ITM Web of Conferences **59**, 03008 (2024) https://doi.org/10.1051/itmconf/20245903008
HMMOCS-II 2023

Generation of a text description of weakly structured optical coherence tomography images

Arthur Zhigalov[1*], *Alexander* Lositsky[2], *Lyubov* Grishina[1], and *Irina* Bolodurina[1]

[1]Research Institute of Digital Intelligent Technologies, Orenburg State University, Orenburg, Russia
[2]S. Fyodorov eye microsurgery federal state institution, Orenburg, Russia

Abstract. Computer vision methods help to automate and improve processes in the field of medicine. In the field of ophthalmology, computer vision algorithms can be used to analyze images obtained using optical coherence tomography OCT, to identify pathologies and changes in the structure of the eye, however, due to the heterogeneity of patterns and configurations of tomographs, a comprehensive solution is needed. Within the framework of this work, an approach to the construction of a system for generating a text description of DICOM images using artificial intelligence methods is presented. To build a system of automatic description of anatomical properties and pathologies, a set of models for detection and classification was built on the OCT image. Data augmentation was performed for the task of recognizing areas with retinal slices in the OCT image. The computational experiment of constructing classification models showed recognition accuracy from 0.75 to 0.93 according to the balanced accuracy metric. Based on the developed models, a web service has been developed to demonstrate the functionality, which provides a report on finding 11 tags on an OCT scan.

1 Introduction

With the latest achievements in the field of image analysis by computer vision methods, it has become possible to automate some processes in the field of medicine. Intelligent algorithms play an important role by providing the means to automate the analysis of medical images in several tasks.

Computer vision algorithms are widely used to classify, diagnose, and automatically interpret medical images, such as X-rays, to detect signs of diseases such as pneumonia or sclerosis. Another task is image segmentation: to highlight specific areas in medical images, such as MRI or CT scans, which can help doctors identify tumors, abnormalities or other changes in tissues.

In the field of ophthalmology, computer vision algorithms can be used to analyze images obtained using optical coherence tomography OCT, to identify pathologies and changes in the structure of the eye.

* Corresponding author: leroy137.artur@gmail.com

OCT is one of the most informative ways to analyze the pathological process and condition of the retina and fundus. This method is a non-invasive way to visualize the structure of the back of the eye in high resolution. With the help of OCT, an ophthalmologist can determine the degree of pathology, as well as anatomical indicators, for example, the thickness of the retinal layers.

There are several options for generating a scan of an OCT image using an optical coherence tomograph, depending on its settings and model. Figure 1 shows 2 examples of OCT templates.

