

DEVELOPMENT OF A RETINAL PATHOLOGY CLASSIFICATION MODEL USING TRANSFER LEARNING AND OPEN DATASETS

Lyubov E. Aksenova

LLC PREDICT SPACE

Kuban State Technological University
Novorossiysk & Krasnodar, Russia
axenovalubov@gmail.com

Kirill D. Aksenov

LLC PREDICT SPACE

Kuban State Technological University
Novorossiysk & Krasnodar, Russia

Anton V. Prisyazhnyuk

LLC PREDICT SPACE

Kuban State Technological University
Novorossiysk & Krasnodar, Russia

Victoria V. Myasnikova

LLC PREDICT SPACE

Maykop State Technological University
Maykop, Republic of Adygea, Russia

Andrey V. Krasov

Saint-Petersburg State University of Telecommunications
St. Petersburg, Russia

Article history:

Received 24.06.2025, Accepted 01.09.2025

Abstract

The retina is a multilayer structure that perceives and transmits visual stimuli to the brain via the optic nerve; however, its complex structure makes it vulnerable to damage. Optical coherence tomography (OCT) is the gold standard for diagnosing retinal pathologies. This study aimed to develop a deep learning model for the classification of retinal pathologies based on OCT data. The study design included the use of transfer learning and training a model with a ResNet-50 architecture on three open datasets containing over 110,000 OCT images and 10 retinal condition classes, including normal and 10 pathological conditions. As a result, the model demonstrated a high classification accuracy of 95 %. Thus, transfer learning and significant dataset expansion provide high classification model accuracy. The study also highlights that open access to data significantly impacts the development of artificial intelligence technologies in healthcare.

Key words

Artificial intelligence, classification, neural network, open dataset, optical coherence tomography (OCT), retina

1 Introduction

The retina is a thin, multilayered structure responsible for receiving, modulating, and transmitting visual stimuli from the external world to the optic nerve and, ultimately, to the visual cortex of the brain [Herzlich et al., 2010]. The complex structure of the retina makes it highly susceptible to various types of damage. In in-

dividuals aged 60 to 95 years, the prevalence of retinal diseases is 52.37 %, with the most common being age-related macular degeneration (AMD) (35.43 %), followed by hypertensive retinopathy (4.35 %), epiretinal membrane (ERM) (3.66 %), branch retinal vein occlusion (BRVO) (2.90%), and diabetic retinopathy (DR) (2.15%) [Thapa et al., 2020]. According to the World Health Organization, 80% of all cases of vision impairment can be prevented or treated [GBD Collaborators, 2021]. This statistic underscores the pressing need for the development of preventive measures to address retinal diseases, including effective methods for their early visualization.

Optical coherence tomography (OCT) is the gold standard for diagnosing most retinal diseases [Turbert, 2023], enabling cross-sectional imaging of the retina with micron-level resolution, providing high detail and accuracy for visualizing various structures and pathologies. The resulting images can be evaluated both qualitatively and quantitatively. Since its inception, OCT has undergone significant advancements, improving image quality, speed, and resolution [Kharousi et al., 2013]. However, the processes of analyzing and interpreting OCT data remain highly complex. One of the most common challenges in interpretation is the failure to identify subtle and early signs of disease or rare pathologies. Furthermore, analyzing such medical data is complicated by the need not only to interpret images but also to describe the results, which can take up to 50% of the time. Meanwhile, there is a global shortage of highly qualified specialists in ophthalmology, not only in Russia but worldwide [Neroev, 2014]. The large number of patients, coupled with a shortage of medical personnel, leads to over-

