Communications in Mathematics and Applications
Vol. 14, No. 3, pp. 1283–1292, 2023
ISSN 0975-8607 (online); 0976-5905 (print)
Published by RGN Publications
DOI: 10.26713/cma.v14i3.2413

http://www.rgnpublications.com

Special Issue
Recent Trends in Mathematics and Applications
Proceedings of the International Conference of Gwalior Academy of Mathematical Sciences 2022
***Editors*: Vinod P. Saxena and Leena Sharma**

Research Article

Early Detection of Diabetic Retinopathy Using Deep Convolutional Neural Network

Rajeswari Kannan, S. R. Vispute, Reena Kharat, Dipti Salunkhe* and N. Vivekanandan

Pimpri Chinchwad College of Engineering (Savitribai Phule Pune University), Pune 411 044, Maharashtra, India
***Corresponding author:** dipti.salunkhe21@pccoepune.org

Received: February 16, 2023 **Accepted:** July 2, 2023

Abstract. The most frequent ailment in diabetics that can cause blindness is diabetic retinopathy. It takes time for the clinician to recognize the stage of diabetic retinopathy and report the findings. As a result, the suggested paper takes little time to analyse and is highly helpful in outlining the stages of the condition look for a way to categorise. This essay compares and contrasts various research papers and methodologies. Convolutional and pooling layers can be utilized to boost accuracy, and ReLU can be employed as an activation function, according to the methods we suggested. In this study, fully connected layers which are also useful for classification are proposed. There are numerous indicators of retinal injury, including the fovea, optic nerve head, effusions, blood vessels, hemorrhages, and ocular micro aneurysms. Classification, pre-processing, feature extraction, segmentation, and detection are among the difficulties. The purpose of this research is to use convolutional neural networks to identify diabetic retinopathy (CNN). We compared various CNN architecture models and found that pre-processing can raise model accuracy by up to 5%.

Keywords. Diabetic Retinopathy (DR), Image classification, Deep Convolutional Neural Network (DCNN), Exudates

Mathematics Subject Classification (2020). 05A15

1. Introduction

Recent decades have seen a sharp rise in the number of people with diabetes, which increases the risk of eye illnesses such diabetic retinopathy. The most common reason for blindness is diabetic retinopathy. One of the most common chronic diseases and the main factor in middle-aged blindness in affluent nations is DR (*Diabetic Retinopathy*). DR manifests as minute variations in retinal capillaries. This causes the mild, non-proliferative form of diabetic retinopathy, or the initial stage of DR the progression of vision impairment can be slowed but not entirely stopped with early identification. However, this is frequently challenging because symptoms frequently show up too late for efficient treatment. Only half of those who have *Diabetic Retinopathy* (DR), which affects an estimated 93 million people worldwide, are aware of it. Generally speaking, diabetic retinopathy has four stages. In its most advanced form, aberrant blood vessels proliferate all over the retina's surface, which can result in cell loss and scarring of the retina is a time-consuming process that can result in treatment delays, misconceptions, etc. even for highly experienced professionals. Automated approaches have been acknowledged as being crucial for the early diagnosis of diabetic retinopathy. The research community has invested a lot of time and energy into developing automated computer screening tools that can quickly identify diabetic retinopathy (Hatanaka *et al*. [2]).

Currently, DR diagnosis is a laborious and time-consuming process requiring skilled medical professionals to evaluate retinal color pictures. Then, professionals place the patient's eye's level of deterioration into one of four categories. Although it takes a lot of time, this method is effective. Results take around 2 days to come in, and it could be challenging to get in touch with the patient after that.

2. Literature Survey

Automating the process of detecting diabetic retinopathy requires pre-processing, features extraction, classification, analysis, and output indicating whether or not the image is diabetic. Studying many papers can help you determine which approach to use for the best accuracy. Wang *et al*. [12] uses picture improvement, scaling, histogram equalization, and green channel extraction to pre-process raw retinal fundus images.

For quantitative analysis, 14 additional features from the pre-processed pictures are also retrieved, including papillary, foveal, vascular, and effusion areas. The Kaggle dataset for diabetic retinopathy is used for the experiments, and the outcomes are assessed taking into account the mean and standard deviation of the retrieved features. The first phase is pre-processing, and the second is image segmentation (Sisodia *et al*. [7]). The SVM classifier is then used to classify a selection of images several classification methods, including fuzzy, c-means clustering, SVM, neural networks, PCA, and straightforward Bayesian methods. Here, we achieve varied feature accuracy of 80-90%. The work of Mobeen-ur-Rehman *et al*. [4] contains two contributions. We first suggested a neural network architecture specifically designed for the task of classifying images in diabetic retinopathy. Compared to conventional feature extraction-based techniques like sparse representation, it performs better classifiers, *Support Vector Machines* (SVM), *K-Nearest Neighbour* (KNN) methods, and *Linear Discriminant Analysis* (LDA).

