"Efficient Classification of Diabetic Retinopathy Stages using VGG-NIN Deep Learning Architecture"

Neelam S. Nikale [1], Dr.Swati Bhavsar[2]

[1] *ME(computer Student , Dept. of Computer Engineering Matoshri College of Engineering & R. C. Nashik Maharastra,India*

[2] *Assistant Professor, Dept. of Computer Engineering Matoshri College of Engineering & R. C. Nashik Maharastra,India*

ABSTRACT

Diabetic retinopathy (DR) is a serious condition that damages retinal blood vessels, potentially leading to blindness. Traditional diagnosis involves manual analysis of colored fundus images by clinicians, which is error-prone and time-consuming. To mitigate these challenges, computer vision techniques have been utilized for automating DR detection. However, existing methods often struggle with computational complexity and inadequate feature extraction for precise DR stage classification. This paper introduces a novel approach for classifying DR stages with minimal learnable parameters to enhance training efficiency and model convergence. The VGG16 architecture, augmented with a spatial pyramid pooling layer (SPP) and network-in-network (NiN) structures, constitutes the VGG-NiN model, capable of effectively processing DR images across different scales due to the adaptability of the SPP layer. Moreover, the incorporation of NiN enhances the model's ability to capture nonlinear features, thereby improving classification accuracy. Experimental findings validate the efficacy of the proposed model, demonstrating superior performance in terms of accuracy and computational efficiency compared to existing techniques.

Keyword : - *CNN, colored fundus images, diabetic retinopathy, deep learning*

---

1. INTRODUCTION

Diabetic retinopathy, a common complication of diabetes, can lead to vision loss if left undetected and untreated. This paper aims to utilize advanced deep learning techniques, specifically the VGG-NIN architecture, for the early detection of diabetic retinopathy from retinal images.

Diabetic Retinopathy (DR) stands as one of the leading causes of blindness worldwide among diabetic patients. Its early detection plays a pivotal role in preventing irreversible vision loss. This paper endeavors to leverage the potential of deep learning techniques, specifically the VGG-NIN architecture, to develop a robust and efficient system for the automated detection of diabetic retinopathy from retinal fundus images.

The introduction of the paper on "Diabetic Retinopathy Detection Using VGG-NIN: A Deep Learning Architecture" delves into the prevalence and impact of diabetes mellitus (DM), specifically on eye health, leading to diabetic retinopathy (DR). It highlights key statistics regarding the global occurrence of DR, its papered rise, and the prevalence within specific regions, emphasizing the urgency of early detection.

The section outlines the progression of DR stages, from non-proliferative diabetic retinopathy (NPDR) to proliferative diabetic retinopathy (PDR), detailing the distinctive characteristics and effects on the human retina.

Emphasizing the importance of early detection, the text mentions symptoms associated with different stages of DR and the challenges in identifying asymptomatic cases.

Highlighting the significance of this paper, the introduction focuses on the need for efficient and accurate detection methods. It mentions the limitations of conventional techniques and introduces convolutional neural networks (CNNs) as a promising approach for image classification tasks.

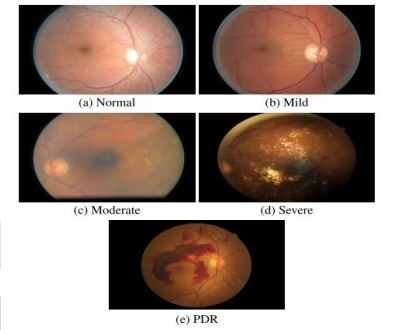

FIGURE : The different stages of DR

2. OBJECTIVES

The primary objective of this paper is to design, develop, and evaluate a deep learning-based system using the VGG-NIN architecture to accurately identify and classify diabetic retinopathy stages from retinal fundus images. This system aims to achieve high accuracy, sensitivity, and specificity in the detection process, providing a reliable tool for healthcare professionals.

The paper's main objective is introduced: to detect all five stages of DR using a minimal set of learning parameters and improve existing classification results by employing a modified VGG16 model coupled with a new version of the Network in Network (NiN) architecture. The use of transfer learning and the application of the proposed architecture to enhance classification results are discussed

3. RELAETD WORK

An exploration of current methodologies and technologies used in diabetic retinopathy detection, emphasizing their strengths, weaknesses, and applicability.

