

Application of convolution neural networks in eye fundus image analysis

N Y Ilyasova^{1,2}, A S Shirokanev^{1,2} and I A Klimov¹

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

²Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

e-mail: ilyasova.nata@gmail.com, alexandrshirokanev@gmail.com,
klimov.ilya.05@gmail.com

Abstract. In this work, we proposed a new approach to analyzing eye fundus images that relies upon the use of a convolutional neural network (CNN). The CNN architecture was constructed, followed by network learning on a balanced dataset composed of four classes of images, composed of thick and thin blood vessels, healthy areas, and exudate areas. The learning was conducted on 12x12 images because an experimental study showed them to be optimal for the purpose. The test error was no higher than 4% for all sizes of the samples. Segmentation of eye fundus images was performed using the CNN. Considering that exudates are a primary target of laser coagulation surgery, the segmentation error was calculated on the exudate class, amounting to 5%. In the course of this research, the HSL color system was found to be most informative, using which the segmentation error was reduced to 3%.

1. Introduction

A severe consequence of diabetic retinopathy (DRP) is vision loss. DRP affects all parts of the retina, leading to macular edema, which in turn causes fast worsening of eyesight [1]. The accurate and early diagnosis alongside an adequate treatment can prevent the vision loss in more than 50 % of cases [2-3]. There are a number of approaches to treating DRP, one of which involves laser photocoagulation [4]. During this procedure, a number of retina areas where edema occurs are cauterized with a laser. The procedure is conducted via coagulating near-edema zones. The development of diagnostic systems enabling an automatic identification of the edema zone is currently a relevant task [5]. For the laser coagulation procedure to be automated [6, 7], objects in the eye fundus image need to be classified [8-11], which can be done in a number of ways [12].

Convolutional neural networks are the choice of preference when dealing with object classification [13]. Such is the conclusion members of the research community involved in medical image analysis have come to in the course of their research. Techniques for medical data analysis are often among research topics at international conferences and symposia. May 2006 has seen the publication of the first issue of IEEE Transaction on. The first detailed review to be published on the use of deep learning for medical image analysis appeared in 2017 [5]. Nowadays an active trend for the development of digital medicine is seen.

In Ref. [14], a classification model based on a convolutional neural network was used for diagnosing the H. Pylori infection. In the work, an architecture specially tailored for a particular task was utilized. The authors came to the conclusion that the particular disease was possible to diagnose based on endoscopic images obtained using CNN. In Ref. [15], diagnosing an early-stage hypertension retinopathy was discussed. The onset of the disease in the eye retina is prompted by blood hypertension. The classifier proposed in Ref. [15] offered a 98.6 percent accuracy. In Ref. [16], a toolkit was developed for the automated analysis of psoriasis-affected skin biopsy images, which is of considerable significance in clinical treatment. The paper is a pioneering attempt into automatic segmentation of psoriasis-affected skin biopsy images. The study resulted in a practical system based on the machine analysis. CNN training on a prepared dataset was demonstrated, intended for further analysis of input images.

In this work, we study a class of eye fundus images with pathological changes that can be found at different stages of the disease. Manifestations of the diabetic retinopathy include exudates, which cause the retina thickening (Figure 1). In general, the image of an eye fundus with pathology contains four classes of objects, such as thick and thin blood vessels, healthy areas and exudate zones.



Figure 1. Illustration of diagnostic eye fundus images: healthy eye (left) and affected eye (right).

2. Training the convolutional neural network

The initial data for analysis contained 11 training datasets of various size. All datasets were balanced and in total contained 534 images. For the purposes of the present work, the CNN training was conducted on four above-described classes of eye fundus images. The initial dataset consisted of 75 percent of training images and 25 percent of test images. To prevent overtraining, a control dataset was also used. A 3x3 convolution kernel was chosen because it is optimal for 12x12 images. The CNN architecture was constructed empirically so as the required accuracy of no less than 96 % is ensured. Table 1 gives architecture of the empirically constructed convolutional neural network. With this architecture, a recognition accuracy of 99.3% was attained, which is the best recognition result for the four above-mentioned classes of images. Figure 2 depicts a plot for training the CNN model in each training epoch.