work and increased errors in interpreting medical examinations by physicians [Hussain and Oestreicher, 2018]. Recent advances in artificial intelligence (AI) have made it possible to analyze various types of medical data—from EEG spectra to high-resolution OCT images—demonstrating the universality and adaptability of modern deep learning approaches [Plotnikov, 2024; Kermany et al., 2018]. According to the literature, automating the identification of pathologies in OCT images using artificial intelligence technologies not only accelerates the diagnostic process but also demonstrates accuracy comparable to that of an ophthalmologist [Kermany et al., 2018]. Thus, the development and implementation of such methods in clinical practice could facilitate early detection of symptoms, enabling treatment to begin at earlier stages. Initial studies related to the automation of OCT data analysis using artificial intelligence technologies involved the development of models based on manual feature extraction and their classification [Chen et al., 2020; Dalal and Triggs, 2005; Rong et al., 2014]. However, this approach is labor-intensive and not always accurate. At the same time, deep learning-based methods can automatically extract features from data, significantly simplifying the process [Larochkin et al., 2025]. Several studies have utilized neural network architectures such as VGG-16, AlexNet, and GoogleNet for automatic feature extraction and their subsequent use in training classifiers like kNN (k-Nearest Neighbors), Random Forest, and SVM (Support Vector Machine) [Awais et al., 2017; Perdomo et al., 2018; Chan et al., 2018; Khan and Khan, 2024]. These models are trained on large datasets and can identify complex relationships that are difficult to detect manually, making them especially effective for classification tasks. However, the successful application of such models largely depends on the availability of sufficient data. Today, there are many open datasets obtained via OCT that are also used to accelerate model development. Umer Sadiq Khan and colleagues trained a model with a concatenation layer for combining weights from two networks with ResNet50 and EfficientNetB0 architectures on the OCT2017 dataset comprising 84,452 OCT images [Mathews and Anzar, 2022]. As a result, the accuracy of detecting four retinal pathologies was 97.50%. In another study, a comparison of ResNet50, ResNet50V2, DenseNet121, and DenseNet169, trained on the same dataset (expanded to 30,964 images through augmentation), achieved accuracies of 97.9%, 96.9%, 98.8%, and 99.4%, respectively [Maurya et al., 2024]. The use of ensemble classification methods combining the results of models based on architectures such as Xception, NASNetMobile, MobileNet-V1, and EfficientNet-B3 resulted in an accuracy of 99.80% on the OCT2017 dataset [Saini et al., 2023].

It is worth noting that several strategies can be used when developing deep machine learning models for specialized tasks. Creating custom architectures involves designing new neural network architectures tailored to

the unique requirements and characteristics of a specific task. For example, DilipKumar Jang Bahadur Saini and colleagues developed an LCNN model that uses lesion attention modules in OCT images to improve the accuracy of disease detection [Fang et al., 2022]. In the study by Leyuan Fang, a described architecture incorporated a Siamese neural network based on ResNet18 (SSPSF) with disease detection accuracies of 97.74% and 98.94% on the RETOUCH and AI Challenger open datasets, respectively [Haloi, 2018]. Custom models are advantageous in the short term, but in the long term and when expanding the scope of tasks or retraining on data of a different modality, such models lack robustness and predictability. Meanwhile, well-known architectures can be fine-tuned on specialized data related to a specific task. An example is the use of networks like VGG16, Inception V3, or ResNet for medical image classification, where the model is pre-trained on large datasets such as ImageNet and then fine-tuned on medical images to recognize specific pathologies [Errabih et al., 2022; Kermany et al., 2018; Kulyabin et al., 2024; Aksenova et al., 2023]. Thus, transfer learning not only saves time and computational resources but also enhances the model's generalization capabilities on medical images. At the previous stages of research, we developed and tested a segmentation model using clinical data to determine the type and evaluate the quantitative parameters of biomarkers in OCT images [Aksenova et al., 2024]. Additionally, we selected the neural network architecture for a model to classify retinal diseases based on OCT images [Gholami et al., 2018]. In the present study, we developed a deep machine learning model for the classification of retinal diseases using transfer learning and open datasets.

2 Materials and methods

In this study, we used open datasets such as OCTDL, OCTID, and OCT2017. All datasets included images obtained using optical coherence tomography and represented B-scans of the retina—linear cross-sectional slices obtained by overlaying consecutive slices, called A-scans. The OCTDL dataset contains more than 2000 OCT images annotated for diseases such as age-related macular degeneration (AMD), diabetic macular edema (DME), epiretinal membrane (ERM), normal (NO), branch retinal vein occlusion (RVO), and vitreoretinal dystrophy (VID) [Kulyabin et al., 2024]. The data were obtained and annotated at Friedrich-Alexander University Erlangen-Nürnberg, Ural Federal University named after the First President of Russia B. N. Yeltsin, and Ural State Medical University. The OCTID dataset is a database of optical coherence tomography images containing more than 500 images grouped by conditions such as normal (NO), age-related macular degeneration (AMD), central serous retinopathy (CSR), and diabetic retinopathy (DR) [Kermany et al., 2018]. The OCT2017 dataset consists of OCT images and chest X-

rays. The OCT images are classified into four groups: choroidal neovascularization (CNV), diabetic macular edema (DME), drusen (Drusen), and normal images (NO). The data were obtained and annotated at the University of California, San Diego [He et al., 2015]. It is worth noting that all images were acquired using different devices—Optovue Avanti RTVue XR, Cirrus HD-OCT machine by Carl Zeiss Meditec, Inc. (Dublin, CA), and Spectralis OCT by Heidelberg Engineering (Germany). Examples of the images are shown in Figure 1.