Review existing methods, algorithms, and technologies employed in diabetic retinopathy detection, highlighting their strengths, limitations, and relevance in comparison to deep learning-based approaches like VGG-NIN.

Gondal et al. [14]: Proposed Referable Diabetic Retinopathy (RDR) CNN for binary classification of DR stages, grouping stages 0/1 and 2/3/4 due to challenges in detecting stage-1 features.

Yang et al. [15]: Implemented a DCNN with global and local networks for two-stage DR classification (NDPR and normal), evaluated using kappa score. Limitation: did not consider the entire dataset of five stages. **Garcia et al. [16]:** Experimented with CNN networks like AlexNet and VGGNet16, focusing on contrast improvement. Achieved best-reported results on VGG16 with 0.54 sensitivity, 0.93 specificity, and 0.83 accuracy without explicitly mentioning DR stages.

Dutta et al. [11]: Proposed three neural network models (Feed Forward NN, Deep NN, Convolutional NN) achieving an accuracy of 0.89 on the training dataset. Preprocessing involved mean, median, and standard deviation calculation.

Authors in [10], [17]–[22]: Various methods proposed for detecting different DR stages.

For instance:

[23]: Achieved 94% accuracy on the DRIVE dataset using deep CNN architecture for DR detection with spatial feature analysis.

[17]: Proposed AI-based disease-staging considering retinal area for decision-making. Used modified Davis staging. Model output had a lower false-negative rate (FNR) than false positive rate (FPR).

[10]: Implemented a CNN model and MA and HE detection but struggled to classify early stages (1 and 2), classifying them as class 0 (negative class).

Pratt et al. [20]: Proposed CNN for classifying five DR stages but struggled with accurate classification of the mild stage due to skewed dataset, resulting in high specificity and low sensitivity.

4.**OTHER NOTABLE WORKS**:

[24]: Used RCNN ResNet for detecting all DR stages.

[25]: Utilized SVM and deep learning with 170 color fundus images for DR classification.

[26]: Claimed to speed up DR detection using deep learning and backpropagation compared to SVM.

[22]: Proposed CNN model with ICA for some DR stages.

[28]: Developed a novel Deep CNN for classifying all five DR stages using retinal fundus images.

Qummar et al. [9]: Employed pretrained models (Resnet50, Dense169, Inceptionv3, Dense121, Xception) achieving 80.8% accuracy for multi-class classification but faced complexity issues in their ensemble model. Did not specify individual model performance on DR stages.

Overall observations: Existing models have limitations in detecting certain DR stages (especially early stages), struggles with multi-class classification, and may overlook crucial features. There's a lack of detailed discussion on model complexity across related works.

5. **SUMMARY OF RELATED WORK**:

Single and Binary Classification:

Gondal et al. [14] proposed a CNN for Referable Diabetic Retinopathy (RDR) classification. They achieved good results in binary classification, grouping stages 0/1 and 2/3/4 due to the challenge of detecting stage-1 features.

Yang et al. [15] used a DCNN with global and local networks for NDPR vs. normal classification, though limited by not considering the entire five-stage dataset.

Garcia et al. [16] focused on improving CNN networks like Alexnet and VGGnet16 for DR but didn't explicitly mention DR stages. Best-reported results were with VGG16, achieving 0.54 sensitivity, 0.93 specificity, and 0.83 accuracy.

Dutta et al. [11] proposed three NN models achieving 0.89 accuracy on the training dataset, employing various preprocessing steps.

Multi-Class Classification:

Various authors proposed methods for detecting the severity of DR across its five stages.

Some studies [23], [17], [10], [20], [24], [25], [26], [22], [28] used different models such as CNNs, RCNN Resnet, SVM, and deep learning for classification but faced challenges in accurately detecting early stages or all five stages of DR.

Recent Significant Work:

Qummar et al. [9] used five pretrained models achieving 80.8% accuracy for multi-class classification on the Kaggle dataset but faced complexity issues with the ensemble model.

6: CONTRIBUTION OF PROPOSED MODEL.

Dataset Division: Split into training (64%), validation (20%), and test (16%) sets.

Learning Rate: Adaptive learning rate from 0.01 to 0.0001 to prevent overfitting.

Image Augmentation: Utilized Keras Image Data Generator for augmentation (rescale, shear, zoom, flip).

