Retinal Image Analysis for Diabetic Retinopathy Grading: Preliminarily Results

Marina Lukashevich
Dept. of Electronic Computing Machines
Belarusian State University of Informatics and Radioelectronics
Minsk, Belarus
lukashevich@bsuir.by

Yuliya Golub
United Institute of Informatics Problems of NAS of Belarus
Minsk, Belarus
6423506@gmail.com

Valery Starovoitov
United Institute of Informatics Problems of NAS of Belarus
Minsk, Belarus
ORCID 0000-0001-7190-761X

Aun Irtaza
Dept. of Computer Science
University of Engineering and Technology
Taxila 47050, Pakistan
ORCID 0000-0001-7757-5839

Abstract. **Many eye diseases manifest themselves in the retina. Advances in Artificial intelligence (AI), especially in deep learning, improve pathological image analysis in routine clinical practice. We have developed a method for retina image analysis, based on image preprocessing stage and deep neural network as machine learning model. This is the preliminary results of big project for retina image analysis. The main focus was done for diabetic retina images. Also the scheme of new technology for retina image analysis was presented.**

Keywords: **diabetic retinopathy, retina image analysis, image preprocessing, machine learning, convolution neural network**

I. Introduction

There are many different eye diseases. The fund's image analysis is used to detect different stages of the 3 main diabetes-based eye diseases, i.e. Diabetic Retinopathy, Diabetic Macular Edema, and Glaucoma. In our research, we have considered Diabetic Retinopathy because it causes blindness (the most severe consequences). Diabetic retinopathy (DR), also known as diabetic eye disease, is a medical condition in which damage occurs to the retina due to diabetes mellitus. It is a leading cause of blindness. Diabetic retinopathy affects up to 80 percent of those who have had diabetes for 20 years or more. Diabetic retinopathy often has no early warning signs. Digital pathology nowadays plays an increasingly important role in basic, translational, and clinical pathology research and in routine clinical practice. Retinal (fundus) photography with manual interpretation is a widely accepted screening tool for diabetic retinopathy, with performance that can exceed that of in-person dilated eye examinations.

The clinical signs of DR include (1) multiple cotton wool spots (accumulations of axoplasmic debris within adjacent bundles of ganglion cell axons); (2) venous beading and/or looping; (3) microaneurysms (deep round and blot haemorrhages); (4) hard exudates (lipid deposits); and (5) intraretinal microvascular abnormalities (dilated preexisting capillaries). Artificial intelligence (AI), particularly machine learning (ML), has been widely applied to the pathological image analysis and has provided significant support for medical research and clinical practice. ML can infer image analysis rules from data representations and typically does not require manual algorithm adaptation to different data sets or images [1, 2].

II. Development of New Technology for Retina Image Analysis

A. The Scheme of New Technology

Development of new technology for retina image analysis is the main goal of our international scientific research. The plan for this issue have been developed. It includes the next stages:

1. *Collecting dataset of retina images* from public datasets. It is necessary to collect a variable dataset of retina images obtained in different conditions and using different equipment. Thanks to this it is possible to develop a new technology, but not a new method or approach suitable for limited amount of images.
2. *Quality image analysis*. This stage will allow to select images for the next stages of image processing and decision making. It will increase the quality of decision making in the framework of retina image analysis.

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

3. *Image preprocessing.* Low quality of retina images is the main reason of preprocessing stage. This stage can include different options. Main of them is image filtering, circle (mask) cropping, brightness correction, etc.

4. *Machine learning model development* for decision making (image classification). At this stage the main direction for our future research is a neural network approach (deep learning). And the promising approach is using the convolution neural network.

5. *Model evaluation* using different metrics. The vast majority of retina image dataset is unbalanced. It needs to find a suitable metric for unbalanced classes.

The scheme of new technology development for retina image analysis is presented on Fig. 1. We have marked those stages that were explored during the first preliminary stage of our project. Also the next stages and directions of scientific researches are presented in this scheme.