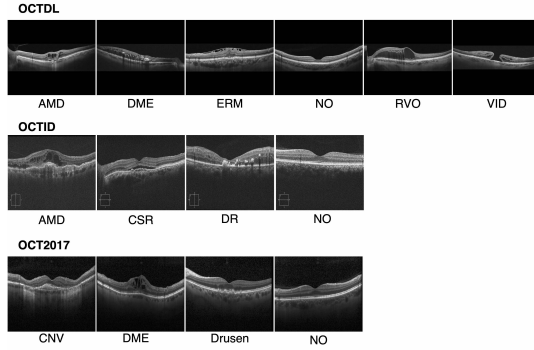


Figure 1. Examples of retinal pathologies from three open datasets: OCTDL, OCTID, and OCT2017.

Table 1 provides a description of the class distribution within the datasets. Before training, the datasets were divided into training, validation, and test sets in a 65/15/20 ratio. In each case, the images were randomly assigned to the respective sets. Thus, the test datasets for the three datasets contained different images.

Table 1. Class distribution in the dataset by disease types

Class	OCTDL	OCTID	OCT2017
AMD	1231	0	0
DME	147	0	11349
ERM	155	0	0
NO	332	206	51140
RVO	101	0	0
VID	76	0	0
CSR	0	102	0
DR	0	110	0
CNV	0	0	37206
Drusen	0	0	8617
Sum	2042	418	108312

Before training the model, we performed preprocess-

ing and data augmentation. Preprocessing included resizing the images to 224x224 pixels. The following augmentation methods were applied: random cropping, horizontal and vertical flipping, rotation, shifting, and Gaussian blur.

To develop a classification model for OCT data related to various diseases, we used a deep neural network with a ResNet50 architecture [Krizhevsky et al., 2017]. This architecture, which was pre-trained on the large ImageNet dataset [Fang et al., 2022], consists of 50 layers organized into 5 blocks, each containing a set of residual blocks (Figure 2). Convolutional layers are the primary components of the ResNet-50 architecture. They apply a convolution operation to the input image to extract spatial features.

Let X_l be the input representation at layer l . The application of a convolutional layer followed by ReLU activation can be written as follows:

$$X_{l+1} = \text{ReLU}(X_l * W_l + b_l) \quad (1)$$

where:

W_l — trainable filters (weights) of layer l ;

b_l — bias of layer l ;

$*$ — convolution operation;

$\text{ReLU}(z)$ — activation function defined as $\max(0, z)$.

Subsampling layers (max pooling) reduce the dimensionality of the input data from the convolutional layer while preserving important features. This helps to decrease the number of parameters, reduce computational costs, and retain significant features. The subsampling process can be expressed as follows:

$$X_{\text{pooled}}(i, j) = \max_{(m, n) \in \mathcal{P}(i, j)} X_l(m, n) \quad (2)$$

where:

$\mathcal{P}(i, j)$ — the receptive field for the position (i, j) in the input representation X_l ;

$X_{\text{pooled}}(i, j)$ — the result of the max pooling operation.

Residual blocks consist of three convolutional layers, each accompanied by a batch normalization layer and a ReLU activation function. The three layers are 1x1, 3x3, and 1x1 convolutions, where the 1x1 layers are responsible for reducing and subsequently restoring (recovering) the dimensions, leaving the 3x3 layer as the bottleneck with smaller input/output dimensions. The residual block can be described as follows:

$$X_{\text{out}} = \text{ReLU}(\mathcal{F}(X_{\text{in}}) + X_{\text{in}}) \quad (3)$$

where:

X_{in} — the input tensor for the residual block;
 $\mathcal{F}(X_{in})$ — the result of three sequential convolutions (1×1 , 3×3 , and 1×1) with normalization and ReLU activation.