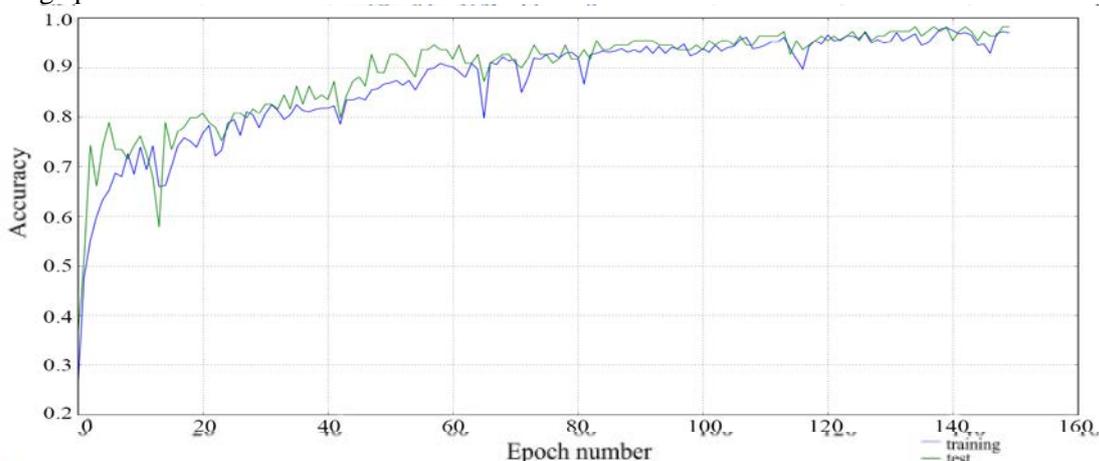


Figure 2. The dependency of learning accuracy on the number of epochs.

Table 1. Architecture of the convolutional neural network.

| Layer number | Layers | Parameters | Layer number | Layers | Parameters |
|--------------|---------------|----------------|--------------|-----------------|----------------|
| 1 | Convolutional | 300 neurons | 3 | Activation | Function: RELU |
| 1 | Activation | Function: RELU | 4 | Convolutional | 150 neurons |
| 2 | Convolutional | 300 neurons | 4 | Activation | Function: RELU |
| 2 | Activation | Function: RELU | 4 | MaxPooling | Size: 2×2 |
| 2 | Dropout | 0.5 | 4 | Dropout | 0.5 |
| 2 | MaxPooling | Size: 2×2 | 5 | Fully-connected | 4 |
| 3 | Convolutional | 150 neurons | 5 | Activation | softmax |

To attain a recognition certainty of 95 %, the CNN was put through 120 training runs on the initial images of all sizes. Figure 3 shows an average training result for each image size.

The results in Figure 2 show that the highest classification accuracy is attained for 12x12 images.

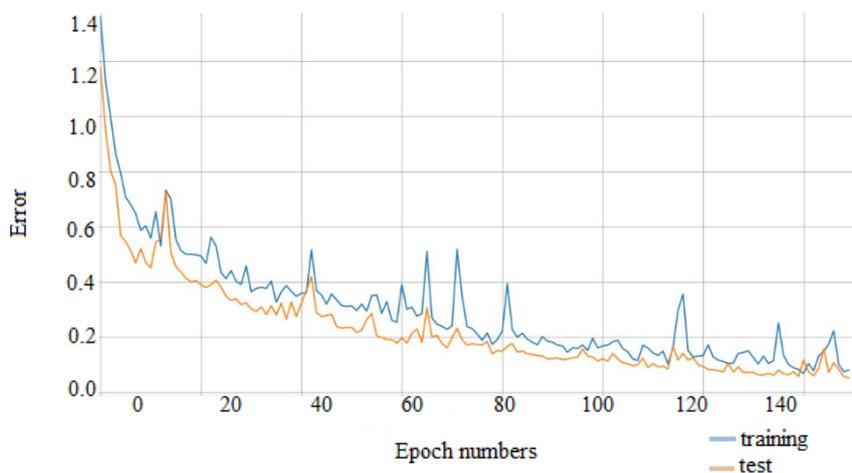


Figure 3. The dependency of learning error on the number of epochs.