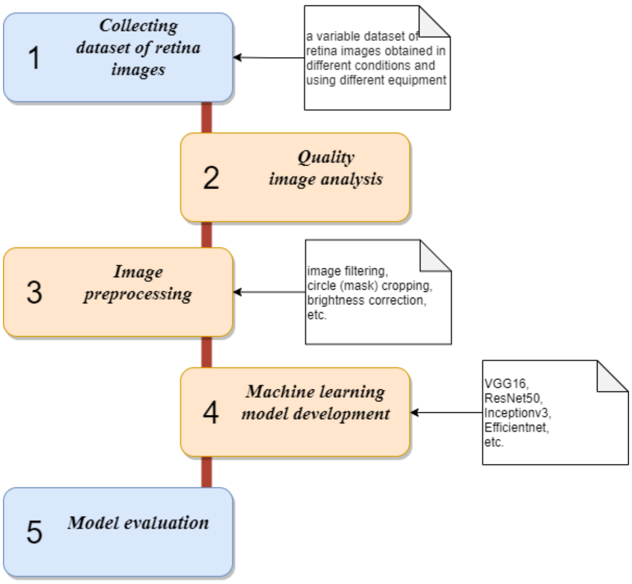

Fig. 1. The scheme of new technology development for retina image analysis

B. Quality Image Analysis

For quality retina image analysis 16 measures described in foreign and domestic literature were selected: BEGH, BISH, BREN, CMO, CURV, FUS, HELM, EBCM, KURT, LAPD, LAPL, LAPM, LOCC, LOEN, SHAR, WAVS [3]. 8 tests were carried out for the correspondence of the obtained values of local quality measures to the normal distribution. Experiments have shown that local estimates do not correspond to the normal distribution of the data. Therefore, Weibull distribution parameters were used to assess the quality of retinal images. The parameter scale (scale) of the FUS measure allows you to divide images into two classes - satisfactory and unsatisfactory quality. Fig. 2 shows examples of retinal images and quality values for each image.

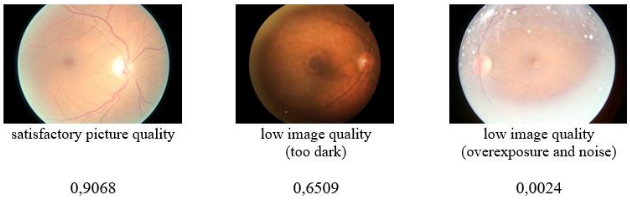

Fig. 2. The examples of retinal images and quality values for each image

III. Experimental Environment

The main goal of our research on this stage was to develop machine learning model for classifying samples from retina image dataset into 5 classes (0 – No DR, 1 – Mild, 2 – Moderate, 3 – Severe, 4 – Proliferative DR). There are two main approaches in image analysis and classification: traditional approach and deep learning approach. In the framework of traditional approach we perform feature extractions and decision making [4]. In deep leaning neural network extracts features and makes decision themselves. Deep neural networks consist of a large number of layers, have complex and difficult for interpretation. But it is a good variant for large datasets with variable samples. We have selected deep learning approach for this research, because it is promising direction for variable data.

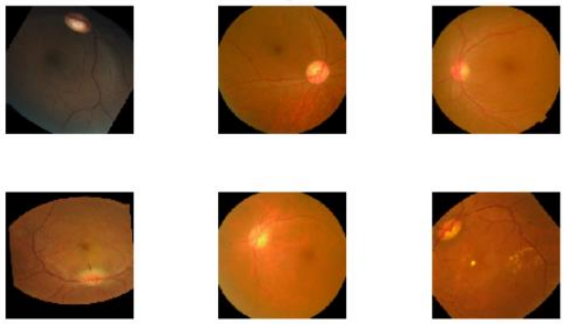

Fig. 3. Examples of the images from examineted dataset

We obtained images for training, validation and testing from a Kaggle competition [5]. It provides a large set of retina images, taken using funds photography under a variety of imaging conditions. A clinic has rated each image for the severity of diabetic retinopathy on a scale of 0 to 4 (0 – No DR, 1 – Mild, 2 – Moderate, 3 – Severe, 4 – Proliferative DR). Like any real-world data set, noise is presented in both the images and labels. Images may contain artifacts, be out of focus, underexposed, or overexposed. The images were gathered from multiple clinics using a variety of

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

cameras over an extended period of time, which will introduce further variation. The examples of the images from this dataset are shown on Fig. 3.