The output of the third convolutional layer is added to the input of the residual block, and the resulting tensor is passed through the ReLU activation function again using "skip connections" (or "shortcut connections"). Skip connections are special connections in neural networks that allow information to bypass certain layers and be passed directly to later processing stages. In the ResNet-50 architecture, there are two types of residual blocks: convolutional blocks and identity blocks (Figure 2). These blocks have a similar structure but differ in how they use "skip connections." The identity block is used when the input and output tensors have the same dimensions. The convolutional block is used when the dimensions of the tensors need to be changed. This is achieved by using a 1×1 convolutional layer in the "skip connection," which adjusts the dimensions of the input tensor to match the dimensions of the output tensor. The final stage of the ResNet50 architecture is the average pooling layer, which reduces the output features to a single vector. Let the output tensor after the last convolutional block have dimensions $H \times W \times C$. The average pooling operation transforms it into a vector X_{pooled} with dimensions $1 \times 1 \times C$:

$$X_{pooled}(c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X(i, j, c) \quad (4)$$

This vector is then passed to a 10-dimensional fully connected layer (fc), which transforms it into 10 output values corresponding to the number of classes:

$$z = W_{fc} \cdot X_{pooled} + b_{fc} \quad (5)$$

where:

W_{fc} — weight matrix of size $10 \times C$;
 b_{fc} — bias vector of size 10;
 z — output vector of size 10, where each component corresponds to one of the 10 classes.

These values are then passed through a softmax function, which converts them into probabilities for each class:

$$\hat{y}_i = \frac{\exp(z_i)}{\sum_{j=1}^{10} \exp(z_j)} \quad (6)$$

This ensures the final prediction of the model.

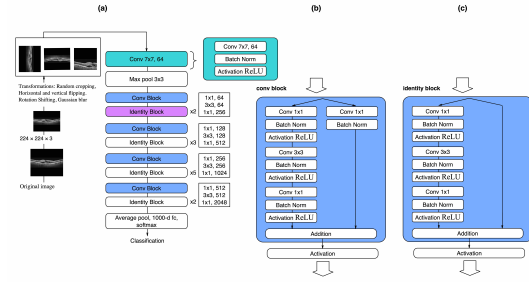


Figure 2. Overview of the ResNet-50 architecture. (A) Transformation of the original OCT images and an overview of the overall structure of ResNet-50, where the symbols "x2", "x3", "x5" indicate the number of repetitions of the corresponding blocks. (B) Structure of the convolutional block (conv block). (C) Structure of the identity block (identity block).

To evaluate the accuracy of the models, the following metrics were measured: Accuracy (7), Precision (8), Recall (9), F1 Score (10), and AUC-ROC.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (8)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (9)$$

where TP — True Positives, TN — True Negatives, FP — False Positives, FN — False Negatives.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (10)$$

AUC-ROC (Area Under the ROC Curve) is a measure of a model's ability to distinguish between classes. The ROC curve (Receiver Operating Characteristic) is a graph that shows the performance of the model across all classification thresholds. The AUC-ROC is calculated as the area under the ROC curve, which is built based on various values of sensitivity (Recall) and specificity (1 - False Positive Rate). These metrics were calculated for each class using the "one vs. rest" method. To evaluate metrics for the entire dataset, the mean value of the obtained scores was calculated.

For training the models, the early stopping algorithm was used. Early stopping occurred when the Accuracy metric, measured using the validation dataset, reached its maximum value.

The training and evaluation of the models were conducted using the PyTorch deep learning framework (version 2.1.1) and Yandex Cloud resources

(<https://yandex.cloud/ru/docs/billing/concepts/bonus-account>), which were provided as part of the Yandex Cloud Boost program. The configuration included vCPUs on the Intel Broadwell platform and NVIDIA® Tesla® V100 GPUs. The developed model was integrated with the large language model (LLM) using containerization technologies and APIs to enable seamless interaction and automation. This integration facilitated the creation of a chatbot application designed to provide clinicians with efficient access to diagnostic insights and decision support for OCT image analysis.

3 Results

In this study, the ResNet-50 model was trained on open datasets OCTDL, OCTID, and OCT2017. The study design is presented in Figure 3.

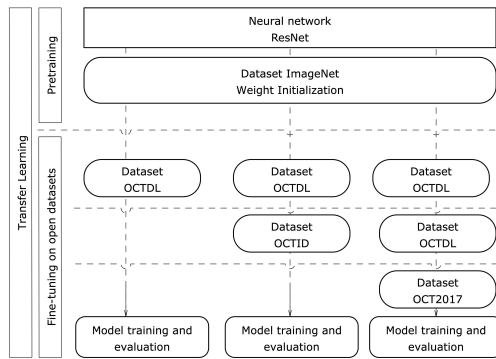


Figure 3. Study Design.