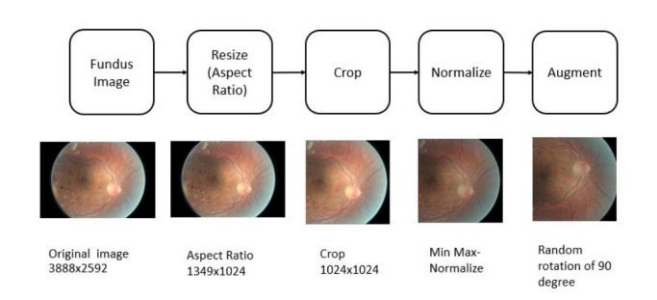

FIGURE : Preprocessing steps.

No existing model efficiently detected all DR stages with minimal learnable parameters. The proposed model aims to achieve this and perform better than current state-of-the-art models.

The paper highlights the total trainable parameters of the proposed model and discusses its complexity, which could potentially lead to reduced time and space complexity compared to other models discussed in related work.

Notes:

Existing models struggle with accurate detection of early/mild stages of DR.

The proposed model aims to address these limitations by efficiently detecting all stages of DR with fewer learnable parameters and enhanced performance compared to existing models.

Comprehensive evaluation and comparison of the proposed model against existing approaches remain essential for validation and establishing its superiority in detecting all stages of DR.

7.PROPOSED WORK

Introducing the proposed approach utilizing VGG-NIN architecture and its potential advantages over existing systems. Discuss the proposed system's novelty, advantages over existing methods, and how the utilization of VGG-NIN architecture addresses the limitations of traditional approaches.

A. Dataset Description

Source: Kaggle dataset organized by EyePACS, comprising 88,702 images.

Labelled Images: 35,126 labelled images for diabetic retinopathy classification.

Classes: Dataset contains five distinct classes based on the severity of DR.

Distribution: Refer to Table 1 in the paper for class distribution details.

B. Preprocessing

Image Resizing: Initially resized images to 1349 × 1024 while maintaining the aspect ratio.

Random Cropping: Images cropped to a fixed size of 1024 × 1024.

Dataset Division: Split into training (64%), validation (20%), and test (16%) sets.

Learning Rate: Adaptive learning rate from 0.01 to 0.0001 to prevent overfitting.

Image Augmentation: Utilized Keras Image Data Generator for augmentation (rescale, shear, zoom, flip).

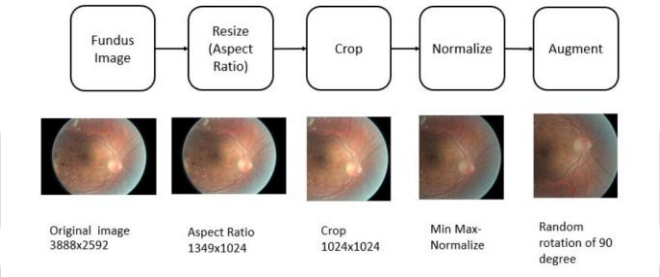

FIGURE : Preprocessing steps.

C. The VGG-NiN Model

Model Architecture: Combination of VGG16, SPP, and NiN.

VGG16 Handling: Accepts RGB images of size 224 × 244; input passed through convolutional layers and fully connected layers.

Spatial Pyramid Pooling (SPP): Used between the last conv layer and the first fully connected layer to generate a fixed-size output vector compatible with subsequent layers' requirements. Handles input feature map variations.

NiN (Network in Network): Added on top of SPP layer, uses mlpconv layers for non-linear feature encoding and high-level abstraction.

Parametric Relu (PRelu): Activation function in mlpconv layers, enabling adaptive rectifier parameter learning and reducing overfitting risk.

Softmax& Categorical Cross-Entropy: Activation and loss function for classification; Stochastic Gradient Descent (SGD) optimizer for training.

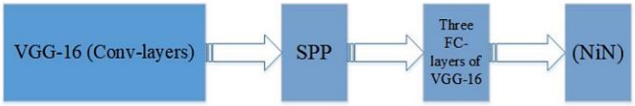

FIGURE: Proposed network architecture

D. Initialization and Hyperparameters

Model Initialization: VGG network's convolutional layers frozen using transfer learning; NiN initialized via the Xavier method.

Hyperparameters: Refer to Table 3 in the paper for different hyperparameter values.

Learning Rate Adaptation: Adaptive learning rate with adjustments if validation loss stagnates over five consecutive iterations. This proposed methodology amalgamates VGG16, SPP, and NiN, leveraging transfer learning, adaptive learning rates, and sophisticated data preprocessing to address the challenge of diabetic retinopathy detection using deep learning techniques.