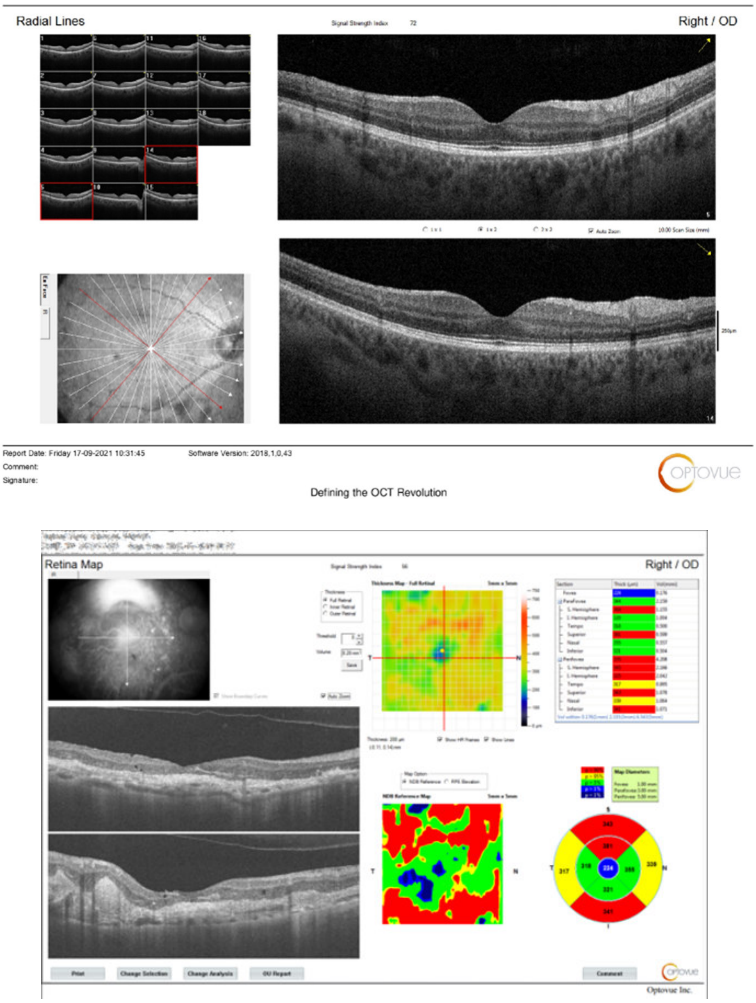

Fig. 1. Examples of OCT snapshot templates.

The heterogeneity of the data structure and templates of the OCT snapshot for building an automated system for generating conclusions requires solving the problem of recognizing the necessary template segments.

In this regard, it is relevant to develop a comprehensive automated system for analyzing OCT images. Within the framework of this work, we propose an approach to building a system for generating a text description of DICOM images using artificial intelligence methods. To build a system of automatic description of the OCT image, the tasks of detecting retinal segments and the task of multi-label classification for the presence of a pathological process and analysis of anatomical properties will be considered.

2 Related work

At the moment, computer vision methods are actively being researched in ophthalmology.

So in [1], the authors proposed a structure that includes expert diagnostics and understanding of the analysis of optical coherence tomography (OCT) using multimodal learning. To demonstrate the proposed approach, the authors create a dataset to improve the classification of diseases using OCT.

The authors of the study [2] used a class of automatic classification algorithms VGG-16, which showed the best results. Its accuracy was 94%.

In [3], the authors used principal component Analysis (PCA) to reduce the size of images before analysis. KNN, SVM and deep learning methods were used to identify the retina affected by diabetes. The system performance is tested using the DRIVE and DDBI datasets.

The technology of transfer learning in the search for pathology was used by the authors of the work [4]. Pre-trained models with the architecture of Inception V3 and Resnet152V2 were used to determine the state of the eyes.

In [5], the authors propose deep learning models that classify OCT images of patients into four categories, such as choroidal neovascularization (CNV), diabetic macular edema (DME), druses and norm. Two different models are proposed. One uses three binary convolutional neural network (CNN) classifiers, and the other uses four binary CNN classifiers.

The authors of the study [6] apply vision transformer (ViT). In the work, the transormer-based model showed a better result on the f1-score metric than the convolutional neural network model.

In the article [7], the authors use three different pre-trained CNN models - VGG16, InceptionV3 and MobileNet. The best of which turned out to be VGG.

The authors of the work [8] based on 8000 images built a classification model for 3 classes of retinal pathologies using the convolutional neural network Xception.

The analysis of the sources showed the relevance of the use of computer vision methods for the recognition of pathologies and the anatomical properties of the retina in OCT images.

3 Methods

3.1 Detection

Before analyzing the image for the presence of a pathological process or a description of the anatomical properties of the retina, it is necessary to isolate the analyzed segments from the OCT image. The main informative blocks on the OCT report are cross-sectional sections of the retina. The main methods for recognizing areas with retinal slices are object detection models.

To solve the problem of detection from the same type of OCT report templates, data augmentation was performed by artificially generating images with annotation files. For different OCT with the same configuration

Since OCT images with the same configuration of the arrangement of retinal segments are in constant positions, it is possible to cut out the desired segments pixel by pixel and generate a file with a random location. Figure 2 shows an example of image generation.

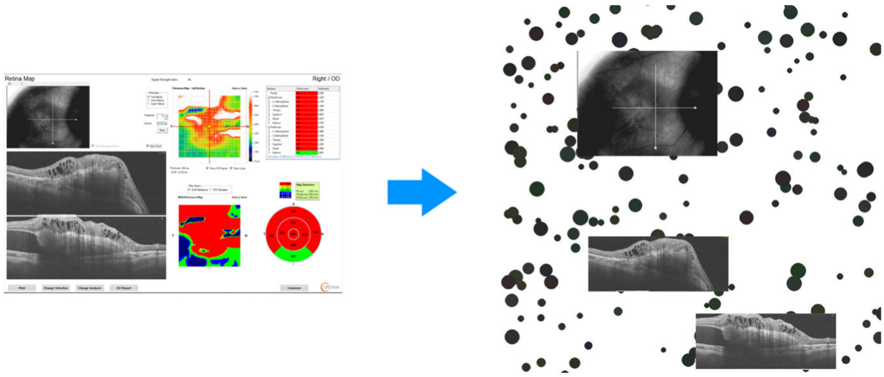

Fig. 2. Example of augmentation of an OCT snapshot template.