To evaluate the effectiveness of the models, accuracy metrics (1–4) and AUC-ROC were used. Table 2 provides the metric values measured for the entire dataset across the three models.

Table 2. Quantitative performance metrics of machine learning models. Data1: Dataset with images from the open OCTDL dataset; Data2: Dataset with images from the combined open OCTDL and OCTID datasets; Data3: Dataset with images from the combined open OCTDL, OCTID, and OCT2017 datasets.

Table 2. Quantitative performance metrics of machine learning models. Data1: Dataset with images from the open OCTDL dataset; Data2: Dataset with images from the combined open OCTDL and OCTID datasets; Data3: Dataset with images from the combined open OCTDL, OCTID, and OCT2017 datasets.

Dataset	Accuracy	F1 Score	Precision	Recall	AUCROC	N_{test}
Data1	0.93	0.89	0.88	0.91	0.99	420
Data2	0.9326	0.8821	0.93	0.85	0.9930	538
Data3	0.9535	0.8787	0.88	0.89	0.9960	19032

sistently show high values across all datasets, confirming the model's effectiveness. The Precision and Recall metrics exhibit divergent trends as the dataset size increases, while the F1 Score decreases with larger datasets.

When evaluating the accuracy for individual classes, metrics Precision, Recall, and F1 Score demonstrate a strong correlation with the number of available samples for each class (tab. 3). Classes with larger datasets, such as DME (2304 images) and NO (5382 images) in Data3, show consistent improvements, with F1 Scores reaching 0.93 and 0.97, respectively. In contrast, smaller classes like VID (15 images) and RVO (29 images) exhibit stable but limited performance due to insufficient data. Precision fluctuates more for underrepresented classes, while Recall improves for larger datasets, reflecting better coverage of true positives. These results emphasize the need for balanced datasets and augmentation to enhance performance for underrepresented classes.

Table 3. Quantitative performance metrics of machine learning models by class.

Class	Precision			Recall			F1 Score			N_{test} images		
	Data1	Data2	Data3	Data1	Data2	Data3	Data1	Data2	Data3	Data1	Data2	Data3
AMD	0.97	0.96	0.88	0.96	0.98	0.93	0.97	0.97	0.91	238	273	246
DME	0.82	0.84	0.92	0.90	0.94	0.94	0.86	0.89	0.93	30	33	2304
ERM	0.86	0.92	0.73	0.91	0.66	0.96	0.89	0.77	0.83	35	35	56
NO	0.92	0.91	0.98	0.88	0.99	0.97	0.90	0.95	0.97	80	109	5382
RVO	0.74	0.88	0.79	0.77	0.72	0.72	0.76	0.79	0.75	22	29	29
VID	0.82	0.93	0.87	0.93	0.87	0.87	0.87	0.90	0.87	15	15	15
CSR	0.00	0.91	0.95	0.00	0.91	0.82	0.00	0.91	0.88	0	22	22
DR	0.00	0.95	0.53	0.00	0.91	0.73	0.00	0.93	0.62	0	22	22
CNV	0.00	0.00	0.97	0.00	0.00	0.97	0.00	0.00	0.97	0	0	8977
Drusen	0.00	0.00	0.89	0.00	0.00	0.87	0.00	0.00	0.88	0	0	1979

The results of the confusion matrix calculations are presented in Figure 4. The high diagonal values, particularly for well-represented classes like AMD, DME, and NO, indicate strong performance and accurate classification in these categories. For underrepresented classes, such as RVO and VID, the off-diagonal values highlight occasional misclassifications, which may be due to insufficient training data or overlap in feature representation with other classes. Additionally, the matrices show that the model struggles with certain classes like CNV and Drusen, where the low Precision values suggest challenges in correctly identifying true positives.

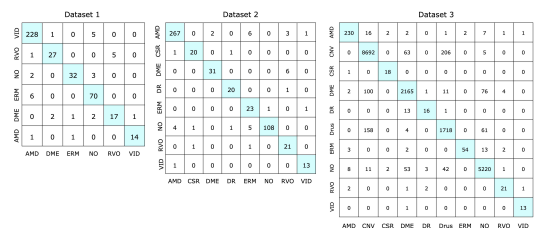


Figure 4. Confusion matrices for 3 datasets.

The resulting metrics, Accuracy and AUC-ROC, con-

To improve the interpretation of the model's results, this study employed a heatmap generation method that uses gradients and activation maps from the last convolutional layer of the neural network to create heatmaps (fig. 5). These maps highlight the regions of the image that are most important for the model's decision-making.