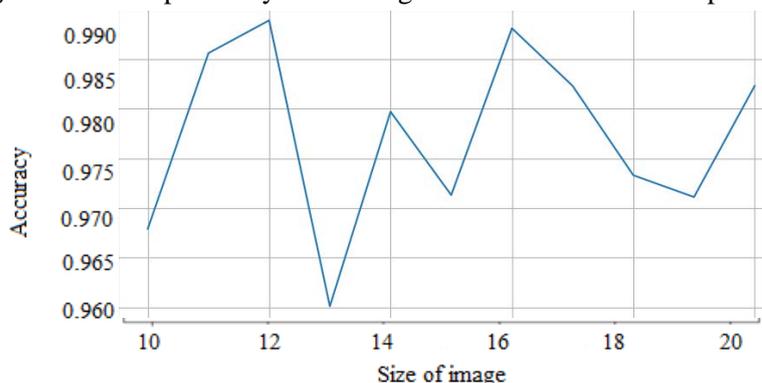


Figure 4. The dependence of accuracy on the size of images in the training set.

3. Experimental study

For the experiments, datasets were formed containing four above-described classes of 12x12 images, using which the best result of CNN testing is achieved (Figure 3). In this study, the segmentation of eye fundus images was conducted via deep learning. Shown in Fig. 4b is the result of CNN-aided image segmentation. With a view of estimating the CNN-aided segmentation error, a manual segmentation by an expert ophthalmologist was introduced as a reference image (Figure 4c). The study was conducted on the exudates class, which had been singled out into a separate image (Figure 4d). The error of CNN-aided segmentation of the said exudate areas was calculated relative to the expert estimate. The result of comparison of the exudation areas highlighted by CNN and the expert is shown in Table 2. Using the data from Table 2, a CNN-aided segmentation error for the exudates was

defined as $E = (k + t) / NM$ and amounted to 7% (where $N \times M$ is the image size, k is the number of expert-highlighted pixels that CNN failed to recognize as exudates, t is the number of exudate pixels recognized by CNN but missing from the expert's image). The error of first kind, defined as $E_1 = l / F$, where l is the number of falsely recognized exudates classes and F is the total number of exudate-containing pixels in the expert's image, amounted to 5%.

In the process of exudates area identification, color plays a key role. The segmentation error can be significantly reduced by operating in particular color spaces. It has been established [17] that color models YUV, RGB and HSL are most close to color perception of the human eye. However, the models RGB and YUV have a number of hardware limitations with certain video-systems. In further research, we used the HSL color model as the one that most adequately renders the color characteristics of blood vessels and exudates. Figure 5 shows pathological areas highlighted by the expert in different HSL color channels.