This study employs the data augmentation methodology, which allows the dataset to be reused by flipping, rotating, and cropping the photos that are already present in the dataset. CNN classification is the following stage. A 95% accuracy rate was achieved after processing 200 test photos and about 800 training images (Roychowdhury *et al*. [5], and Rufaida and Fanany [6]).

Here, characteristics like hard exudates, red lesions, and blood vessels displayed accuracy of 89-95%. Three CNN architectures — AlexNet, VGG16, and InceptionNet V3 — were examined in this research (Verma *et al*. [9]), and a few setups were created to leverage these CNNs for DR set categorization. A total of 166 fundus samples from the public Kaggle dataset, which was provided by EyePACS and was carefully chosen by specialists, were used in this experiment. This dataset was balanced throughout all phases of DR. Based on inspection photographs of Andonová *et al*. [1], it was completed. Following a 5-fold cross-validation, V3 is 37.43%, 50.3%, and 63.23%. We found an improvement of 11% for the flat CNN compared to GoogleNet after applying the correction factor to VGG16. To hasten the convergence of the model, optimization approaches such stochastic gradient descent with momentum were applied. Transfer learning started with database optimization. The photographs were distorted and lost fidelity when they were resized to maintain uniformity across the database. Three CNN architectures — AlexNet, VGG16, and InceptionNet V3 — were examined in this research, and a few setups were created to leverage these CNNs for DR set categorization. A total of 166 fundus samples from the public Kaggle dataset, which was provided by EyePACS and was carefully chosen by specialists, were used in this experiment. This dataset was fairly balanced throughout all phases of DR. Based on inspection photographs, it was completed. Following a 5-fold cross-validation, V3 is 37.43%, 50.3%, and 63.23%. We found an improvement of 11% for the flat CNN compared to GoogleNet after applying the correction factor to VGG16. To hasten the convergence of the model, optimization approaches such stochastic gradient descent with momentum were applied. Transfer learning started with database optimization. The photographs were distorted and lost fidelity when they were resized to maintain uniformity across the database. Fundus pictures from the MESSIDOR database, which had already been pre-processed, transformed, and normalized, were identified by Kanungo *et al*. [3] as bright lesions (soft green and hard blue effusions).

We were able to discriminate between healthy and sick eyes as a result. In order to improve the contrast between the picture and the background, additional techniques such as *Adaptive Histogram Equalization* (AHE), addition of Gaussian noise, transformation of grayscale images, green channelling, and image flipping were used. The CNN with four layers has been defined. Max pooling was employed for down sampling. By joining the four CNN, two completely interconnected layers were added. The accuracy formula was used to confirm the method's effectiveness. Each subset is tested once after five iterations, while the other four sets are used for convolutional training. This approach had an accuracy rate of 82, and for mixed data, this precision is the best. Working with isolated areas of a picture rather than the complete image was the main focus of this research paper.

A three-step technique is suggested by Yu *et al*. [16] to automatically identify and assess the severity of diabetic retinopathy from fundus images. The green plane (green) in each fundus image in JPEG format is used to extract information.

To find regions including the optic disc and blood vessels as the image matching background, the minimum intensity and maximum solidity methods were used in conjunction with image segmentation, blood vessel recognition, and one-step categorization of bright red lesions growth. Green's Build a comprehensive three-stage automated system in stage two that can identify retinopathy lesions and assign each fundus image a DR severity score. This work, at level 2 presented a brand-new two-level hierarchical classification method, at level 3, examined the value of feature ranking and feature reduction, and at level 3, combined the number of lesions, suggested a mechanism for automatically detecting DR on another public record (Yu *et al*. [16]). In level 3, a combination function is used to count and combine the quantity of red and bright lesions. Messidor and Diaretdb1 are used with this dataset. Extracted are numerous haemorrhages, micro aneurysms, hard exudates, and cotton wool patches. The counts for the aforementioned features obtained per image using (CWS) are calculated once the area corresponding to the retinopathy lesion is acquired. The top 30 features were chosen utilizing all 89 photos from the DIARETDB1 dataset, which offered a better feature selection technique than the 28-image training set. One study's automated method for diabetic retinopathy saved retinal fundus photos with dimensions of 1500×1152 pixels and pixel lengths of 24 bits in JPG format. Exudates, micro aneurysms, hemorrhages, blood vessels, and optical discs are characteristics that are used as input datasets and processing procedures.