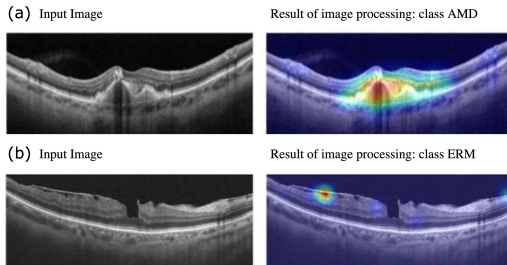


Figure 5. Results of the heatmap generation method. (A) In an image of age-related macular degeneration (AMD), the highlighted region corresponds to the detachment of the retinal pigment epithelium, a biomarker of the disease. (B) Correct identification of the presence of an epiretinal membrane (ERM) is accompanied by activation of the part of the image where the membrane is located.

An example of using a chatbot based on the developed model together with LLM is shown in Figure 6.

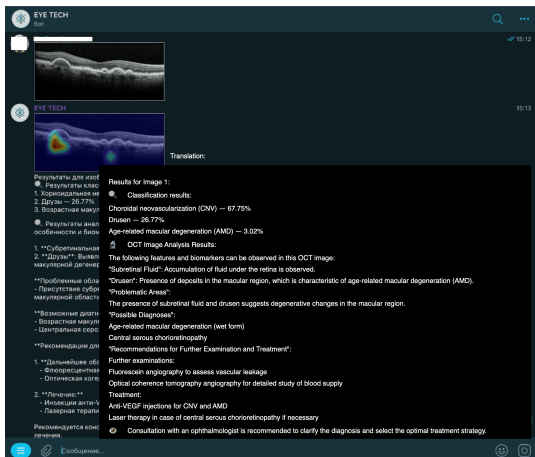


Figure 6. Chatbot based on the model and LLM.

4 Discussion

In this study, a deep machine learning model was developed for the classification of retinal pathologies. The development process involved the application of transfer learning as well as the use of open OCT image datasets. The best accuracy achieved for the classification of 12 retinal pathology classes was 0.95 using a model based on the ResNet50 architecture. Studies [Kulyabin et al., 2024; Peng et al., 2023; Fang et al., 2022] also report

results from the use of datasets such as RETOUCH, OCTDL, OCTID, and Kermany.

Table 4. Comparison of classification metrics for retinal pathology detection across studies and datasets.

Ref.	Dataset	Network	AUCROC	Accuracy	Precision
[?]	RETOUCH dataset	ResNet18	0.873	-	-
[?]	OCT-2017 dataset	ResNet 50	-	94.0	94.5
[?]	OCTDL	ResNet 50	0.988	0.846	0.898
	OCTID	ResNet 50	0.979	0.923	0.932
	Kermany	ResNet 50	0.999	0.998	0.998
	OCTDL + OCTDL	ResNet 50	0.996	0.957	0.954
	Kermany + OCTDL	ResNet 50	0.963	0.833	0.823
Present study	OCTDL	ResNet 50	0.990	0.930	0.880
	OCTDL + OCTID	ResNet 50	0.993	0.9326	0.930
	OCTDL + OCTID + OCT2017	ResNet 50	0.996	0.9535	0.880

It can be noted that the AUC-ROC metric shows higher values for the ResNet50 architecture when using one, two, and three datasets in this study, compared to study [Fang et al., 2022], which used the ResNet18 architecture. The Accuracy metric for the same neural network architecture and the OCT2017 dataset is higher in study [Peng et al., 2023] when the network was trained on OCTDL and OCTDL combined with OCTID, but lower when trained on OCTDL, OCTID, and OCT2017. When comparing the metrics with study [Kulyabin et al., 2024], it can be observed that the results after training on the OCTDL dataset were better in terms of AUC-ROC and Accuracy in this study, but 0.02 lower for Precision. After training on OCTDL and OCTID datasets, the metrics in this study were lower; however, after training on OCTDL, OCTID, and OCT2017, AUC-ROC and Accuracy became equal, while Precision values were lower compared to study [Kulyabin et al., 2024]. These differences may be attributed to variations in the ratio of training, validation, and test datasets, as all other parameters were equivalent in both studies. Moreover, this study successfully trained the model to classify 10 retinal conditions—9 pathologies and the normal state—whereas study [Kulyabin et al., 2024] included only 7 classes. The larger number of classes could have contributed to the relatively lower Precision metric. Thus, it can be concluded that transfer learning and the significant increase in dataset size ensure high accuracy for classification models. Differences in datasets, preprocessing methods, hyperparameters, model architecture, and the division ratio of datasets into training, validation, and test sets can significantly influence OCT image classification results. An important aspect is that open access to data has a substantial impact on the development of artificial intelligence technologies in healthcare. However, such data must be accurately labeled to minimize the impact of noise on accuracy. Additionally, transfer learning reduces the time and resources required for training models.