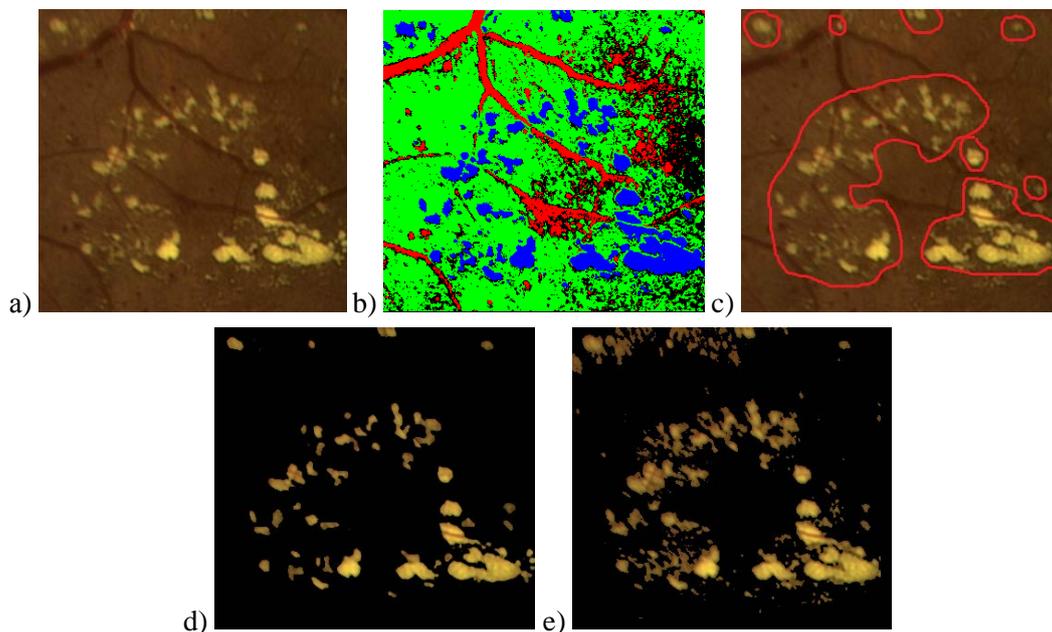


Figure 4. (a) The original eye fundus image; (b) four classes of objects highlighted in the image using CNN; (c) exudates areas manually outlined by an expert; exudates areas highlighted by (d) an expert and (e) using the CNN technique.

Table 2. Percentage of exudates areas in the image.

| Areas | Percentage of exudates area, % |
|--|--------------------------------|
| Exudates area in the expert's image | 9 |
| Exudates area in the CNN-aided image | 15 |
| Total exudates area | 95.6 |
| Expert's exudates areas omitted by CNN | 0.4 |
| CNN-highlighted exudates areas missing in the expert's image | 6 |

Veracity of CNN-aided exudate highlighting has been confirmed by comparison of histograms of CNN-aided and expert's images (Figure 4), which were superimposed for each corresponding channel of HSL color system, with the expert-based histograms marked as green bars, and the CNN-based histograms marked red (Figure 6). The expert-based histograms define an interval of values for the affected fundus areas. From the histograms, the CNN-aided interval of exudates area is seen to be narrower than that obtained based on expert's estimates. The histogram regions corresponding to the false CNN-aided classification are within intervals shown by rectangles (Figure 6). Table 3 gives segmentation errors calculated for each channel of the HSL color model. The data in Table 3 suggest that the H channel is the most informative channel with the least segmentation error.