SVM and ANN are then employed for classification (Yu *et al*. [16]). The MESSIDOR, Diabetes DBI data set was used in this study. The image is resized during the preparation stage, followed by challenges with colour space conversion, image restoration, and image enhancement. The input colour fundus image is converted to HIS by colour space conversion. Histogram Equalization and Contrast Enhancement follow Pre-processing Module. To discover micro aneurysms and hard exudates, feature extraction performs morphological operations like erosion, dilation, etc. To remove optical discs, a clever edge detecting technique is employed. To create the bright papillae in this instance, a mask picture is created, and the mask image is then subtracted from the edge detection image. All features are computed and sent to SVM and ANN classifiers for classification. After training and testing using a given dataset with a 70 : 30 ratio. We can see that SVM produces superior outcomes to ANN. When the SVM classifier reaches 85.60% accuracy and the ANN classifier achieves 55.17% accuracy, accuracy percentages are presented. The study by Xu *et al*. [13], a convolutional neuron was proposed (CNN) to detect diabetic retinopathy. The classification model in the figure includes transfer learning and hyper parameter adjustment. These AlexNet, VggNet-s, GoogleNet, and ResNet models are examined and put to the test here for DR picture categorization. The amount of training instances for the best testing and data normalization for noise reduction is increased by this classification analysis of pre-processing activities like data augmentation. In this study, the model was evaluated for sensitivity, specificity, accuracy, *Receiver Operating Characteristic* (ROC), and *Area Under the Curve* (AUC). Kaggle is the dataset in use here. On a scale of

0 – no DR, 1 – mild, 2 – moderate, 3 – severe, and 4 – proliferative DR, markers are provided by DR in each photograph. The following describes the transfer learning experimental setup: The original fundus image data is expanded 20 times, and after 30 training iterations, the learning rate is linearly varied between stochastic gradient descent and [0.0 0 01-0.1]. An improved procedure for updating weight values. The result is computed using 5-fold cross-validation. His VggNet model classification accuracy in experiments is 95.68%. The accuracy of the various models is 90.07% for AlexNet, 92.3% for GoogleNet, and 93.03% for ResNet (Szegedy *et al*. [8]).

3. Proposed Methodology

Step 1: Datasets on Kaggle in Mesidor, various datasets are provided without charge. Colour images ranging from hundreds to thousands of pixels wide and tall make up both types of data. Select the one that best fits your calculations based on your computer's processing capacity and the type of image it is based on.

Step 2: Research CNN architectures — Different CNN designs should be compared and analysed. We apply 22 levels of precision to our focus on GoogleNet. Pooling comes after the procedure. Up until a dense exit node is reached, the network is flattened to a single dimension. The model is shielded against over fitting and mistakes by regularization.

Step 3: Pre-processing — This stage entails cropping the image. Otsu's technique, for instance, isolates a circular colour image from the retina. By removing the minimum pixel intensity from each colour image, images are normalized. The picture's dimensions are changed. Edge detection is carried out after filters have been applied. This makes the eye image more distinct from the backdrop image. To increase real-time network location capabilities and decrease over fitting, data augmentation is frequently used.

Step 4: Model training and testing — The data obtained after the calculations are further split for cross-validation. Below are some common techniques used in resumes.

- *Training a test split approach*: This method involves randomly dividing the total amount of data into training and testing datasets. After that, use the training set for model training, and the test set for validation. The ideal data split is 70 : 30 or 80 : 20.

  When there is a lack of data, this method can be highly biased because it obscures information about the data that was not used for training. This method is acceptable if the data set is very large and the distribution of the test and training patterns is the same.
- *K-folds cross-validation*: K-fold is a general, simple model that, in comparison to other approaches, typically produces a less biased model. This is since both the training and test sets can contain all of the original dataset's observations. When you have little input data, this is one of the better strategies.

Step 5: Accuracy check — The pertinent accuracy parameters are chosen based on the size and dispersion of the data.

The proportion of accurate predictions to all input samples is known as accuracy.

For diabetic retinopathy, deep convolutional neural networks can be used. The biological processes that under pin CNN. Images may be analysed using convolutional neural networks to detect patterns (Wan *et al*. [10]).