One of the popular models for solving the problem of recognizing objects in an image is the Faster R-CNN detection algorithm.

Faster R-CNN is an object detection algorithm that combines the architecture of a convolutional neural network with methods of processing regions (Region Proposal Networks, RPN) for faster and more accurate detection of objects in images.

3.2 Classification

The description of anatomical properties and pathological processes is solved as a multi-level classification problem with 11 non-overlapping classes. There are a total of 13,000 images in the dataset. The names of classes and their number in the data are presented in Table 1.

Table 1. Description and number of classes.

label	description	count
fovea_1	Complete posterior vitreous detachment	737
fovea_3	Incomplete detachment of the posterior hyaloid membrane with adhesion in the macula is observed	1107
fovea_4	Incomplete detachment of the posterior hyaloid membrane with adhesion in the fovea is observed	440
fovea_8	Vitreomacular traction syndrome (vertical traction) is observed	228
fovea_9	Epiretinal fibrosis is observed	2612
fovea_11	The retinal profile is deformed due to tangential horizontal traction (fovea is smoothed)	554
fovea_12	The epiretinal membrane partially separated spontaneously	306
fovea_13	The retinal profile has a wave-like deformation after peeling	392
square_1	Retina in the macular region without visible pathology	1283
square_2	Insufficient transparency of optical media for a reliable assessment of changes in the structure of the retina	451
square_4	There is a gross violation of the architectonics of the retina	290

To solve the problem, 11 models were used that solve the problem of binary classification for the presence of a particular feature. The Dense-Net 121 architecture was chosen as the neural network model. Unlike conventional convolutional neural networks, in which each layer is connected only to the previous and subsequent layers, in Dense-Net each layer is directly connected to all previous layers in the deep network. This creates an extremely tight connection between the layers and allows you to effectively reuse the features.

4 Results

4.1 Detection

The main informative segments are horizontal cross-sectional images in the OCT image. Using the example from augmentation (Figure 2), they represent the 2 lower slices.

A data set of 500 samples was generated for the task, consisting of both original templates of OCT images and artificially created ones. The Faster R-CNN model with backbone RetinaNet was used to recognize the bounding-box horizontal sections of the retina. The accuracy indicator for the mAP metric (MS-COCO) was 0.9271.

4.2 Classification

Weighted binary cross-entropy is used as a loss function to account for data imbalance. The results of training models using the balanced accuracy metric on training and test datasets are shown in Figure 3.

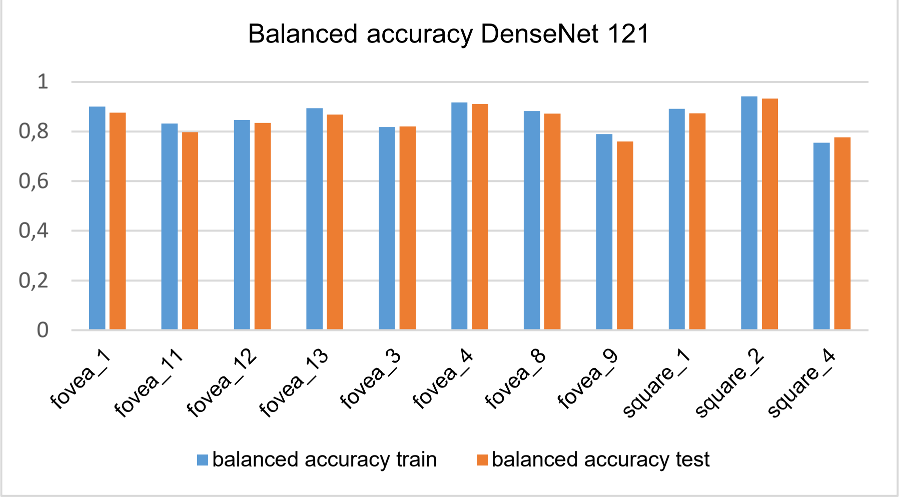

Fig. 3. Результаты balanced accuracy по каждой из модели.

