1605.08695v2
[44] Diederik P. Kingma, Jimmy Ba., Adam: A method for stochastic optimization, arXiv preprint arXiv 1412 (2014) 6980.
[45] Chollet F et al (2015) Keras. https://github.com/fchollet/keras
[46] Rossum G (1995) Python reference manual. Tech rep., Amsterdam
[47] Z. Zhao, K. Chopra, Z. Zeng and X. Li, "Sea-Net: Squeeze-And-Excitation Attention Net For Diabetic Retinopathy Grading," 2020 IEEE International Conference on Image Processing (ICIP), 2020, pp. 2496–2500. https://doi.org/10.1109/ICIP40778.2020.9191345
[48] S. Kajan, J. Goga, K. Lacko and J. Pavlovičová, "Detection of Diabetic Retinopathy Using Pretrained Deep Neural Net-works," 2020 Cybernetics & Informatics (K&I), 2020, pp. 1–5, doi: 10.1109/KI48306.2020.9039793. Budiman, R. (2013). Utilizing Skype for providing learning support for Indonesian distance learning students: A lesson learnt. Procedia - Social and Behavioral Sciences, 83: 5–10. https://doi.org/10.1016/j.sbspro.2013.06.002

7 Authors

S. Geetha, Vellore Institute of Technology, School of Computer Science and Engineering, Chennai, India; E-mail: geetha.s@vit.ac

Mansi Parashar, Vellore Institute of Technology, School of Computer Science and Engineering, Chennai, India; E-mail: mansiparashar00@gmail.com

JS Abhishek, Vellore Institute of Technology, School of Computer Science and Engineering, Chennai, India; E-mail: jsabhishek2001@gmail.com

Raj Vishal Turaga, Vellore Institute of Technology, School of Computer Science and Engineering, Chennai, India; E-mail: vishalturaga@gmail.com

Isah A. Lawal, Department of Applied Data Science, Noroff University College, Kristiansand, Norway; E-mail: Isah.Lawal@noroff.no

Seifedine Kadry, Department of Applied Data Science, Noroff University College, Kristiansand, Norway

Article submitted 2022-02-08. Resubmitted 2022-03-19. Final acceptance 2022-03-21. Final version published as submitted by the authors.