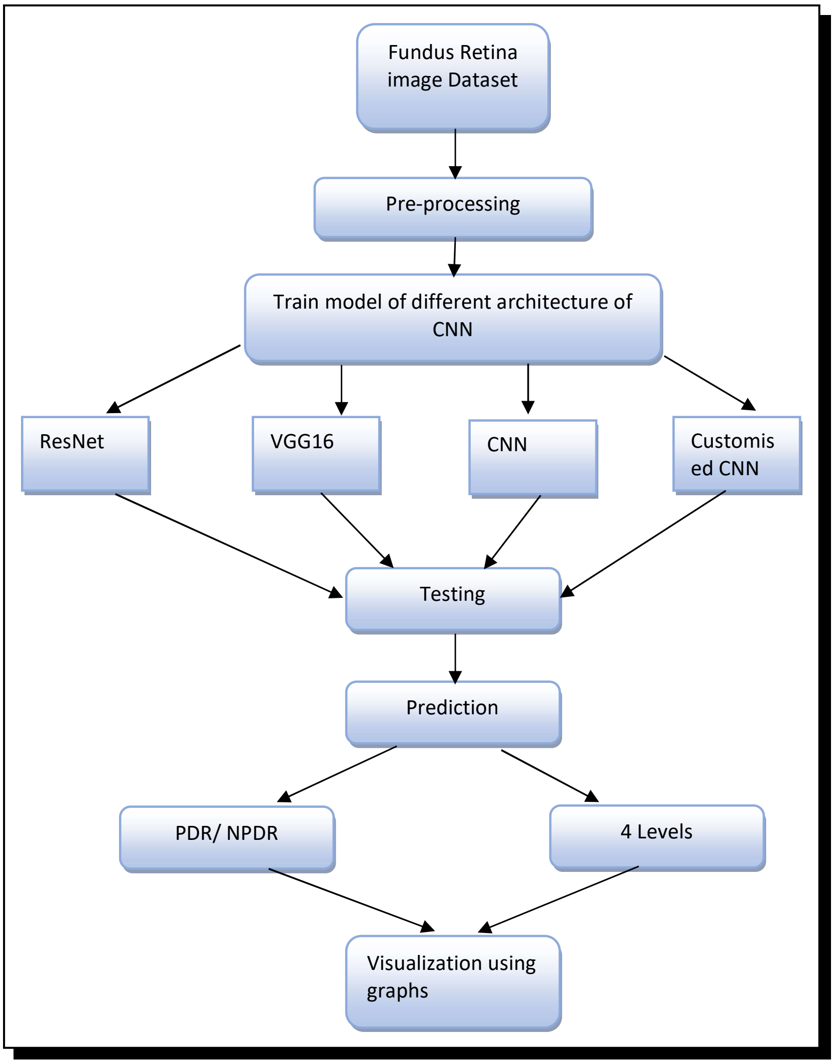

Figure 1. Flow diagram

3.1 Feature Extraction

3.1.1 Convolutional Layer

Convolutions search for patterns by applying filters to the image as they are processed. Filtering is the purpose of a convolutional layer. We effectively search for patterns in each area of a picture as we walk over it. The results produced by the convolution are multiplied by filters, which are stacks of weights expressed as a vector. These weights alter during image training, so when it comes time to analyse an image, these weights yield high values. If it believes it is spotting a trend it has seen previously, values. The network can estimate the content of an image thanks to the combinations of high weights from several filters (Wang and Yang [1]).

[1]Z. Wang and J. Yang, Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation, arXiv:1703.10757 [cs.CV], (2017), 9 pages, DOI: 10.48550/arXiv.1703.10757.

If the convolution feature is present at the specified point, the convolution operation generates a high value, otherwise, it outputs a low value. More specifically, we multiply each kernel cell value by its matching image pixel value, which overlaps the kernel cell, element by element, at a certain place in the convolution kernel, and then we add the results. The precise value is determined using the following formula:

$$h_{i,j} = \sum_{k=1}^{m} \sum_{l=1}^{m} w_{k,l} x_{i+k-1,j+l-1}, \tag{3.1}$$

where m is the width and height of the kernel, h is output of the convolution, x is input, and w is convolution kernel.