The primary limitation of this study is the use of only open datasets for model training. Although these

datasets are relatively large, they may not encompass all possible variations of retinal pathologies, which could reduce the model's generalizability to new data. Moreover, classification results may depend on data quality, further affecting model accuracy. It should also be noted that some classes are underrepresented, causing imbalances during model training and testing. While the model performs effectively overall, further improvements are needed for rare or underrepresented classes, possibly through data augmentation or more balanced datasets.

The integration of the model into a chatbot facilitates its application in ophthalmology practice. This could support independent clinical studies of the model's accuracy and simplify the process of implementing the system into healthcare workflows.

5 Conclusion

This study evaluated the feasibility of using transfer learning and open datasets to train a deep neural network with the ResNet50 architecture. Transfer learning and the significant expansion of the dataset ensured high accuracy for classification models. Additionally, the use of three datasets increased the number of classes the model could identify to 10. The AUC-ROC and Accuracy metrics were comparable to those reported in similar studies. The study also highlights that open access to data plays a crucial role in enabling the development of artificial intelligence technologies in healthcare.

References

- Aksenova, L.E., Aksenov, K.D., Kozina, E.V., et al. (2023). Automated System for Analysis of OCT Retina Images Development and Testing. *Doklady Mathematics*, 108(Suppl 2), S310–S316. doi: 10.1134/S1064562423701545.
- Aksenova, L.E., Aksenov, K.D., Prysyazhnyuk, A.V., Myasnikova, V.V., and Krasov, A.V. (2024). Development of an OCT data classification model for determining the presence and type of ophthalmic diseases. *Computing, Telecommunications and Control*, 17(3), 103–113. doi: 10.18721/JCSTCS.17310.
- Awais, M., et al. (2017). Classification of SD-OCT images using a deep learning approach. In *Proceedings of the IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*. IEEE, pp. 489–492. doi: 10.1109/ICSIPA.2017.8120661.
- Chan, G.C.Y., et al. (2018). Fusing results of several deep learning architectures for automatic classification of normal and diabetic macular edema in optical coherence tomography. In *Proceedings of the IEEE Conference*, Honolulu, HI. pp. 670–673.
- Chen, S., Zhong, S., et al. (2020). Iterative scale-invariant feature transform for remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 59, 3244–3265. doi: 10.1109/TGRS.2020.3008609.
- Dalal, N., and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. doi: 10.1109/CVPR.2005.177.
- Errabih, A., Boussarhane, M., Nsiri, B., Sadiq, A., El Yousfi Alaoui, M.H., Oulad Haj Thami, R., et al. (2022). Identifying Retinal Diseases on OCT Image Based on Deep Learning. *International Journal of Online and Biomedical Engineering*, 18, 141–159. doi: 10.3991/ijoe.v18i15.33639.
- Fang, L., Guo, J., He, X., et al. (2022). Self-supervised patient-specific features learning for OCT image classification. *Medical and Biological Engineering and Computing*, 60, 2851–2863. doi: 10.1007/s11517-022-02627-8.
- GBD 2019 Blindness and Vision Impairment Collaborators; Vision Loss Expert Group of the Global Burden of Disease Study. (2021). Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight. *The Lancet Global Health*, 9(2), e144–e160. doi: 10.1016/S2214-109X(20)30489-7.
- Gholami, P., et al. (2018). OCTID: Optical Coherence Tomography Image Database. *Computers and Electrical Engineering*, 81. Elsevier Ltd.
- Halo, M. (2018). Towards Ophthalmologist Level Accurate Deep Learning System for OCT Screening and Diagnosis. arXiv preprint.
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep Residual Learning for Image Recognition. *arXiv preprint arXiv:1512.03385*.
- Herzlich, A.A., Patel, M., Sauer, T.C., and Chan, C.C. (2010). Retinal anatomy and pathology. In *Retinal Pharmacotherapy* [Internet]. Elsevier. pp. 5–14. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9781437706031000077> doi: 10.1016/B978-1-4377-0603-1.00007-7.
- Hussain, A., and Oestreicher, J. (2018). Clinical decision-making: heuristics and cognitive biases for the ophthalmologist. *Survey of Ophthalmology*, 63(1), 119–124. doi: 10.1016/j.survophthal.2017.08.007.
- Kermany, D.S., Goldbaum, M., Cai, W., Valentim, C.C.S., Liang, H., Baxter, S.L., et al. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, 172(5), 1122–1131.e9. doi: 10.1016/j.cell.2018.02.010.
- Kermany, D., Zhang, K., and Goldbaum, M. (2018). Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification. *Mendeley Data*, Vol. 2.