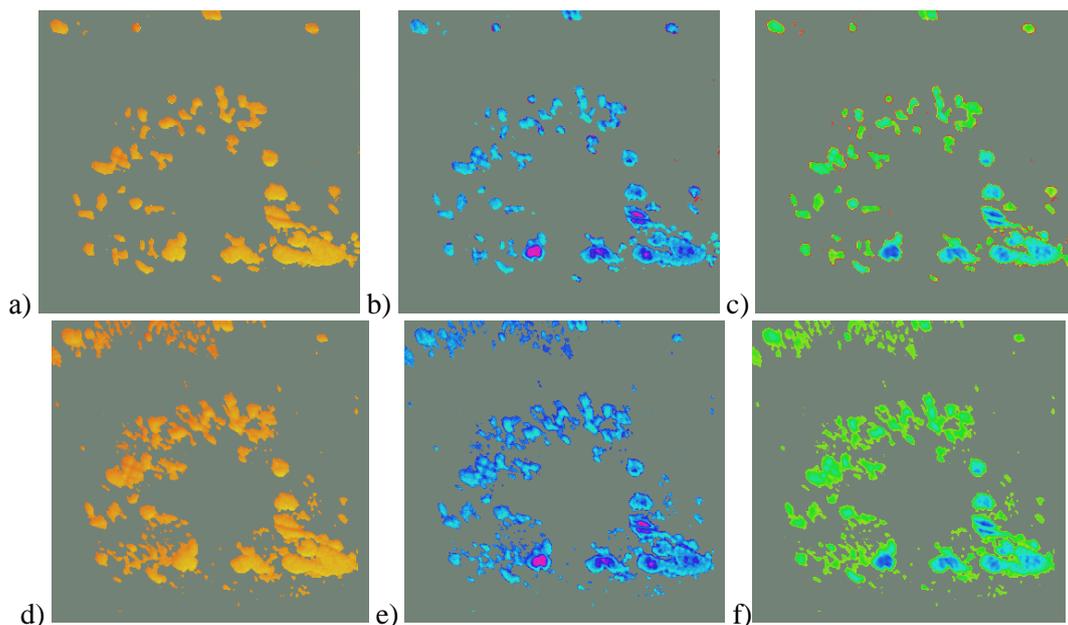


Figure 5. Expert-highlighted affected fundus areas for the HSL color model for the channel (a) H, (b) S, and (c) L; CNN-highlighted affected fundus areas for the HSL color model for the channel (d) H, (e) S, and (f) L.

Table 3. Segmentation error for each channel.

| Channel | H | S | L |
|--|---|---|-----|
| Segmentation error for the exudates class, % | 3 | 9 | 9.8 |

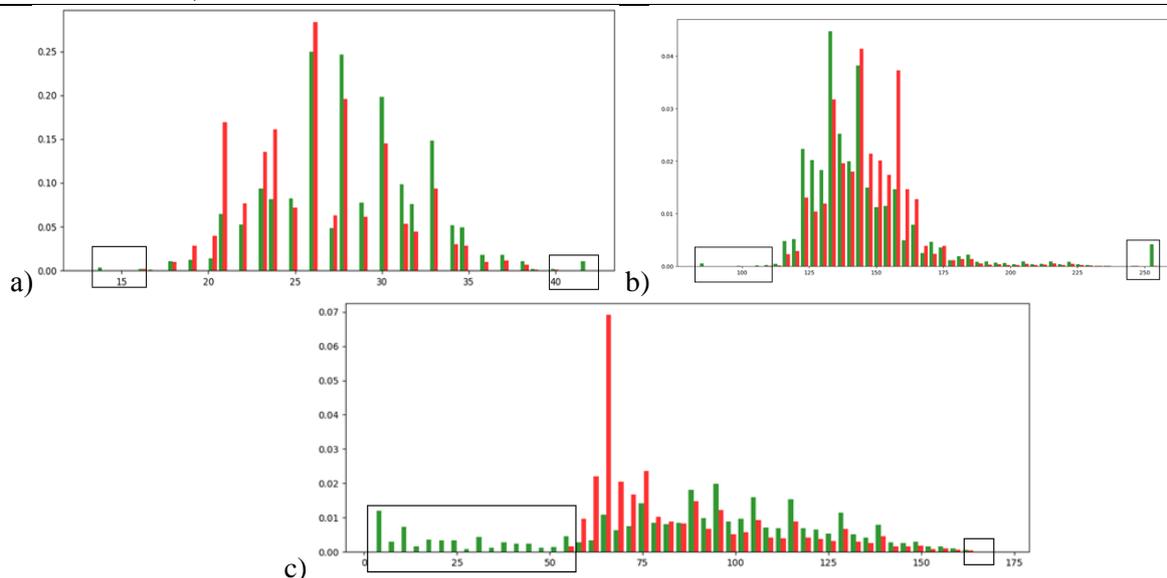


Figure 6. Image histograms obtained by using an expert opinion and CNN: a) H channel, b) S channel, and c) L channel.