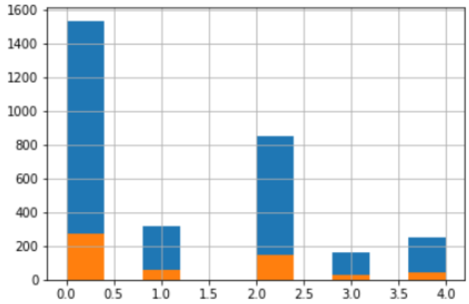

Fig. 4. The amount of samples from each classes (5 classes)

The dataset is highly imbalanced, with many samples for level 0, and very little for the rest of the levels. It was divided into two parts for training and testing machine learning model, Fig. 4.

IV. Methodology of Experiments

A. Retina Images Preprocessing

As we mentioned above retina images contain different types of noises and artefacts. That is why the first stage of the proposed method is image preprocessing. Improvement of the quality of input data can improve the performance of machine learning model. This stage includes reducing lighting-condition effects and cropping uninformative area. The registrations of digital retina images were conducted with many lighting conditions. Some images are very dark and difficult for visualization.

Image smoothing techniques help in reducing the noise. Using different image prospecting libraries, image smoothing (also called blurring) could be done in many ways. In this method, we have performed image smoothing using the Gaussian filter with sigma parameter equal 10. Gaussian filters have the properties of having no overshoot to a step function input while minimizing the rise and fall time. In terms of image processing, any sharp edges in images are smoothed while minimizing too much blur [6].

Cropping is a quite typical step for such kind of images. To solve our case, one method would be to look for rows and columns that have at least one pixel along rows and columns that is greater than some lower limit or threshold as a pixel value. So, if we are sure that the *black* areas are absolute zeros, we can set that threshold as *0*. Thus, if *img* represents the image data, we would have correspondingly two boolean arrays: *(img>tol).any(1)* and *(img>tol).any(0).*

The results of this stage are presented on Fig. 5 (several samples of images before preprocessing) and Fig. 6 (several samples of images after preprocessing).

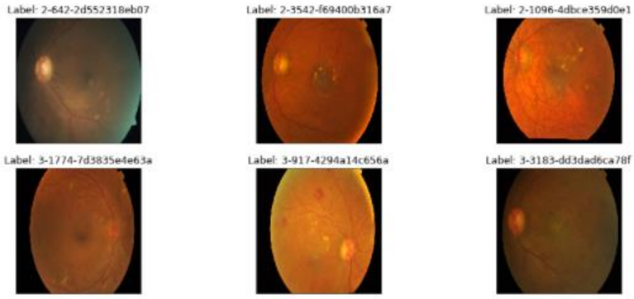

Fig. 5. Samples of retina images before preprocessing

B. Experimental Details

Colaboratory from Google and Keras framework were used for experiments. Colaboratory, or "Colab" for short, allows to write and execute Python in browser, with zero configuration required, free access to GPUs and easy sharing. It is a good choice for scientific recaches in the fild of AI [7]. Keras is a deep learning API written in Python, running on top of the machine learning platform TensorFlow. It was developed with a focus on enabling fast experimentation.