The greatest balanced accuracy is achieved when recognizing the square_2 label, the metric indicator is 0.9317. While for fovea_9 of all classes, the accuracy index is the lowest and is equal to 0.7592.

4.3 Service

Based on the developed models, a web service has been developed to demonstrate the functionality, available at http://retinadeepai.site. To get the result, you need to upload an image or choose from existing examples. After clicking on the "Find objects" button, the service will launch the detection algorithm and give the result with the found retinal slices.

If at least 1 retinal slice is found in the image, after clicking on the "Markup Analysis" button, the algorithm for recognizing anatomical and pathological processes on the found segment will start.

After analyzing the system, it will display a report with the found labels and the degree of confidence in its presence as a percentage.

5 Conclusion

Within the framework of this work, an approach to the construction of a system for generating a text description of DICOM images using artificial intelligence methods is proposed. To build a system of automatic description of anatomical properties and pathologies, a complex of models for detection and classification into 11 labels was built on the OCT image. Data augmentation was performed for the task of recognizing areas with retinal slices in the OCT image. The computational experiment of constructing DenseNet-121 classification models showed recognition accuracy from 0.75 to 0.93 according to the balanced accuracy metric.

Based on the constructed models, a web service has been developed with which you can upload a scan report on OCT.

The research was carried out with the financial support of the scholarship of the President of the Russian Federation for young scientists and postgraduates (SP-919.2022.5).

References

1. Y. -Y. Logan, K. Kokilepersaud, G. Kwon, G. AlRegib, C. Wykoff, H. Yu, *Multi-Modal Learning Using Physicians Diagnostics for Optical Coherence Tomography Classification*, in 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), Kolkata, India, pp. 1-5 (2022).
   https://www.doi.org/10.1109/ISBI52829.2022.9761432
2. J.M. Jerry, T. Thomas, R. Pugalenthi, T. Thomas, *Automatic Classification of Retinal Fundus Images for Diabetic Retinopathy Detection*, in 2023 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI), Chennai, India, pp. 1-5 (2023). https://www.doi.org/10.1109/RAEEUCCI57140.2023.10134335
3. A.S. Jadhav, R.V. Pawar, P.B. Patil, *Segmentation and Classification of Retina Images using SVD Features*, in 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT), Mysuru, India, pp. 712-716, (2021).
   https://www.doi.org/10.1109/ICEECCOT52851.2021.9707985
4. D.S. Renuga, G. Gopalakrishnan, K. Gayathri, *Detection of Eye Strain using Retina Medical Images through CNN*, in Smart Technologies, Communication and Robotics (STCR), Sathyamangalam, India, pp. 1-5 (2021). https://www.doi.org/10.1109/STCR51658.2021.9589024
5. J. Kim, L. Tran, *Retinal Disease Classification from OCT Images Using Deep Learning Algorithms,* in 2021 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), Melbourne, Australia, pp. 1-6 (2021). https://www.doi.org/10.1109/CIBCB49929.2021.9562919
6. A.M. Mutawa, S. Sruthi, *Diabetic Retinopathy Classification using Vision Transformer*, in 6th European Conference on Electrical Engineering & Computer Science (ELECS),

Bern, Switzerland, pp. 25-30 (2022).
https://www.doi.org/10.1109/ELECS55825.2022.00012

7. P.K. Das, S. Pumrin, *CNN Transfer Learning for Two Stage Classification of Diabetic Retinopathy using Fundus Images*, in 2023 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON), Phuket, Thailand, pp. 443-447 (2023).
https://www.doi.org/10.1109/ECTIDAMTNCON57770.2023.10139437
8. M.T. Do, H.N. Huynh, T.N. Tran, T.L. Hoang, *Prediction of Retina Damage in Optical Coherence Tomography Image Using Xception Architecture Model*, in 2023 IEEE 5th Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS), Tainan, Taiwan, pp. 58-61 (2023).
https://www.doi.org/10.1109/ECBIOS57802.2023.10218586.