- Khan, U.S., and Khan, S.U.R. (2024). Boost diagnostic performance in retinal disease classification utilizing deep ensemble classifiers based on OCT. *Multimedia Tools and Applications*. doi: 10.1007/s11042-024-19922-1.
- Kharousi, N., Wali, U.K., and Azeem, S. (2013). Current Applications of Optical Coherence Tomography in Ophthalmology. In *Optical Coherence Tomography*. IntechOpen.
- Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60, 84–90. doi: 10.1145/3065386.
- Kulyabin, M., Zhdanov, A., Nikiforova, A., et al. (2024). OCTDL: Optical Coherence Tomography Dataset for Image-Based Deep Learning Methods. *Scientific Data*, 11, 365. doi: 10.1038/s41597-024-03182-7.
- Larochkin, P., Kotina, E., Girdyuk, D., and Ostroumov, E. (2025). Deep learning methods for right ventricle segmentation in radionuclide imaging. In *Cybernetics and Physics*, 14(1), 62–66. doi: 10.35470/2226-4116-2024-14-1-62-66.
- Mathews, M.R., and Anzar, S.M. (2022). Residual Networks and Deep-Densely Connected Networks for the Classification of Retinal OCT Images. In *2022 International Conference on Connected Systems & Intelligence (CSI)*, Trivandrum, India, pp. 1–7. doi: 10.1109/CSI54720.2022.9923993.
- Maurya, R., Pandey, N.N., Joshi, R.C., and Dutta, M.K. (2024). MacD-Net: An automatic guided-ensemble approach for macular pathology detection using optical coherence tomography images. *International Journal of Imaging Systems and Technology*, 34, e22954. doi: 10.1002/ima.22954.
- Neroev, V.V. (2014). Organization of ophthalmological care for the population of the Russian Federation. *Vestnik Oftalmologii*, 6, 8–12.
- Peng, J., et al. (2023). Multi-Scale-Denoising Residual Convolutional Network for Retinal Disease Classification Using OCT. *Sensors*, 24(1), 150. Multidisciplinary Digital Publishing Institute.
- Perdomo, O., et al. (2018). OCT-NET: A convolutional network for automatic classification of normal and diabetic macular edema using SD-OCT volumes. In *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, pp. 1423–1426. doi: 10.1109/ISBI.2018.8363839.
- Plotnikov, S. (2024). Adaptive observer for parameter estimation in neural mass models using EEG recordings of evoked potentials. *Cybernetics and Physics*, 13(4), 296–301. doi: 10.35470/2226-4116-2024-13-4-296-301.
- Rong, W., Li, Z., Zhang, W., and Sun, L. (2014). An improved Canny edge detection algorithm. In *Proceedings of the 2014 IEEE International Conference on Mechatronics and Automation (ICMA)*. IEEE, pp. 577–582. doi: 10.1109/ICMA.2014.6885761.
- Saini, D.J.B., Sivakami, R., Venkatesh, R., Raghava, C.S., Dwarkanath, P.S., Anwer, T.M.K., et al. (2023). Convolution neural network model for predicting various lesion-based diseases in diabetic macula edema in optical coherence tomography images. *Biomedical Signal Processing and Control*, 86, 105180. doi: 10.1016/j.bspc.2023.105180.
- Thapa, R., Khanal, S., Tan, H.S., Thapa, S.S., and van Rens, G.H.M.B. (2020). Prevalence, pattern and risk factors of retinal diseases among an elderly population in Nepal: The Bhaktapur Retina Study. *Clinical Ophthalmology*, 14, 2109–2118. doi: 10.2147/OPTH.S262131
- Turbert, D. (2023). What Is Optical Coherence Tomography? American Academy of Ophthalmology. Published April 27, 2023. Available from: <https://www.aao.org/eye-health/treatments/what-is-optical-coherence-tomography>