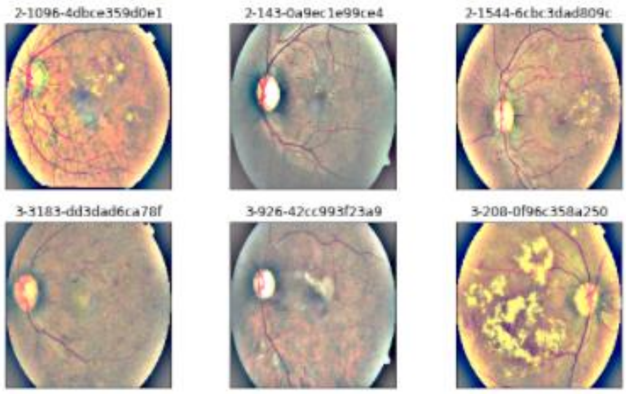

Fig. 6. Samples of retina images after preprocessing

C. Machine Learning Model Development and Evaluation

Deep learning approach has become more popular in digital image processing. There are a lot of promising model for classification tasks. Since AlexNet won the 2012 Image Net competition, convolution neural networks have become more accurate by going bigger. While the 2014 ImageNet winner GoogleNet achieves 74.8% top-1 accuracy with about 6.8M parameters, the 2017 ImageNet winner SENet (achieves 82.7% top-1 accuracy with 145M parameters. Recently, GPipe further pushes the state-of-the-art ImageNet top-1 validation accuracy to 84.3% using 557M parameters: it is so big that it can only be trained with a specialized pipeline parallelism library by partitioning the network and spreading each part to a different accelerator. There are several state-of-art CNN models suitable for our

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

task: VGG16 [8], ResNet50, Inceptionv3 and EfficientNet. We are planning to exam all of these architectures. In the framework of our preliminary stage we have built machine learning model using VGG16.

The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) is an annual computer vision competition [9]. Each year, teams competed on two tasks. The first is to detect objects within an image coming from 200 classes (object localization). The second is to classify images, each labeled with one of 1000 categories (image classification). This model won the 1st and 2nd place on the above categories in 2014 ILSVRC challenge. This model achieves 92.7% top-5 test accuracy on ImageNet dataset which contains 14 million images belonging to 1000 classes. VGG-16 architecture map is presented on Fig. 7.

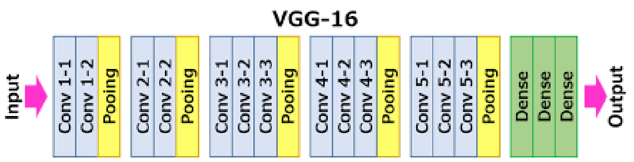

Fig. 7. VGG-16 architecture map

The input to the network is image of dimensions (224, 224, 3). The first two layers have 64 channels of 3*3 filter size and same padding. Then after a max pool layer of stride (2, 2), two layers which have convolution layers of 256 filter size and filter size (3, 3). This followed by a max pooling layer of stride (2, 2) which is same as previous layer. Then there are 2 convolution layers of filter size (3, 3) and 256 filter. After that there are 2 sets of 3 convolution layer and a max pool layer. Each has 512 filters of (3, 3) size with same padding. This image is then passed to the stack of two convolution layers. In these convolution and max pooling layers, the filters use filters with the size (3, 3). In some of the layers, it also uses (1, 1) pixel which is used to manipulate the number of input channels. There is a padding of 1-pixel (same padding) done after each convolution layer to prevent the spatial feature of the image.

After the stack of convolution and max-pooling layer, there is a (7, 7, 512) feature map. We flatten this output to make it a (1, 25088) feature vector. After this there are 3 fully connected layer, the first layer takes input from the last feature vector and outputs a (1, 4096) vector, second layer also outputs a vector of size (1, 4096) but the third layer output a number of classes. Then after the output of 3rd fully connected layer is passed to softmax layer in order to normalize the classification vector. All the hidden layers use ReLU as its activation function. ReLU is more computationally efficient because it results in faster learning and it also decreases the likelihood of vanishing gradient problem [10].