Batch size	8
Initial learning rate	0.01
Momentum	0.9
Minimum learning rate	0.000,1
Number of epoch	50

TABLE 1: Hyper-parameters

SYSTEM STRUCTURE

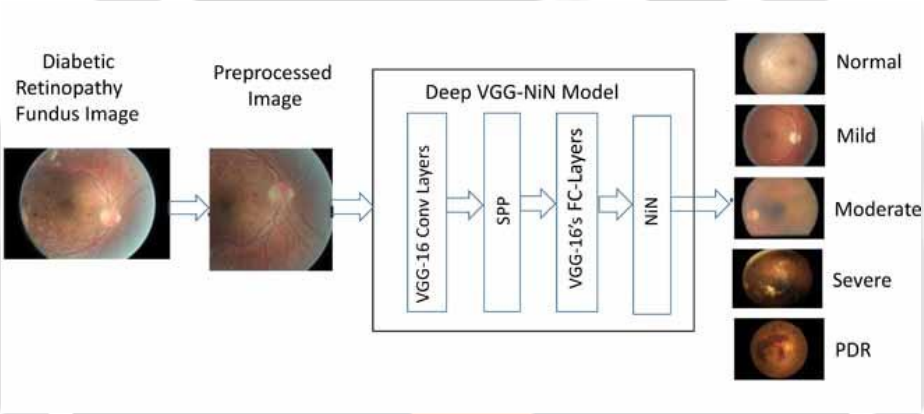

Figure: Deep VGG-NiN model for the classification of Diabetic Retinopathy Stages using Fundus Images.

The architecture diagram for the Deep VGG-NiN model used in the classification of Diabetic Retinopathy Stages using Fundus Images, as described in the IEEE paper titled "Diabetic Retinopathy Detection Using VGG-NIN: a Deep Learning Architecture," would likely involve a combination of two well-known deep learning architectures: VGG (Visual Geometry Group) and NiN (Network in Network).

VGG-16 architecture:

The VGG-16 is a version of the popular convolutional neural network called VGG-Net. VGG-16 consists of several layers, including 13 convolutional layers and 3 fully connected layers. He must therefore learn the weights of 16 diapers. It takes as input a color image of size 224××224 px and classifies it in one of the 1000 classes. It therefore returns a vector of size 1000, which contains the probabilities of belonging to each of the classes. The architecture of VGG-16 is illustrated in the figure. Each convolutional layer uses color filters of size 3××3 px, moved with a step of 1 pixel. The zero-padding is 1 pixel so that the input volumes have the same output dimensions. The number of filters varies depending on the "block" in which the layer is located. In addition, a bias parameter is introduced into the convolution product for each filter. Each convolutional layer has the function of activating a ReLU. In other words, there is always a ReLU correction layer after a convolution layer. The pooling operation is performed with cells of size 2××2 px and a step of 2 px, so the cells do not overlap. The first two fully connected layers each compute a vector of size 4096 and are each followed by a ReLU layer. The last returns the probability vector of size 1000 (the number of classes) by applying the Softmax function. In addition, these three layers use a bias parameter for each element of the output vector. In our case, we kept the same architecture with the addition of Softmax as an activation function.

