

POSSIBILITIES OF USING VISUAL TRANSFORMERS IN THE CLASSIFICATION OF OPHTHALMOLOGIC DISEASES

Volkov E.N.¹, Averkin A.N.¹

¹ Dubna State University, Institute of System Analysis and Management,
Department of System Analysis and Management, 141982, Russia,
Dubna, 19, Universitetskaya str., e-mail envolkoff1998@yandex.ru

The possibilities of using artificial intelligence technologies in diagnostics of various diseases are expanding every year. This tendency also takes place in ophthalmology, where the use of artificial neural networks of various types has brought the diagnostics of diseases to a whole new level. Traditionally, convolutional neural networks (CNN) have been used to analyze ophthalmic images, namely retinal camera images, as has been demonstrated in many papers [1]. However, the advent of visual transformer models in 2020, based on an improved attention mechanism, has yielded significant gains in classification and segmentation accuracy in medical image analysis [2, 3].

Although transformer applications are still inferior to CNN in analyzing ophthalmic images, selected studies [4, 5] show their promise in this direction. The created architecture of the visual transformer model for the task of glaucoma and diabetic retinopathy classification, based on fundus images, shows a high result on the test sample (F1-0.81; AUC-0.87), which indicates the possibility of improving the result both by fine-tuning the network and by creating an ensemble of models in further studies.

Литература

1. E. N. Volkov and A. N. Averkin, "Possibilities of Explainable Artificial Intelligence for Glaucoma Detection Using the LIME Method as an Example," 2023 XXVI International Conference on Soft Computing and Measurements (SCM), Saint Petersburg, Russian Federation, 2023, pp. 130-133. DOI: 10.1109/SCM58628.2023.10159038.
2. Zhang Y., Wang J., Gorriz J. M. et al. Deep Learning and Vision Transformer for Medical Image Analysis // Journal of Imaging. MDPI AG, 2023. Vol. 9, № 7. P. 147. DOI: 10.3390/jimaging9070147.
3. Azad R., Kazerouni A. et al. Advances in Medical Image Analysis with Vision Transformers: A Comprehensive Review // arXiv preprint arXiv:2301.03505. – 2023.
4. Yu S., Ma K., Bi Q. et al. Mil-vt: Multiple Instance Learning Enhanced Vision Transformer for Fundus Image Classification // Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VIII 24. – Springer International Publishing, 2021. – P. 45-54. DOI: 10.1007/978-3-030-87237-3_5.
5. Sun R., Li Y., Zhang T. et al. Lesion-aware Transformers for Diabetic Retinopathy Grading // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2021. – P. 10938-10947.