The pre-trained model was trained on ImageNet and not on medical images. In experiments all layers for image feature extraction are used. The decision making layers of pre-trained model were deleted. Batch normalization, flatten and dense layers were added in a new model, Fig. 8.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
vgg16 (Functional)           (None, 7, 7, 512)         14714688
_________________________________________________________________
batch_normalization (BatchNo (None, 7, 7, 512)         2048
_________________________________________________________________
flatten (Flatten)            (None, 25088)             0
_________________________________________________________________
dense (Dense)                (None, 5)                 125445
=================================================================
Total params: 14,842,181
Trainable params: 2,486,277
Non-trainable params: 12,355,904
_________________________________________________________________
```

Fig. 8. Used neural network model

The initial image dataset was divided into two sets for training, and testing (70% and 30 %). The validation set consists of 100 images. Details about training processing are presented on Fig. 9, 10. Numerical results are presented in Table I.

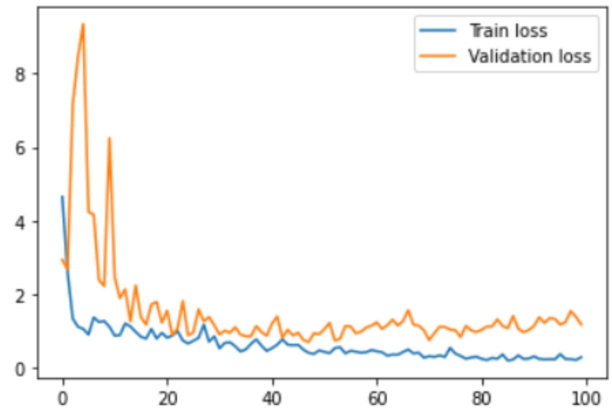

Fig. 9. Model Loss on Training and Validation Datasets

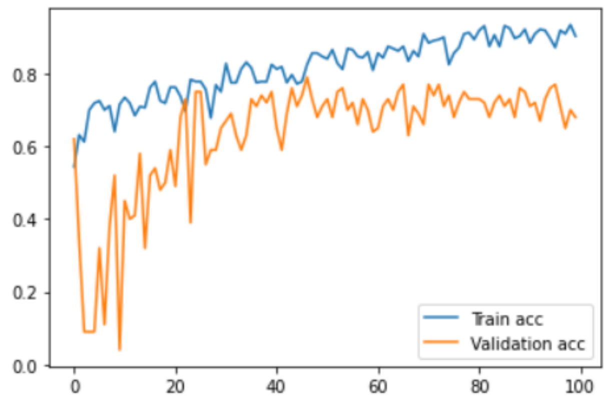

Fig. 10. Model Accuracy on Train and Validation Dataset

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.

TABLE I. TABLE TYPE STYLES

Model	Pre-trained dataset	Precision	Recall
VGG16	ImageNet	0.9230	0.8763

V. DISCUSSIONS

We have obtained preliminary acceptable results from the first phase of the study. It is obvious that the deep learning approach is a priority in computer vision in general and in our applied problem in particular. The neural network architecture used is not the best, but it allows you to get preliminary results and understand the range of problems. It is also necessary to carefully implement the stages of image preprocessing (noise reduction), since this may impair the classification results.

VI. CONCLUSIONS

The method for retina image analysis has been developed, based on image preprocessing stage and deep neural network as machine learning model. This is the preliminary results of big projects for retina image analysis. The main focus was done for diabetic retina images. Also the scheme of new technology for retina image analysis was presented.

ACKNOWLEDGMENT

The study was supported in part by the Belarusian Republican Foundation for Fundamental Research under Grant F21PACG-001.

REFERENCES

[1] V. Gulshan, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs, Jama, 2016, 316 (2), pp. 2402-2410.

[2] R. Gargeya and T. Leng, “Automated identification of diabetic retinopathy using deep learning,” pp. 962–969, 2017.

[3] Y. I. Golub, F. V. Starovoitov, V. V. Starovoitov Comparative analysis of no-reference measures for digital image sharpness assessment. Doklady BGUIR. 2019; 7(125): 113-120.

[4] M. Haloi, S. Dandapat, and R. Sinha, “A gaussian scale space approach for exudates detection, classification and severity prediction,” arXiv preprint arXiv:1505.00737, 05 2015.

[5] Kaggle. APTOS 2019 Blindness Detection (Data); 2019. Available from: https://www.kaggle.com/c/aptos2019-blindness-detection.

[6] L. G. Shapiro, G. C. Stockman, "Computer Vision", Prentice Hall, 2001, pp. 137, 150.

[7] Welcome To Colaboratory - Google Research. Available from: https://research.google.com/colaboratory/

[8] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition - arXiv preprint arXiv:1409.1556, 2014.

[9] The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) Available from: https://www.image-net.org/challenges/LSVRC/

[10] Abien Fred Agarap, Deep Learning using Rectified Linear Units (ReLU) - arXiv preprint arXiv: 1803.08375, 2018.

Pattern Recognition and Information Processing (PRIP'2021) : Proceedings of the 15th International Conference, 21–24 Sept. 2021, Minsk, Belarus. – Minsk : UIIP NASB, 2021. – 246 p. – ISBN 978-985-7198-07-8.