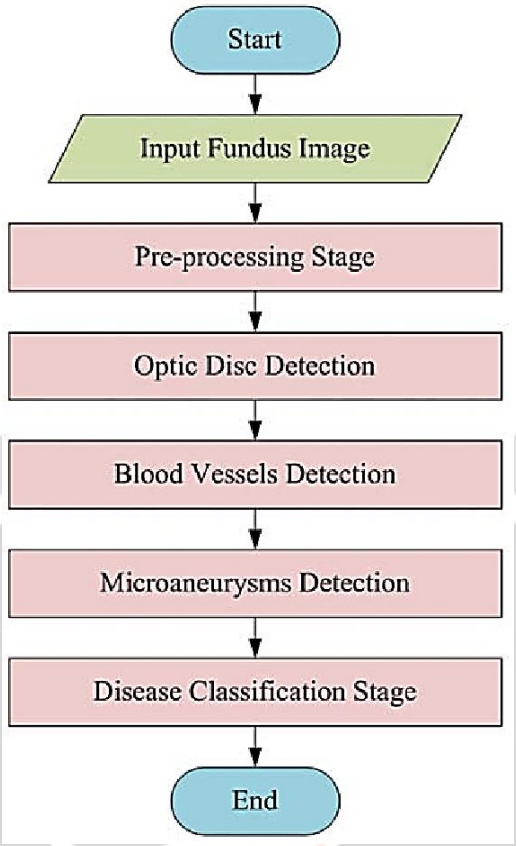

Figure: General Flowchart of Detection of Diabetic Retinopathy

8. CONCLUSIONS

In this paper, the authors introduced modifications to the existing Convolutional Neural Network (CNN) architecture, aiming to improve the efficiency and accuracy of diabetic retinopathy (DR) stage classification in color fundus images while reducing the number of learnable parameters compared to their previous ensemble model. They utilized imbalanced versions of the Kaggle dataset to validate their proposed model's performance measures. The results demonstrated that their proposed model achieved superior performance compared to both state-of-the-art ensemble and non-ensemble methods, while also being computationally efficient. The authors aim to continue enhancing their model's architecture and preprocessing techniques to further improve its performance, especially in classifying the early stages of DR

6. REFERENCES

[1] S. Jan, I. Ahmad, S. Karim, Z. Hussain, M. Rehman, and A. A. Shah, ‘‘Status of diabetic retinopathy and its presentation patterns in diabetics at ophthalomogy clinics,’’ J. Postgraduate Med. Inst. (Peshawar-Pakistan), vol. 32, no. 1, 2018.
[2] L. Math and R. Fatima, ‘‘Adaptive machine learning classification for diabetic retinopathy,’’ Multimedia Tools Appl., vol. 80, pp. 5173–5186, Oct. 2020.
[3] A. He, T. Li, N. Li, K. Wang, and H. Fu, ‘‘CABNet: Category attention block for imbalanced diabetic retinopathy grading,’’ IEEE Trans. Med. Imag., vol. 40, no. 1, pp. 143–153, Jan. 2021.
[4] L. Andersen and P. Andersson, ‘‘Deep learning approach for diabetic retinopathy grading with transfer learning,’’ Tech. Rep., 2020.
[5] N. Congdon, Y. Zheng, and M. He, ‘‘The worldwide epidemic of diabetic retinopathy,’’ Indian J. Ophthalmol., vol. 60, no. 5, p. 428, 2012.
[6] W. R. Memon, B. Lal, and A. A. Sahto, ‘‘Diabetic retinopathy,’’ Prof. Med. J., vol. 24, no. 2, pp. 234–238, 2017.
[7] R. Sarki, K. Ahmed, H. Wang, and Y. Zhang, ‘‘Automatic detection of diabetic eye disease through [2].