3.1.2 Pooling Layer

Pooling layer comes after convolutional layer. It divides the input image into a number of non-overlapping rectangles and outputs a value for each of these sub-regions. The assumption is that a feature's approximate placement in relation to other features is more significant than its precise location. Max-pooling and average pooling are the two primary pooling layers:

- *Max-Pooling* — It produces the sub-highest region's value.
- *Average-Pooling* — It produces the sub-average region's value.

The depth is not reduced by the pooling layer, only the spatial dimensions are

$$h_{i,j} = \max\{x_{i+k-1,j+l+1} \ \forall\ 1 \le k \le m \text{ and } 1 \le l \le m\}. \tag{3.2}$$

3.1.3 Activation Layer

A function that compresses a value into a range is passed a value. ReLU activation is the most used activation method. It accepts the input "x" and either returns "x" if it is positive or "0" otherwise. ReLU function is utilized since it is economical to execute. For all positive values, ReLU is linear, while for all negative values, it is zero. As there is no difficult arithmetic involved, computation costs are low. The training and running of the model can be sped up as a result. It is only lightly engaged. Any given unit may not activate at all because ReLU is zero for all negative inputs.

3.2 Classification

Fully connected network of CNN is used for the classification:

3.2.1 Fully Connected Input Layer (Flatten)

This layer "flattens" the output from earlier layers so that it can be used as an input for the subsequent layer.

3.2.2 First Fully Connected Layer

This layer takes inputs from the feature analysis and applies weights to predict correct label.

3.2.3 Fully Connected Output Layer

This layer gives a fully connected layer's purpose is to use the pooling process's outcomes to label the image by classifying it. To select the weights that are the most correct, this layer follows its own procedure. Weights are assigned to each neuron based on which label is the most appropriate. The classification choice is then made once each label has received a "vote" from the neurons.

4. Dataset and Results

We have used Kaggle image dataset of retina where we took dataset for 300 images for the first time and then 900 images and lastly 3600 images. Table 1 shows comparison between different CNN architectures.

Table 1. Evaluation matrices comparison with and without pre-processing

Accuracy	CNN architecture		
	Custom CNN	DenseNet	Resnet
With pre-processing	70.30%	86.88%	56.62%
Without pre-processing	69.14%	82.46%	48.97%

From Table 1, we get to know that after pre-processing, the accuracy definitely gets better.

Table 2. Evaluation matrices comparison with number of epochs

Epoch	Custom CNN	Densenet
1	70%	80%
5	53.3%	85%
10	54.8%	89%

The accuracy of custom CNN rapidly declines with the number of epoch sizes, as shown in Table 2. The cause of this is over fitting. However, when compared to other architectures, DenseNet performs well and uses a lot of resources. In order to acquire accurate findings, a large number of training photos were provided.

Table 3. Comparison based on database size and 4 level classifications

Number of images	Accuracy	Architecture
3600	89.10%	DenseNet
900	70.30%	Custom CNN
100	75%	Custom CNN

Table 3 shows that the performance of a custom CNN decreases as the quantity of the dataset increases. Over fitting is to blame for this. The model that was initially trained on 100 photos offers good fitting results, but poor testing accuracy. The accuracy was found to be low as the number of samples increased. This was caused by the dataset's imbalance, where the majority of the photos are from stage 0. As a result, custom CNN's accuracy decreased.

5. Conclusion

After reading numerous publications, we can say that deep CNN outperforms other ML techniques in terms of outcomes and efficiency. Using binary classifiers, we train CNNs to perform at the cutting edge, model performance declines as the number of classes increases. The GoogleNet architecture offers the best accuracy, and our architecture's dropout approach can cut down on computation. Pre-processing methods will depend on the dataset, and the optimal feature extraction can be achieved by using better filters. In CNN, we can employ many layers to improve accuracy. Convolutional and pooling layers can be used for greater accuracy, and the ReLU function can be used as an activation function, according to our suggested methodology. Effectively classifies layers that are fully connected. This study can be expanded to include effective multiclass classification, where it is possible to obtain the stage of diabetic retinopathy. By pre-processing, expanding the epochs, and increasing the number of images, we can improve accuracy by 5%.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

All the authors contributed significantly in writing this article. The authors read and approved the final manuscript.

References

[1] M. Andonová, J. Pavlovičová, S. Kajan, M. Oravec and V. Kurilová, Diabetic retinopathy screening based on CNN, in: *2017 International Symposium ELMAR*, Zadar, Croatia (2017), pp. 51 – 54, DOI: 10.23919/ELMAR.2017.8124433.