4. Conclusion

In this work, a convolutional neural network (CNN) has been applied to the analysis of an eye fundus image. CNN architecture has been constructed, allowing a testing error of no more than 4% to be attained. Based on a 3x3 convolution kernel, CNN training was conducted on 12x12 images, thus enabling the best result of CNN testing to be achieved. CNN-aided segmentation of the input image conducted in this work has shown the CNN to be capable of identifying all training dataset classes

with high accuracy. The segmentation error was calculated on the exudates class, which is key for laser coagulation surgery. The segmentation error on the exudates class was 7 %, with the error of first kind being 5 %.

In the study, we utilized the HSL color model because it renders color characteristics of eye blood vessels and exudates most adequately. We have demonstrated the *H* channel to be most informative, with the segmentation error amounting to 3 %.

5. References

- [1] Shadrichev F E Diabetic retinopathy *Territorial Diabetes Center* 8-11
- [2] Ilyasova N Yu 2014 Evaluation of geometric features of the spatial structure of blood vessels *Computer Optics* **38(3)** 529-538
- [3] Khorin P A, Ilyasova N Yu and Paringer R A 2018 Informative feature selection based on the Zernike polynomial coefficients for various pathologies of the human eye cornea *Computer Optics* **42(1)** 159-166 DOI: 10.18287/2412-6179-2018-42-1-159-166
- [4] Astakhov Y S, Shadrichev F E, Krasavira M I and Grigotyeva N N 2009 Modern approaches to the treatment of diabetic macular edema *Ophthalmologic sheets* **4** 59-69
- [5] Litjens G A 2017 Survey on deep learning in medical image analysis *Medical Image Analysis* **42** 60-88
- [6] Ilyasova N, Kirsh D, Paringer R, Kupriyanov A and Shirokanev A 2017 Coagulate map formation algorithms for laser eye treatment *IEEE Xplore* 1-5
- [7] Shirokanev A S, Kirsh D V, Ilyasova N Yu and Kupriyanov A V 2018 Investigation of algorithms for coagulate arrangement in fundus images *Computer Optics* **42(4)** 712-721 DOI: 10.18287/2412-6179-2018-42-4-712-721
- [8] Ilyasova N, Kirsh D, Paringer R and Kupriyanov A 2017 Intelligent feature selection technique for segmentation of fundus images *Seventh International Conference on Innovative Computing Technology (INTECH)* 138-143
- [9] Shirokanev A S, Ilyasova N Yu and Paringer R A 2017 A smart feature selection technique for object localization in ocular fundus images with the aid of color subspaces *Procedia Engineering* **201** 736-745
- [10] Ilyasova N, Paringer R and Kupriyanov A 2016 Regions of interest in a fundus image selection technique using the discriminative analysis methods *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* **9972** 408-417
- [11] Ilyasova N Yu, Kupriyanov A V and Paringer R A 2014 Formation of features for improving the quality of medical diagnosis based on discriminant analysis methods *Computer Optics* **38(4)** 851-855
- [12] Ilyasova N Yu 2013 Methods for digital analysis of human vascular system. Literature review *Computer Optics* **37(4)** 517-541
- [13] Guido S, Andreas C 2017 Introduction to machine learning with python. *O'Reilly Media* 392
- [14] Shichijo S 2017 Application of Convolutional Neural Networks in the Diagnosis of Helicobacter pylori Infection Based on Endoscopic Images *The LANCET* **25** 106-111
- [15] Bambang K T 2017 The Classification of Hypertensive Retinopathy using Convolutional Neural Network *ICCSCI* 166-173
- [16] Anabik P 2018 Psoriasis skin biopsy image segmentation using deep convolutional neural network *Computer Methods and Programs in Biomedicine* **159** 59-69
- [17] Nikitaev V G 2004 Experimental study of color models in automated image analysis tasks *Scientific session MIFI* **1** 253-254

Acknowledgements

This work was financially supported by the Russian Foundation for Basic Research under grant # 19-29-01135, # 17-01-00972 and by the Ministry of Science and Higher Education within the State assignment to the FSRC "Crystallography and Photonics" RAS.