deep learning using fundus images: A survey,'' IEEE Access, vol. 8, pp. 151133–151149, 2020.
[8] R. E. Putra, H. Tjandrasa, and N. Suciati, ''Severity classification of non-proliferative diabetic retinopathy using convolutional support vector machine,'' Int. J. Intell. Eng. Syst., vol. 13, no. 4, pp. 156–170, Aug. 2020.
[9] S. Qummar, F. G. Khan, S. Shah, A. Khan, S. Shamshirband, Z. U. Rehman, I. Ahmed Khan, and W. Jadoon, ''A deep learning ensemble approach for diabetic retinopathy detection,'' IEEE Access, vol. 7, pp. 150530–150539, 2019.
[10] R. Ghosh, K. Ghosh, and S. Maitra, ''Automatic detection and classification of diabetic retinopathy stages using CNN,'' in Proc. 4th Int. Conf. Signal Process. Integr. Netw. (SPIN), Feb. 2017, pp. 550–554.
[11] S. Dutta, B. C. Manideep, S. M. Basha, R. D. Caytiles, and N. C. S. N. Iyengar, ''Classification of diabetic retinopathy images by using deep learning models,'' Int. J. Grid Distrib. Comput., vol. 11, no. 1, pp. 89–106, Jan. 2018.
[12] X. Zeng, H. Chen, Y. Luo, and W. Ye, ''Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network,'' IEEE Access, vol. 7, pp. 30744–30753, 2019.
[13] T. Nazir, A. Irtaza, Z. Shabbir, A. Javed, U. Akram, and M. T. Mahmood, ''Diabetic retinopathy detection through novel tetragonal local octa patterns and extreme learning machines,'' Artif. Intell. Med., vol. 99, Aug. 2019, Art. no. 101695.
[14] W. M. Gondal, J. M. Kohler, R. Grzeszick, G. A. Fink, and M. Hirsch, ''Weakly-supervised localization of diabetic retinopathy lesions in retinal fundus images,'' in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2017, pp. 2069–2073.
[15] Y. Yang, T. Li, W. Li, H. Wu, W. Fan, and W. Zhang, ''Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks,'' in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Springer, 2017, pp. 533–540.
[16] G. García, J. Gallardo, A. Mauricio, J. López, and A. D. Carpio, ''Detection of diabetic retinopathy based on a convolutional neural network using retinal fundus images,'' in Proc. Int. Conf. Artif. Neural Netw. Springer, 2017, pp. 635–642.
[17] H. Takahashi, H. Tampo, Y. Arai, Y. Inoue, and H. Kawashima, ''Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy,'' PLoS ONE, vol. 12, no. 6, Jun. 2017, Art. no. e0179790.
[18] Z. Wang, Y. Yin, J. Shi, W. Fang, H. Li, and X. Wang, ''Zoom-in-net: Deep mining lesions for diabetic retinopathy detection,'' in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Springer, 2017, pp. 267–275.
[19] Q. Abbas, I. Fondon, A. Sarmiento, S. Jiménez, and P. Alemany, ''Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features,'' Med. Biol. Eng. Comput., vol. 55, no. 11, pp. 1959–1974, Nov. 2017.
[20] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, ''Convolutional neural networks for diabetic retinopathy,'' Procedia Comput. Sci., vol. 90, pp. 200–205, 2016.
[21] N. B. Prakash, D. Selvathi, and G. R. Hemalakshmi, ''Development of algorithm for dual stage classification to estimate severity level of diabetic retinopathy in retinal images using soft computing techniques,'' Int. J. Elect. Eng. Inform., vol. 6, no. 4, 2014.
[22] J. de la Torre, A. Valls, D. Puig, and P. Romero-Aroca, ''Identification and visualization of the underlying independent causes of the diagnostic of diabetic retinopathy made by a deep learning classifier,'' 2018, arXiv:1809.08567. [Online]. Available: http://arxiv.org/abs/1809.08567
[23] C. T. R. Kathirvel, ''Classifying diabetic retinopathy using deep learning architecture,'' Int. J. Eng. Res., vol. V5, no. 6, pp. 19–24, May 2016.
[24] A. Kind and G. Azzopardi, ''An explainable AI-based computer aided detection system for diabetic retinopathy using retinal fundus images,'' in Proc. Int. Conf. Comput. Anal. Images Patterns. Springer, 2019, pp. 457–468.
[25] E. Bhatti and P. Kaur, ''DRAODM: Diabetic retinopathy analysis through optimized deep learning with multi support vector machine for classification,'' in Proc. Int. Conf. Recent Trends Image Process. Pattern Recognit. Springer, 2018, pp. 174–188.
[26] P. N. S. Kumar, R. U. Deepak, A. Sathar, V. Sahasranamam, and R. R. Kumar, ''Automated detection system for diabetic retinopathy using two field fundus photography,'' Procedia Comput. Sci., vol. 93, pp. 486–494, 2016.
[27] A. Hyvärinen and E. Oja, ''Independent component analysis: Algorithms and applications,'' Neural Netw., vol. 13, nos. 4–5, pp. 411–430, 2000.
[28] S. Muhammad Saiful Islam, M. Mahedi Hasan, and S. Abdullah, ''Deep learning based early detection and grading of diabetic retinopathy using retinal fundus images,'' 2018, arXiv:1812.10595. [Online]. Available: http://arxiv.org/abs/1812.10595

[29] K. Simonyan and A. Zisserman, ‘‘Very deep convolutional networks for large-scale image recognition,’’ 2014, arXiv:1409.1556. [Online]. Available: http://arxiv.org/abs/1409.1556
[30] K. He, X. Zhang, S. Ren, and J. Sun, ‘‘Spatial pyramid pooling in deep convolutional networks for visual recognition,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 9, pp. 1904–1916, Sep. 2015.
[31] M. Lin, Q. Chen, and S. Yan, ‘‘Network in network,’’ 2013, arXiv:1312.4400. [Online]. Available: http://arxiv.org/abs/1312.4400