[2] Y. Hatanaka, K. Ogohara, W. Sunayama, M. Miyashita, C. Muramatsu and H. Fujita, Automatic microaneurysms detection on retinal images using deep convolution neural network, in: *2018 International Workshop on Advanced Image Technology* (IWAIT), Chiang Mai, Thailand (2018), pp. 1 – 2, DOI: 10.1109/IWAIT.2018.8369794.

[3] Y. S. Kanungo, B. Srinivasan and S. Choudhary, Detecting diabetic retinopathy using deep learning, in: *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology* (RTEICT), Bangalore, India (2017), pp. 801 – 804, DOI: 10.1109/RTEICT.2017.8256708.

[4] Mobeen-ur-Rehman, S. H. Khan, Z. Abbas and S. M. D. Rizvi, Classification of diabetic retinopathy images based on customised CNN architecture, in: *2019 Amity International Conference on Artificial Intelligence* (AICAI), Dubai, United Arab Emirates (2019), pp. 244 – 248, DOI: 10.1109/AICAI.2019.8701231.

[5] S. Roychowdhury, D. D. Koozekanani and K. K. Parhi, DREAM: Diabetic Retinopathy analysis using machine learning, *IEEE Journal of Biomedical and Health Informatics* **18**(5) (2014), 1717 – 1728, DOI: 10.1109/JBHI.2013.2294635.

[6] S. I. Rufaida and M. I. Fanany, Residual convolutional neural network for diabetic retinopathy, in: *2017 International Conference on Advanced Computer Science and Information Systems* (ICACSIS), Bali, Indonesia (2017), pp. 367 – 374, DOI: 10.1109/ICACSIS.2017.8355060.

[7] D. S. Sisodia, S. Nair and P. Khobragade, Diabetic retinal fundus images: Preprocessing and feature extraction for early detection of diabetic retinopathy, *Biomedical & Pharmacology Journal* **10**(2), 615 – 626, DOI: 10.13005/bpj/1148.

[8] C. Szegedy, S. Ioffe, V. Vanhoucke and A. Alemi, Inception-v4, Inception-ResNet and the impact of residual connections on learning, *Proceedings of the AAAI Conference on Artificial Intelligence* **31**(1) (2017), 4278 – 4284, DOI: 10.1609/aaai.v31i1.11231.

[9] K. Verma, P. Deep and A. G. Ramakrishnan, Detection and classification of diabetic retinopathy using retinal images, in: *2011 Annual IEEE India Conference*, Hyderabad, India (2011), pp. 1 – 6, DOI: 10.1109/INDCON.2011.6139346.

[10] S. Wan, Y. Liang and Y. Zhang, Deep convolutional neural networks for diabetic retinopathy detection by image classification, *Computers & Electrical Engineering* **72** (2018), 274 – 282, DOI: 10.1016/j.compeleceng.2018.07.042.

[11] Y.-X. Wang, D. Ramanan and M. Hebert, Growing a brain: Fine-tuning by increasing model capacity, in: *2017 IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), Honolulu, HI, USA (2017), pp. 3029 – 3038, DOI: 10.1109/CVPR.2017.323.

[12] X. Wang, Y. Lu, Y. Wang and W.-B. Chen, Diabetic retinopathy stage classification using convolutional neural networks, in: *2018 IEEE International Conference on Information Reuse and Integration* (IRI), Salt Lake City, UT, USA (2018), pp. 465 – 471, DOI: 10.1109/IRI.2018.00074.

[13] K. Xu, D. Feng and H. Mi, Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image, *Molecules* **22**(12), 2054, DOI: 10.3390/molecules22122054.

[14] 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: *Medical Image Computing and Computer Assisted Intervention* (MICCAI 2017), M. Descoteaux, L. Maier-Hein, A. Franz, P. Jannin, D. Collins and S. Duchesne (editors), *Lecture Notes in Computer Science*, Vol. **10435**, Springer, Cham. (2017), DOI: 10.1007/978-3-319-66179-7_61.

[15] S. Yu, D. Xiao and Y. Kanagasingam, Exudate detection for diabetic retinopathy with convolutional neural networks, in: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (EMBC), Jeju, South Korea (2017), pp. 1744 – 1747, DOI: 10.1109/EMBC.2017.8037180.

[16] F. Yu, J. Sun, A. Li, J. Cheng, C. Wan and J. Liu, Image quality classification for DR screening using deep learning, in: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (EMBC), Jeju, South Korea (2017), pp. 664 – 667, DOI: 10.1109/EMBC.2017.8036912.